

Analiza struktury przestrzennej  
i czasoprzestrzennej  
maksymalnych opadów dobowych  
w Polsce  
w latach 1956-1980



UNIwersytet IM. ADAMA MICKIEWICZA W POZNANIU  
SERIA GEOGRAFIA NR 85

ALFRED STACH

**Analiza struktury przestrzennej  
i czasoprzestrzennej  
maksymalnych opadów dobowych  
w Polsce  
w latach 1956-1980**



POZNAŃ 2009

ABSTRACT. Stach Alfred, *Analiza struktury przestrzennej i czasoprzestrzennej maksymalnych opadów dobowych w Polsce w latach 1956-1980* [Analysis of the spatial and spatial-temporal structure of maximum daily precipitation in Poland in the years 1956-1980]. Adam Mickiewicz University Press, Poznań 2009. Seria Geografia nr 85. Pp. 323. Figs 147. Tabs 30. ISBN 978-83-232-1988-0. ISSN 0554-8128. Polish text with a summary in English.

The main goal of the work was to analyse regularities underlying spatial and temporal variations in monthly and annual maximum daily precipitation totals (MDPTs) in the years 1956-1980 from the area of Poland. The MDPT field in Poland is usually the sum of the outcomes of three types of processes operating at various spatial scales: local (<10-20 km), regional (50-150 km), and supra-regional (>200 km). The parameters characterising the spatial structure of MDPTs display a marked seasonal cyclicity and variability depending on the precipitation level. In multiannual period, the changes found to occur in the elements of the spatial structure, show that the hypothesis - as to a potential effect of changes in the spatial structure of extreme precipitation events on their point-registered frequency - is a promising research perspective and should continue to be tested. Based on variation in the level and gradient of the probability of occurrence of annual MDPTs in various rainfall-amount classes seasonal subdivision was done. Dates of occurrence of annual MDPTs differ markedly regionally. Many of the newly found regularities need a deeper genetic interpretation, which is hard, if not impossible, to conduct without additional data and a wider context of their analysis.

Alfred Stach - Uniwersytet im. Adama Mickiewicza w Poznaniu, Wydział Nauk Geograficznych i Geologicznych, Instytut Paleogeografii i Geoekologii, Zakład Geoekologii, ul. Dziegiełkowa 27, 61-680 Poznań, Collegium Geographicum; e mail: frdstach@amu.edu.pl

Recenzent: prof. dr hab. Andrzej Kędziora

© Copyright by Alfred Stach 2009

© Copyright for this edition by Wydawnictwo Naukowe UAM, Poznań 2009

Redaktor: Aleksandra Jędrzejczak

Redaktor techniczny: Dorota Borowiak

Łamanie komputerowe: Bartosz Dziamski

ISBN 978-83-232-1988-0

ISSN 0554-8128

WYDAWNICTWO NAUKOWE UNIwersytetu IM. ADAMA MICKIEWICZA W POZNANIU

61-734 Poznań, ul. F. Nowowiejskiego 55, tel. 061 829 39 85, fax 061 829 39 80

e-mail: [press@amu.edu.pl](mailto:press@amu.edu.pl)

[www.press.amu.edu.pl](http://www.press.amu.edu.pl)

Wydanie I. Ark. wyd. 25,00. Ark. druk. 20,25 + płyta DVD

Druk i oprawa: PPHU Totem s.c., Inowrocław, ul. Jacewska 89



# Spis treści

---

Podziękowania .....	9
<b>I. Wprowadzenie</b> .....	11
<b>II. Zarys problematyki</b> .....	15
<b>III. Cel, zakres i metody pracy</b> .....	22
1. Cel i zakres pracy .....	22
2. Podstawy metodyki .....	24
2.1. Wprowadzenie .....	24
2.2. Miary ciągłości lub zmienności przestrzennej .....	24
2.3. Struktura przestrzenna w klasach natężenia analizowanej cechy .....	30
2.4. Przykład analizy danych kodowanych .....	32
2.5. Modelowanie struktury przestrzennej danych .....	38
2.5.1. Wprowadzenie do problematyki modelowania struktury przestrzennej .....	38
2.5.2. Proste i złożone, dopuszczalne funkcje używane przy modelowaniu struktury przestrzennej .....	40
2.5.3. Specyfika matematycznego modelowania struktury przestrzennej ..	44
2.5.4. Kryterium jakości dopasowania modelu .....	45
2.5.5. Optymalizacja modelu struktury przestrzennej .....	46
2.5.6. Modelowanie struktury przestrzennej MSDO w programie ISA-TIS .....	49
2.5.7. Modelowanie struktury przestrzennej danych znormalizowanych ..	49
2.5.8. Modelowanie struktury przestrzennej danych kodowanych .....	52
2.5.9. Nieciągłość i asynchroniczność danych MSDO a ich struktura przestrzenna .....	57
<b>IV. Charakterystyka pluwiotermiczna lat 1956-1980</b> .....	59
<b>V. Dane pomiarowe maksymalnych sum opadów dobowych</b> .....	63
1. Źródła i charakter danych .....	63
2. Błędy i niepewność danych .....	64
3. Zmienność liczby stanowisk pomiarowych .....	67
4. Rozmieszczenie przestrzenne stanowisk pomiarowych .....	70
5. Maksymalne sumy dobowe opadów a źródłowe serie pomiarowe. Teoretyczne dystrybuanty danych .....	83
6. Charakterystyka danych źródłowych za pomocą globalnych statystyk opisowych .....	86

7. Charakterystyka danych źródłowych za pomocą lokalnych statystyk opisowych .....	96
8. Relacje czasowe .....	103
9. Podsumowanie rozdziału .....	107
<b>VI. Analiza struktury przestrzennej MSDO opartej na semiwariogramach danych znormalizowanych .....</b>	<b>111</b>
1. Wprowadzenie .....	111
2. Częstość występowania i charakterystyka elementarnych modeli .....	112
3. Sezonowa i wieloletnia zmienność struktury przestrzennej MSDO .....	115
4. Klasyfikacja struktury przestrzennej MSDO – liczba i charakter składowych ..	122
5. Charakterystyka składowych złożonych modeli struktury przestrzennej MSDO .....	123
6. Klasyfikacja struktury przestrzennej MSDO – wartość semiwariancji .....	135
7. Sezonowa zmienność parametrów modeli struktury przestrzennej .....	150
8. Wieloletnia zmienność parametrów modeli struktury przestrzennej .....	157
9. Zmienność bezwzględnych wartości wariancji składowych przestrzennych ...	161
10. Podsumowanie rozdziału .....	169
<b>VII. Analiza struktury przestrzennej MSDO na podstawie semiwariogramów danych kodowanych .....</b>	<b>173</b>
1. Wprowadzenie .....	173
2. Częstość elementarnych modeli i czynniki ją kształtujące .....	176
3. Zróznicowanie sezonowe .....	178
4. Zróznicowanie zależne od względnej wysokości opadu .....	182
5. Zróznicowanie zależne od bezwzględnej wysokości opadu – miesiące .....	187
6. Sezonowe zróznicowanie zależne od bezwzględnej wysokości opadu .....	195
7. Zróznicowanie zależne od bezwzględnej wysokości opadu – lata .....	201
8. Zmienność wieloletnia .....	205
9. Podsumowanie rozdziału .....	206
<b>VIII. Sezonowość, wydajność i zmienność regionalna rocznych MSDO .....</b>	<b>211</b>
1. Zmienność czasowa występowania i wydajności rocznych MSDO .....	211
2. Estymacja gęstości skupień .....	213
3. Sezonowa zmienność wydajności rocznych MSDO na terenie Polski .....	217
4. Zmienność przestrzenna terminu występowania rocznych MSDO .....	220
5. Testowanie istotności różnic regionalnych terminu występowania rocznych MSDO .....	227
6. Autokorelacja przestrzenna terminów rocznych MSDO .....	231
7. Podsumowanie rozdziału .....	234
<b>IX. Podsumowanie i wnioski .....</b>	<b>236</b>
<b>X. Dodatek: Szczegółowe zagadnienia metodyczne .....</b>	<b>240</b>
1. Problem zmian ilości i rozkładu przestrzennego danych .....	240
2. Estymacja, symulacja i optymalizacja .....	243
3. Anizotropia struktury przestrzennej MSDO .....	254
4. Normalizacja danych .....	261

4.1. Ogólna charakterystyka procedur normalizacji danych wykorzystywanych w geostatystyce .....	261
4.2. Normalizacja parametryczna z użyciem wielomianów Hermite .....	263
4.3. Przykład normalizacji danych MSDO .....	267
4.4. Wpływ normalizacji danych na ocenę ich struktury przestrzennej .....	269
5. Maskowanie danych w trakcie analizy struktury przestrzennej .....	271
5.1. Wpływ danych ekstremalnych i odstających na określenie struktury przestrzennej zjawiska .....	271
5.2. Metodyka procedury interaktywnego czyszczenia semiwariogramu .....	272
5.3. Statystyki procedury czyszczenia semiwariogramu dla wszystkich analizowanych zbiorów danych MSDO .....	276
6. Rozdzielczość modeli przestrzennych .....	280
6.1. Ogólne kryteria wyboru rozdzielczości rastrowych modeli przestrzennych .....	280
6.2. Ocena optymalnej rozdzielczości modeli z wykorzystaniem charakterystyk analizowanego zbioru danych .....	281
<b>XI. Literatura .....</b>	<b>285</b>
<b>XII. Załączniki .....</b>	<b>295</b>
1. Zestawienie stacji meteorologicznych i posterunków opadowych zlokalizowanych na terytorium Niemiec, z których dane sum dobowych opadów zostały wykorzystane w niniejszej pracy .....	295
2. Opis zawartości załączonego dysku DVD .....	298
2.1. Wstęp .....	298
2.2. Używane formaty plików .....	298
2.3. Spis zawartości dysku .....	301
2.3.1. Dane źródłowe .....	301
2.3.2. Numeryczne i graficzne wyniki analiz .....	302
2.3.3. Wersje elektroniczne kolorowych rycin i tabel .....	303
2.3.4. Pliki instalacyjne programów komputerowych .....	303
<b>XIII. List of figures and tables .....</b>	<b>305</b>
<b>Analysis of the spatial and spatial-temporal structure of maximum daily precipitation in Poland in the years 1956-1980 (S u m m a r y) .....</b>	<b>320</b>



# Podziękowania

---

Tak duże przedsięwzięcie, jakim było przygotowanie niniejszej rozprawy, z pewnością by się nie powiodło, gdyby nie życzliwa atmosfera i pomoc wielu osób. Miłym obowiązkiem autora jest więc złożenie wszystkim najserdeczniejszych podziękowań.

Od początku swojej pracy zawodowej jestem członkiem zespołu kierowanego przez prof. dr. hab. Andrzeja Kostrzewskiego (obecnie Zakład Geokologii UAM). Stworzone przez Pana Profesora doskonałe warunki pracy i twórcza atmosfera dyskusji naukowych stanowiły fundament, na którym wszystko zostało zbudowane.

W problematykę zmienności przestrzennej opadów atmosferycznych wprowadził mnie prof. dr. hab. Jan Tamulewicz. Naszą harmonijną współpracę przerwała choroba i śmierć Janka. To, co jest tematem niniejszej rozprawy planowaliśmy zrobić razem. Wszystkie jej niedoskonałości to skutek tej wielkiej straty.

Czasochłonne, i wymagające także znaczących środków finansowych, prace nad analizą ekstremalnych opadów nie byłyby pewnie możliwe w tak szerokim zakresie i tak szybko, gdyby nie mój udział w projekcie badawczym koordynowanym przez prof. dr. hab. Jacka Janię z Uniwersytetu Śląskiego. Inspirujące i pouczające dyskusje w trakcie roboczych spotkań członków zespołu realizującego projekt miały duży wpływ na końcowy rezultat mojej pracy.

Pierwszą, bardzo jeszcze niedoskonałą jej wersję przeczytał cierpliwie i wnikliwie prof. UAM dr. hab. Waldemar Ratajczak. Jego uwagi, w pierwszej chwili niekiedy irytujące, najczęściej po dłuższym zastanowieniu okazywały się jednak cenne i trafne. Dzięki temu mogłem wyeliminować znaczną część braków i niejasności.

Związała i niezwykle konkretna recenzja oraz dyskusje prowadzone z prof. dr. hab. Andrzejem Kędziorą stanowiły ważny impuls do przeprowadzenia wielu zmian i uzupełnień. Między innymi, pod wpływem sugestii Pana Profesora zmieniłem układ pracy na bardziej, mam nadzieję, przejrzysty.

Pani mgr Anna Dmowska poświęciła znaczną część wakacyjnego czasu na przeprowadzenie długotrwałych obliczeń w programie IKSIM i żmudne ich przygotowanie.

Wiele osób pomagało mi na etapie tworzenia i weryfikowania bazy danych. Niniejsza praca niezwykle również zyskała dzięki licznym dyskusjom prowadzonym z koleżankami i kolegami z Zakładu oraz Instytutu. Nie wymieniam ich wszystkich po imieniu i nazwisku, ale zawsze będę pamiętał ich pomoc oraz życzliwe zainteresowanie.


Nie byłoby jednak ani tego pomysłu, ani jego realizacji, gdyby nie codzienne wsparcie i motywacja w chwilach kryzysu udzielana mi zawsze przez Najbliższych.

---

## Wprowadzenie\*

Jednym z bardziej niepokojących wyników prognoz wykonywanych przy zastosowaniu globalnych (ang. GCM) i regionalnych modeli klimatycznych (Bell i in. 2004, Christensen, Christensen 2004, Fowler i in. 2005, Frei i in. 1998, Hennessy i in. 1997, Kharin i in. 2007, Palmer, Räisänen 2002, Semmler, Jacob 2004, Wehner 2004) jest wysokie prawdopodobieństwo wzrostu częstości i natężenia ekstremalnych opadów. Koresponduje to z wieloma opublikowanymi analizami archiwalnych serii pomiarowych wskazującymi na pojawianie się w drugiej połowie XX wieku takich tendencji (Dore 2005, Groisman i in. 2005, Kiktev i in. 2003, Klein Tank, Können 2003, Koning, Franses 2005, Kundzewicz i in. 2005, Kunkel 2003, Kunkel i in. 1999, New i in. 2001, Trömel, Schönwiese 2007, Zolina i in. 2005). Poznanie prawidłowości powtarzalności i zasięgu przestrzennego ekstremalnych opadów ma oczywiście bardzo duże znaczenie zarówno gospodarcze, jak i społeczne (Douglas, Barros 2003, Durran, Kirby 2004, Faulkner, Prudhomme 1998, Frei i in. 2000, Kothavala 1997, Kundzewicz 2005, Rezacova i in. 2005, Ulbrich i in. 2003). Pełne możliwości badania tych problemów umożliwia jedynie łączenie punktowych pomiarów naziemnych<sup>1</sup> z obszarowymi danymi teledetekcyjnymi<sup>2</sup> (Haberlandt 2007, Nesbitt i in. 2004, New i in. 2001). Niestety, stosunkowo krótki okres dostępności danych teledetekcyjnych oraz trudności w dokonywaniu na ich podstawie precyzyjnej oceny intensywności opadów utrudnia formułowanie uogólnień. Dlatego, niezależnie od wprowadzania nowych technik pomiarowych należy dokonywać analizy i reinterpretacji archiwalnych danych, korzystając z możliwości stwarzanych przez rozwój GIS (Dobesch i in. 2001, Tveito i in. 2006).

---

\* Ponad 70 ze 147 rycin zamieszczonych w niniejszej publikacji zostało przygotowane w oryginale w postaci kolorowej. Ponieważ zostały one wydrukowane w odcieniach szarości utrudnia to, a w niektórych przypadkach wręcz uniemożliwia pełną interpretację ich treści. Zamieszczenie elektronicznej wersji kolorowych rycin na załączonym DVD ma na celu częściowe zrekompensowanie tej niedogodności. Ryciny, których kolorowe kopie są na DVD, oznaczono w tekście symbolem dyskietki .

<sup>1</sup> Rejestracje pluwiograficzne i pomiary za pomocą deszczomierzy: ilościowe, bezpośrednie, ciągłe i/lub okresowe.

<sup>2</sup> Radary meteorologiczne i obrazowanie satelitarne: pośrednie pomiary ilościowe.

Dotychczasowe analizy maksymalnych opadów dobowych, występujących na obszarze Polski, z punktu widzenia współczesnych możliwości i potrzeb, mają istotne słabości (Bogdanowicz, Stachý 1998, Fal, Plenzler 1981, Kłysik, Fortuniak 1993, Ustrnul, Czekierda 2000). Po pierwsze, wykonywano je na podstawie niewielkiej części dostępnych zbiorów danych. Przyczyną tego stanu rzeczy było z jednej strony traktowanie bardzo rygorystycznie warunków kompletności danych i równomierności pokrycia, a także ówczesne trudności „techniczne” przetwarzania dużych ich zbiorów. Po drugie, wykorzystywany aparat matematyczny skoncentrowany był na analizie czasowej, najczęściej powtarzalności opadów z określonych klas wysokości. Tworzone w końcowym etapie owych opracowań mapy dawały w związku z tym bardzo zgeneralizowany obraz zmienności przestrzennej.

Większość modeli stosowanych do symulacji globalnego systemu klimatycznego i prognozowania skutków antropogenicznych zmian bilansu promieniowania jest wykorzystywana głównie do uzyskiwania stosunkowo prostych wskaźników klimatycznych, głównie o charakterze czasowym, jak na przykład: średnie temperatury roczne czy sezonowe, sumy roczne lub sezonowe opadów, przeciętna wydajność opadów itp. Symulowane wartości tych wskaźników wykazują już bardzo dużą zgodność z danymi pomiarowymi, co zwiększa zaufanie do długookresowych prognoz klimatycznych. Często jednakże, kiedy takiemu porównaniu poddawane są bardziej skomplikowane parametry klimatyczne, między wynikami symulacji a rzeczywistością istnieją duże rozbieżności. Trenberth i in. (2003) jako główną myśl swojego artykułu umieścili następujące zdanie: *W efekcie zmian klimatycznych w reżimie opadów najprawdopodobniej zmianie ulegnie intensywność, frekwencja i czas trwania epizodów opadowych, lecz te charakterystyki są bardzo rzadko analizowane zarówno na podstawie danych pomiarowych, jak i w modelach*<sup>3</sup>. Tezę tę zilustrowali następnie kilkoma przykładami. Należałoby ją uzupełnić jednak o stwierdzenie, że w wypadku nieciągłych elementów meteorologicznych – jakimi są opady – zmiany dotyczyć mogą zarówno ich charakterystyk czasowych, jak i przestrzennych. Zmiany rejestrowanej punktowo frekwencji opadów w różnych klasach intensywności, czy czasu trwania, mogą bowiem wiązać się zarówno z rzeczywistą zmianą częstości ich występowania, jak i zmianami ich zasięgu przestrzennego. Tak więc, obserwowany wzrost „ilości” ekstremalnych deszczy może się wiązać jedynie z ich większym zasięgiem przestrzennym, bez żadnych zmian reguł ich powtarzalności<sup>4</sup>. Omówiony powyżej problem wymaga większej uwagi niż mu do tej pory poświę-

---

<sup>3</sup> „As climate changes, the main changes in precipitation will likely be in the intensity, frequency, and duration of events, but these characteristics are seldom analyzed in observations or models”. Trenberth i in. 2003: s. 1.

<sup>4</sup> Można sobie wyobrazić nawet taką sytuację, że rejestrowana punktowo frekwencja opadów z danej klasy intensywności/czasu trwania może wzrosnąć, nawet jeśli częstość procesów je generujących zmaleje. Większe znaczenie bowiem będzie miał wzrost ich zasięgu przestrzennego.



cano. Analizy zmian charakterystyk klimatu, które do tej pory koncentrowały się prawie wyłącznie na ich aspekcie czasowym, powinny także uwzględniać potencjalną ewolucję struktury przestrzennej. Konieczne jest zatem opracowanie na podstawie archiwalnych danych wieloletnich podstawowych charakterystyk tej struktury, które mogłyby być punktem odniesienia dla testowania możliwości istnienia ewentualnych trendów ich zmian.

Celem niniejszej rozprawy jest przeprowadzenie wszechstronnej analizy struktury przestrzennej wyselekcjonowanego zbioru wysokich opadów, którymi są publikowane w rocznikach opadowych PIHM i IMGW z lat 1954-1981 maksymalne sumy dobowe (określane w dalszej części niniejszej pracy jako **MSDO**), rejestrowane dla poszczególnych miesięcy w posterunkach opadowych i stacjach meteorologicznych<sup>5</sup>. Pojęcie „struktura przestrzenna” jest rozumiane w sensie zdefiniowanym na gruncie geostatystyki (Chilès, Delfiner 1999, Cressie 1993, Deutsch, Journel 1998, Goovaerts 1997, Olea 1999, Wackernagel 2003, Webster, Oliver 2001), jako prawidłowości zmian zasięgu i siły autokorelacji przestrzennej analizowanej cechy.

Podjęty temat ma również aspekt praktyczny. Rejestracje opadów wykonywane za pomocą radarów meteorologicznych lub satelitów mają charakter pośredni. Aby można je wykorzystać do oceny rzeczywistej dostawy wody do powierzchni ziemi muszą zostać skalibrowane. Wykonuje się to bardzo często z użyciem metod geostatystycznych, na podstawie synchronicznych z pomiarami teledetekcyjnymi, punktowych pomiarów naziemnych (Grimes i in. 1999, Haberlandt 2007, New i in. 2001, Sun i in. 2003). Takie procedury są między innymi testowane również w Polsce (Jurczyk i in. 2004, 2007). Podstawą każdej geostatystycznej metody estymacji i symulacji jest określenie w postaci parametrycznej modelu struktury przestrzennej analizowanej cechy. Stworzenie bazy danych takich modeli oraz wykazanie prawidłowości ich powtarzalności może mieć w tym kontekście duże znaczenie.

Autor niniejszej pracy od blisko dziesięciu lat specjalizuje się w aplikacjach metod analizy przestrzennej do badania i rozwiązywania problemów z zakresu nauk o Ziemi (Górska-Zabielska, Stach 2008, Stach 1998, 1999, 2002a, 2002b, 2002c, 2005, 2006a, 2007c, Stach, Podsiadłowski 2000, 2001, 2002, Stach, Tamulewicz 2003, Stach i in. 2003a, 2003b). Z tego właśnie względu, choć głównym motywem podjęcia tematu była chęć zmierzenia się z konkretnym zagadnieniem klimatologicznym, aspekt metodyczny jest traktowany jako równie ważny. Przedstawiono bowiem w tym zakresie szereg nowych, nie tylko na gruncie polskim, propozycji analizy przestrzennej danych ekstremalnych opadów. Autor skupił się bardziej na szerokim, wszechstronnym omówieniu rezultatów wykonanych obliczeń, zachowując dużą

---

<sup>5</sup> Pełne omówienie celu rozprawy znajduje się w podrozdziale III.1, a szczegółową charakterystykę analizowanych danych zamieszczono w rozdziale V.

ostrożność w formułowaniu daleko idących uogólnień. Powściągliwość w interpretacji wynika także z innej przyczyny. Mimo bardzo wnikliwej kwerendy literatury tematu nie udało się znaleźć znaczących publikacji, które mogłyby stanowić punkt odniesienia do wyników uzyskanych w ramach niniejszej pracy. Świadczy to albo o oryginalności niniejszej rozprawy, albo o braku zainteresowania badaczy marginalnym zagadnieniem. Problem braku wypracowanych wzorców interpretacji był jednak dla autora z jednej strony ekscytującym źródłem motywacji do intensywniejszej pracy intelektualnej, ale z drugiej także powodem wielu frustracji. Dołączony do pracy materiał faktograficzny powinien jednakże umożliwić zainteresowanym badaczom pogłębienie i poszerzenie problematyki w tych zakresach, gdzie być może została ona poruszona zbyt powierzchownie. Autor ma nadzieję, że niniejsze opracowanie okaże się także pożyteczne z punktu widzenia praktyki w takich dziedzinach, jak gospodarka wodna, ochrona środowiska, rolnictwo czy planowanie przestrzenne.

Rozprawę wykonano w trakcie realizacji zamawianego projektu badawczego PBZ-KBN-086/P04/2003: „Ekstremalne zdarzenia meteorologiczne i hydrologiczne w Polsce (ocena zdarzeń oraz prognozowanie ich skutków dla środowiska życia człowieka)”, kierowanego przez prof. dr. hab. Jacka Janię z Uniwersytetu Śląskiego. Zaprezentowana w niej problematyka jest częścią realizowanego przez autora znacznie szerszego tematu badawczego. Obejmował on obok analizy i modelowania struktury przestrzennej maksymalnych sum opadów dobowych także estymację i symulację pola prawdopodobieństwa MSDO. Uzyskane w tym zakresie wyniki miałyby, oprócz dużych wartości dla różnych dziedzin praktyki, istotne znaczenie dla lepszego zrozumienia mechanizmów generujących MSDO w różnych skalach czasoprzestrzennych oraz dla regionalizacji klimatycznej kraju. Autor rozwinął tę problematykę także w kierunku szczególnie go interesującym – probabilistycznej oceny zmienności na terytorium Polski potencjału erozyjnego ekstremalnych opadów. W trakcie pracy stwierdzono jednak wkrótce, że jej zakres tematyczny, a przede wszystkim „objętość”, przekracza znacznie ramy typowej rozprawy habilitacyjnej. Zdecydowano zatem wydzielić tę jej część, która ma większy ładunek metodyczny i poznawczy, nie rezygnując, oczywiście, z realizacji całości planu. Część wstępnych wyników została zresztą zaprezentowana już na konferencjach naukowych i opublikowana (Stach 2006b, 2006c, 2006d, 2006e, Stach 2007a, 2007b). Zamieszczone powyżej wyjaśnienie ma istotne znaczenie, ponieważ szereg ważnych decyzji metodologiczno-metodycznych podjęto z uwagi na cały zaplanowany zakres tematu, a nie tylko zawartą w niniejszej rozprawie jego część. Z tego względu czytelnik powinien poznać w ogólnym zarysie owo inicjalne założenie w celu zorientowania się, jakie możliwości ono otwiera, a jednocześnie jakie ograniczenia wprowadziło do sposobu opracowania głównego problemu niniejszej rozprawy.

## II

---

### Zarys problematyki

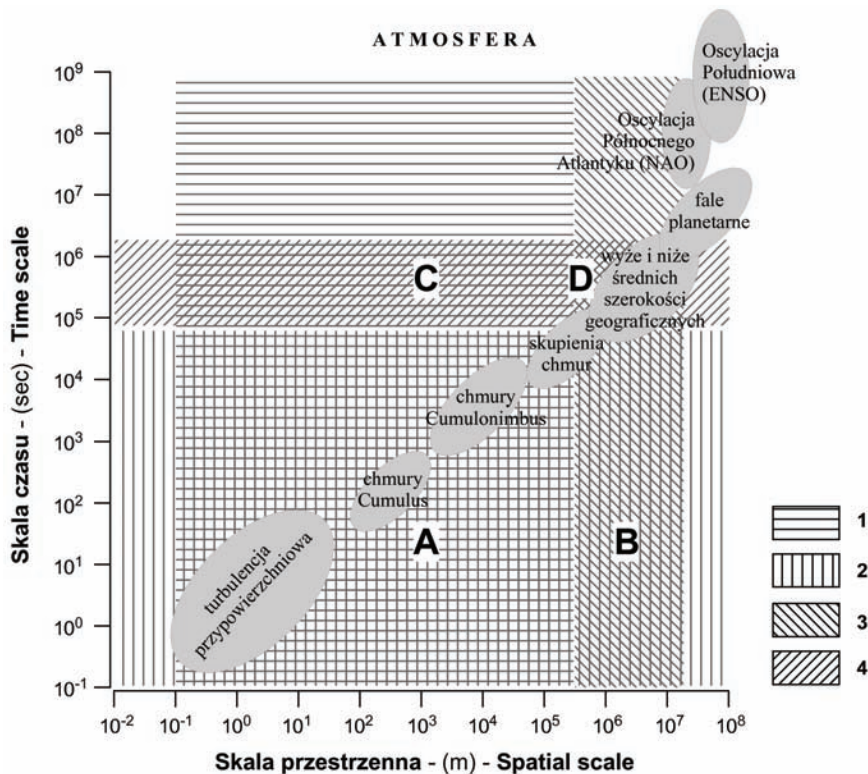
W średnich szerokościach geograficznych, w odróżnieniu od stref tropikalnych, zaburzenia synoptyczne wraz z zachodzącymi w ich efekcie ekstremalnymi opadami, mają swoją genezę głównie w strefach charakteryzujących się silnymi południkowymi gradientami temperatury i naprężeń baroklinicznych ponad oceanami (Hense, Friederichs 2006).

Rycina 1 prezentuje zestawienie wybranych typowych zjawisk atmosferycznych w układzie skal przestrzeni i czasu. Są to procesy turbulencyjne, chaotyczne charakteryzujące się możliwością generowania ekstremalnych wartości. Turbulencyjny charakter wynika z nieliniowych interakcji charakterystyk przepływów (temperatury, prędkości, masy), jak też mających postać procesów progowych kondensacji i parowania wody. Ekstrema zachodzące w różnych skalach zazwyczaj łączą się, na przykład:

- ekstremalne duże pionowe gradienty entropii prowadzą do niestabilności konwekcyjnej w skali przestrzennej  $\approx 100$  km. Będąca ich skutkiem silna konwekcja może wywołać ekstremalne prędkości pionowe powietrza o skali  $\approx 1$  km. W obrębie tych struktur cyrkulacji pionowej mogą powstawać bardzo intensywne opady (także gradu), które w zależności od ich natężenia i wydajności mogą powodować poważne zniszczenia;

- olbrzymie poziome gradienty temperatur mogą w średnich szerokościach geograficznych powodować niestabilności barokliniczne w skalach przestrzennych  $\approx 1000$  km. Będące ich efektem cyklony - turbulencje atmosferyczne - powodują powstawanie ekstremalnych prędkości horyzontalnych powietrza wzdłuż ich stref frontalnych. Tarcie przypowierzchniowe wywołuje tam silny pionowy przepływ ścinający, powodujący dynamicznie indukowaną turbulencję z ekstremalnymi porywami wiatru lub wichurami.

W przeciwieństwie do wiatru, ekstrema opadów wykazują regularne relacje czasoprzestrzenne (prawa skali), które są prawdopodobnie efektem bardziej ogólnych prawidłowości zmienności czasoprzestrzennej opadów.



**Ryc. 1.** Domeny czasu i przestrzeni analizy MSDO w odniesieniu do ważniejszych zjawisk atmosferycznych (Hense, Friederichs 2006, zmienione): 1 – bezpośredni zakres przestrzenny danych MSDO, 2 – bezpośredni zakres czasowy danych MSDO, 3 – pośredni zakres przestrzenny danych MSDO, 4 – pośredni zakres czasowy danych MSDO. Domeny określone za pomocą „A”, „B”, „C” i „D” są efektem kombinacji wyżej wymienionych zakresów. Szczegółowe objaśnienia w tekście

Rycina 1 umożliwi również ocenę reprezentatywności, wykorzystanych w niniejszej rozprawie, danych MSDO w odniesieniu do skal czasowych i przestrzennych procesów atmosferycznych powiązanych z powstawaniem opadów. W deszczomierzach o powierzchni zbiorczej  $0,02 \text{ m}^2$ , usytuowanych na powierzchni o rozciągłości około  $6,0 \times 10^5 \text{ m}$ , skumulowane były opady z okresu doby ( $0 \div 8,64 \times 10^4 \text{ s}$ ). Te wartości graniczne przestrzeni i czasu określają zakres bezpośredniej reprezentatywności danych MSDO (powierzchnia określona literą A na rycinie 1). Obejmuje ona całkowicie zasięgi procesów turbulencji przypowierzchniowej, formowania chmur *Cumulus* i *Cumulonimbus* oraz częściowo procesy mezoskalowe (tworzenie skupień komórek konwekcyjnych). Trzeba jednakże pamiętać (patrz rozdz. V), że rze-

czywista rozdzielczość przestrzenna danych pluwiometrycznych jest związana z ich zagęszczeniem na analizowanym obszarze i wynosi od 1 do 3 km. Pomiary MSDO rejestrowały jednak również efekty procesów rozgrywających się w skalach czasoprzestrzennych przekraczających jedną dobę i rozciągłość terytorium Polski. Zmierzona maksymalna suma dobowa opadu może stanowić bowiem część epizodu opadowego trwającego wiele dni i obejmującego swoim zasięgiem nawet cały sektor atlantycko-europejski półkuli północnej (zakresy B, C i D na ryc. 1). Uwzględnić wówczas należy możliwość oddziaływań dużych struktur cyrkulacji atmosferycznej: układów niżowych i wyżowych, a nawet fal planetarnych.

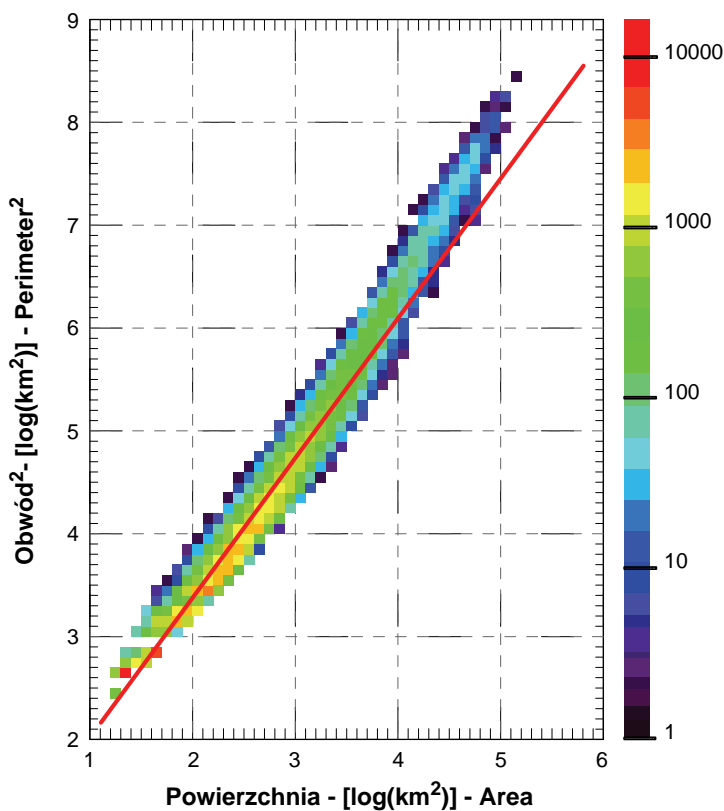
O czasoprzestrzennej zmienności reżimu opadowego Polski decydują zasadniczo procesy synoptyczne zachodzące nad europejskim sektorem umiarkowanych szerokości geograficznych (Stach, Tamulewicz 2005a). Natura opadogenezy, dynamika jej przebiegu oraz efektywność mierzona sumami opadu zależą bowiem od rodzaju układów barycznych, właściwości mas powietrznych oraz procesów na frontach atmosferycznych. Dynamiczne cechy klimatu Polski są zazwyczaj kształtowane przez peryferyjne części niżów i wyżów; trajektorie poruszania się ich centrów znajdują się poza naszymi granicami. Położenie geograficzne układów barycznych, a ściślej ich przestrzenna orientacja względem obszaru Polski nadaje kierunek ruchom powietrza w naszym regionie, zaś wartości poziomego gradientu ciśnienia wpływają na prędkość przemieszczania się mas powietrznych. Z kierunkiem ich napływu związany jest transport pary wodnej oraz ciepła. Intensywność opadów, czas ich trwania wiążą się natomiast z charakterem stref kontaktu różnych mas powietrznych, czyli właściwościami stref frontowych.


Opady w atmosferze są efektem końcowym długiej kaskady oddziaływań zachodzących w coraz to większej skali (Hense, Friederichs 2006). Kaskada rozpoczyna się od interakcji pojedynczych cząsteczek wody, kontynuowana jest w skali kropelek w chmurach i kropel opadu, a później przechodzi na poziom małoskalowych turbulencji, chmur i zróżnicowania mezoskalowych przepływów stanowiących element cyrkulacji w skalach synoptycznych i planetarnych. To co dzieje się na każdym z wymienionych poziomów decyduje, czy w danym momencie czasu i punkcie przestrzeni powstaje opad. Chmury, a więc także opady, mają w związku z tym bardzo złożoną i bogatą strukturę czasoprzestrzenną. Część z niej jest efektem przyjętego sposobu charakteryzowania opadów, i wynikającego z tego faktu, że nie można w tym przypadku mówić o ekstremach rozumianych dosłownie. Są one bowiem odnoszone do arbitralnie przyjętego interwału czasu, w obrębie którego dokonywana jest akumulacja opadu. Z fizycznego punktu widzenia opad to pionowy transport masy ciekłej lub zamrożonej wody (przy

założeniu, że ignorowana jest jego nieciągła natura: krople deszczu, grudki gradu, czy płatki śniegu). W standardowym systemie pomiarowym możliwa jest jedynie rejestracja masy skumulowanej w określonym przedziale czasu i na określonej powierzchni. Zarówno ów przedział czasu, jak i powierzchnia muszą być na tyle duże w porównaniu z rozmiarem kropli, aby zebrana próbka opadu umożliwiła dokładne określenie natężenia owego transportu masy. Typowe minimalne wartości wynoszą około jednej minuty i kilkaset centymetrów kwadratowych. Światowa Organizacja Meteorologiczna (WMO 1994 za: Hense, Friederichs 2006) przygotowała raport zawierający zestawienie ekstremów opadowych ( $R_T$ ) z szerokiego zakresu czasów ich agregacji  $T$  (od 1 minuty do 2 lat:  $6,0 \times 10^1 \div 6,3 \times 10^7$  s). Wahają się one od 38 do 40 768 mm. Relacja między obiema zmiennymi ma charakter potęgowy ( $R_T = 6,5T^{0,49}$ ) i jest bardzo silna ( $R = 99,7\%$ ). W pracy Lovejoya i Schertzer (1995) zostało wykazane, że przedstawiona wyżej zależność jest powiązana z bardziej ogólną relacją skali opadów względem czasu (podobne prawo skali obowiązuje w odniesieniu do przestrzeni). Można ją teoretycznie uzasadnić za pomocą losowego modelu układu kaskadowego. Jasne jest w kontekście tych danych, że ekstremalność opadów można definiować jedynie w odniesieniu do ustalonego czasu ich akumulacji, a także w dużym stopniu – geograficznie. Suma 10 mm opadu w lipcu na terenie Polski stanowi skrajnie wysoką wartość, jeśli zmierzono ją w ciągu 1 minuty. Te same 10 mm na dobę to opad wysoki, lecz daleki od ekstremum. W skali miesiąca zaś jest to wartość skrajnie niska.

Pierwsze dane sugerujące istnienie regularnych przestrzennych relacji skali opublikował w roku 1982 S. Lovejoy. Opierając się na analizie obrazów satelitarnych i radarowych stwierdził on występowanie silnej relacji pomiędzy obwodem ( $U$ ) a powierzchnią ( $F$ ) obszarów chmur i opadów. Wyraził ją w postaci  $U^2 \sim F^D$ , gdzie  $D = 1,35$ . Wykładnik  $D$  jest wymiarem fraktalnym chmur i obszarów opadów (Lynch 2007, Peitgen i in. 1997, Ratajczak 1998, Sornette 2006). Nowsze opracowanie tego zagadnienia opublikowali w 2003 roku Hauf i Theusner (ryc. 2). Źródłem danych było 3258 kompozytowych obrazów radarowych o rozdzielczości 2 na 2 km z 39 dni w okresach letnich lat 1997 i 1998, kiedy nad terytorium Niemiec i sąsiednich krajów (analizowany obszar miał 920 na 920 km) przemieszczały się zimne fronty. Na obrazach tych zidentyfikowano około 140 000 „obiektów”. Średnio 28% z nich to były skupienia chmur (komórek konwekcyjnych), a 72% – pojedyncze komórki. Zróżnicowanie tych proporcji było bardzo niewielkie, bo wynosiło jedynie  $\pm 3\%$ . Ta stała relacja wskazuje na istnienie stanu równowagi procesów tworzenia się skupień chmur w wyniku łączenia pojedynczych komórek i ich zaniku. Jest to niezależne od ilości istniejących w danej chwili komórek. Powierzchnia pojedynczych chmur wynosiła średnio 240 km<sup>2</sup>. Ich

skupienia miały powierzchnię zróżnicowaną w zależności od pory doby: od 2500 km<sup>2</sup> w nocy do 4500 km<sup>2</sup> w południe. Kiedy znano ilość komórek tworzących skupienie, można było oszacować ich średnią powierzchnię. Z obliczeń tych wynikało, że komórki stanowiące składniki skupień są około 2,5 razy większe od istniejących samodzielnie. Ich powierzchnia w trakcie dnia mieściła się bowiem w zakresie od 520 do 700 km<sup>2</sup>. Przyczyna tego stanu rzeczy nie jest jeszcze znana. Autorzy cytowanej pracy (Hauf i Theusner 2003) sugerują, że może się na to składać efekt dwóch zjawisk. Po pierwsze, tworzenie skupień najczęściej obejmuje największe z izolowanych komórek, po drugie – po połączeniu prawdopodobnie ulegają one jeszcze powiększeniu. Opisana wyżej, bardzo istotna różnica powierzchni komórek izolowanych i stanowiących elementy skupień zwraca uwagę na znaczenie skupień chmur.



**Ryc. 2.** Relacja między powierzchnią ( $F$ , km<sup>2</sup>) a obwodem ( $U^2$ , km<sup>2</sup>) pojedynczych chmur *Cumulonimbus* i ich skupień generujących opady konwekcyjne na terenie Niemiec (Hauf, Theusner 2003). Kolorem przedstawiono ilość uwzględnionych chmur/skupień w poszczególnych klasach powierzchni i obwodu. Dokładne objaśnienia w tekście 

Rycina 2 powstała na podstawie opisanych wyżej danych (Hauf i Theusner 2003). Zakres zmienności  $U^2$  względem  $F$  obejmuje cztery rzędy wielkości. W zakresie pierwszego rzędu wielkości relacja pomiędzy obiema zmiennymi na wykresie podwójnie logarytmicznym jest linią prostą odpowiadającą funkcji  $U^2 \sim F^D$ , gdzie  $D = 1,359$ . Jest wartość bardzo zbliżona do określonej przez Lovejoya (1982), co sugeruje jej potencjalnie bardziej ogólne znaczenie<sup>6</sup>. Bardziej szczegółowy ogłąd wykresu ujawnia jednak istnienie nieznaczej, ale wyraźniej „krzywizny” danych. Według Haufa i Theusnera (2003) wskazywałoby to na bardziej skomplikowaną zależność. Badacze ci ustalili, że użycie modelu składającego się z dwóch segmentów – jednego dla komórek izolowanych i złożonych z maksymalnie trzech pojedynczych oraz drugiego dla większych i bardziej złożonych układów – redukuje znacznie rozrzut danych. Pierwszy segment ma wartość  $D$  wynoszącą około 1,25, drugi – około 1,54. Wyniki te wskazywałyby że reguły skalowania przestrzennego nie mają charakteru uniwersalnego wymiaru fraktalnego, ale multifraktalny (Ferraris i in. 2003). Konkluzją autorów omawianego opracowania (Hauf i Theusner 2003) jest stwierdzenie, że odpowiedzialne są za to zarówno istotne różnice dotyczące struktury pojedynczych komórek i ich skupień, jak i procesy odpowiedzialne za ich powstawanie i ewolucję. Przywołując rozważania przedstawione we wprowadzeniu, na temat potencjalnego wpływu zmian struktury przestrzennej opadów na ich rejestrowaną punktowo częstość, można wskazać, że efekt taki wywołałaby na przykład nawet nieznaczna zmiana owych podanych powyżej współczynników skalujących relacje między powierzchnią a obwodem chmur.

Strukturę przestrzenną opadów z użyciem różnych miar autokorelacji przestrzennej analizowano już wielokrotnie (Bacchi, Kottogoda 1995, Berne i in. 2006, Ciach, Krajewski 2006, Ferraris i in. 2003, Goodrich i in. 1995,

---

<sup>6</sup> Podczas przygotowywania niniejszej rozprawy do druku ukazała się kolejna publikacja (Lovejoy i in. 2008) podkreślająca uniwersalność owych relacji. O ile analizowane w poprzednich pracach dane miały zazwyczaj zasięg regionalny i wąskie ramy czasowe, to cytowani powyżej autorzy przeprowadzili obliczenia oparte na pomiarach wykonanych czujnikiem PR w trakcie kolejnych pełnych 1176 okrążeń satelity TRMM. Stwierdzili oni, że stochastyczny multifraktalny trójparametryczny model kaskadowy wyjaśnia obserwowaną zmienność przestrzenną tłumienia sygnału radarowego z dokładnością rzędu  $\pm 4,6\%$  dla skal przestrzennych od 4,3 do około 30 tys. km. Wartość indeksu multifraktalnego oceniono na 1,5 oraz stwierdzono, że odbicie sygnału może być modelowane jako czysty proces multiplikatywny, co znaczy, że jest on zachowany w kolejnych skalach przestrzennych. Autorzy pokazali również, że rozszerzenie stosowalności modelu do skali, w której następuje rozdzielanie turbulencji i opadu (ok. 40 cm w wypadku słabego deszczu) daje prognozy częstości występowania mierzalnych opadów o zadowalającej dokładności. Taki wynik stoi w wyraźnej sprzeczności z klasycznym ujęciem przedstawionym na ryc. 1, które łączy morfologię z mechanizmem zjawisk i dzieli zakres 4 do 20 tys. km na szereg podzakresów, z których każdy jest zdominowany przez inny typ zjawisk.



Haberlandt 2007, Hauf, Theusner 2003, Joseph i in. 2000, Moszkowicz 2000, Skaugen 1997, Yoo, Ha 2007). Jako materiał źródłowy w niektórych z nich wykorzystywano ciągle rejestracje pluwiograficzne lub teledetekcyjne, co umożliwiało wykonanie pełnej oceny struktury przestrzennej dla opadów o różnej wysokości i czasie trwania.

Dobrym przykładem, ważnym przede wszystkim ze względów geograficznych, jest praca Moszkowicza (2000). Autor dysponował dwoma źródłami danych: specjalną siecią cyfrowych deszczomierzy rejestrujących (z korytkami wywrotnymi) oraz wynikami pomiarów wykonanych z pomocą radaru meteorologicznego zlokalizowanego w Legionowie pod Warszawą. Deszczomierzy było 16, zlokalizowanych w odległościach od 10 m do 5 km, a najkrótszy interwał rejestracji wynosił 10 minut. Wykorzystane w pracy obrazy radarowe miały rozdzielczość 2 na 2 km, a ich zasięg, w celu zminimalizowania potencjalnych błędów, ograniczono do 100 km. Efektem przetworzenia oryginalnych wartości tłumienia sygnału radarowego, pozyskiwanych co 10 minut, były szacowane wartości intensywności opadu, określenie jego typu (konwekcyjny lub frontalny) oraz wektorów przemieszczania się. Cytowany autor (Moszkowicz 2000) wykorzystał w swoim opracowaniu pomiary naziemne z jednego sezonu rocznego z wyłączeniem 5 miesięcy zimowych oraz jedynie 7 przypadków opadów zarejestrowanych za pomocą radaru. Korzystając z danych pochodzących z deszczomierzy, mógł zatem przeprowadzić analizę struktury przestrzennej opadów na bardzo krótkich dystansach i przy pełnym spektrum czasu ich trwania. Była ona jednak ograniczona, ze względu na liczbę i rozmieszczenie stanowisk, poprzez założenie izotropowości i maksymalny zasięg 5 km. Z kolei, obrazy radarowe umożliwiały przeprowadzenie analizy anizotropowej osobno dla opadów konwekcyjnych i frontalnych, ale ich ograniczeniem była mała liczba analizowanych przypadków, rozdzielczość przestrzenna i zasięg ograniczony do kilkudziesięciu kilometrów.

Zarówno w pracy Moszkowicza (2000), jak też w innych przytoczonych wyżej nie dysponowano tak dużym, jak w niniejszym opracowaniu, zbiorem danych i to wyłącznie ekstremalnych opadów. Zazwyczaj także analiza ograniczała się do znacznie mniejszych obszarów i krótszych przedziałów czasu. Na podstawie materiału pomiarowego, którym dysponowano, można się było pokusić o zbadanie zmienności sezonowej, a po obliczeniu semiwariogramów dla sekwencji wartości progowych – także zróżnicowania w obrębie poszczególnych klas wysokości opadów. Należy jednak uwzględnić również, że analizowane zbiory danych stanowią nierozdzieloną „mieszani-nę” różnych genetycznie opadów, zarejestrowanych w różnych terminach.

# III

---

## Cel, zakres i metody pracy

### 1. Cel i zakres pracy

Głównym celem pracy była analiza prawidłowości przestrzennej i czasowej zmienności miesięcznych oraz rocznych maksymalnych dobowych sum opadów (MSDO).

Specyfika tych danych polega na tym, że oprócz typowej dla opadów nieciągłości przestrzennej są także niesynchroniczne. Podana wartość miesięcznej MSDO mogła wystąpić w dowolnym stanowisku w którymkolwiek dniu miesiąca. O ile można z uproszczeniem przyjmować, że pole sum miesięcznych, a zwłaszcza rocznych czy wieloletnich ma charakter ciągły, to w tym wypadku założenie to nie ma uzasadnienia. Mierzone punktowo MSDO nie są jednakże efektem niezależnych, całkowicie losowych zjawisk. Każdy epizod opadowy trwa przez pewien czas i ma określony zasięg przestrzenny. Istnieje zatem niezerowe prawdopodobieństwo, że zostanie zarejestrowany w większej liczbie stanowisk i przynajmniej w części z nich będzie zaklasyfikowany jako okresowa MSDO. Założeniem metodologicznym niniejszej pracy była hipoteza (potwierdzona w dalszym toku postępowania), że prawdopodobieństwo wystąpienia MSDO charakteryzuje się autokorelacją przestrzenną, a zatem ciągłością przestrzenną. Logiczną konsekwencją takiego podejścia jest możliwość zastosowania do analizy tych danych metod geostatystycznych (Chilés, Delfiner 1999, Cressie 1993, Goovaerts 1997, Isaaks, Srivastava 1989, Namysłowska-Wilczyńska 2006, Webster, Oliver 2001, Wackernagel 2003, Zawadzki 2005).

Główny cel rozprawy był realizowany poprzez następujące cele szczegółowe:

- określenie głównych cech statystycznych analizowanych zbiorów miesięcznych i rocznych MSDO, a w tym także zmienności sezonowej i wieloletniej;

- identyfikację i typologię struktury przestrzennej miesięcznych oraz rocznych MSDO;
- określenie charakteru i przypuszczalnej genezy zjawisk generujących MSDO;
- ilościową ocenę udziału poszczególnych zjawisk w całkowitych rozmiarach MSDO;
- zbadanie potencjalnej zmienności sezonowej struktury przestrzennej MSDO;
- weryfikację hipotezy o zróżnicowaniu struktury przestrzennej MSDO w różnych klasach wysokości opadów;
- sprawdzenie, czy struktura przestrzenna MSDO wykazuje istotne tendencje w całym analizowanym okresie;
- określenie sezonowego zróżnicowania prawdopodobieństwa wystąpienia i wysokości rocznych MSDO;
- zbadanie zmienności regionalnej terminu występowania rocznych MSDO.

Ten szeroko zakrojony program wymagał wykonania wielu czasochłonnych prac, związanych z utworzeniem i weryfikacją źródłowej bazy danych, selekcją i przetwarzaniem numerycznym różnorodnych podzbiorów MSDO, a później analizą oraz interpretacją uzyskanych wyników. Najważniejsze z nich to:

- analiza statystyczna zmian ilości i rozmieszczenia przestrzennego sieci punktów danych w analizowanym wieloleciu;
- globalna i lokalna statystyczna charakterystyka różnorodnych podzbiorów bazy danych MSDO;
- budowanie modeli wielomianowych potrzebnych do normalizacji każdego z 325 podzbiorów miesięcznych i rocznych MSDO;
- binarne kodowanie każdego podzbioru według 13 wartości progowych, określonych na podstawie empirycznych funkcji skumulowanego rozkładu;
- maskowanie, w każdym analizowanym podzbiorze, wartości odstających, zaburzających obraz struktury przestrzennej;
- obliczanie 325 semiwariogramów empirycznych danych normalizowanych i 4225 danych kodowanych;
- interaktywne modelowanie matematyczne 4550 semiwariogramów;
- klasyfikacja uzyskanych modeli struktury przestrzennej; testowanie istotności ich zróżnicowania w zależności od sezonu, względnej i bezwzględnej wysokości opadów oraz miejsca w wieloletnim ciągu pomiarowym;
- optymalizacja algorytmu interpolacji prawdopodobieństwa wystąpienia i wysokości rocznych MSDO;
- interpolacja statystyk terminów wystąpienia rocznych MSDO.

Podstawy zastosowanej metodyki przedstawiono w kolejnym podrozdziale (III.2). Szczegółowe, specyficzne dla analizowanego zbioru danych zagadnienia metodyczne opisano w dodatkach (rozdz. X). Sposoby rozwiązania szeregu drobniejszych problemów metodycznych omawiane są także w rozdziałach analitycznych, wraz z prezentacją uzyskanych wyników.

## 2. Podstawy metodyki

### 2.1. Wprowadzenie

Podstawę metodyki niniejszej analizy zmienności przestrzennej maksymalnych sum dobowych opadów na obszarze Polski stanowi geostatystyka. Ponieważ w ostatnich latach ukazały się drukiem dwa obszernie opracowania, które po polsku prezentowały teorię i zastosowania tej dziedziny statystyki przestrzennej (Namysłowska-Wilczyńska 2006, Zawadzki 2005), autor czuje się zwolniony z obowiązku szczegółowego omówienia jej podstaw. Zostaną one przedstawione jedynie w takim zakresie, jaki jest niezbędny dla uzasadnienia zastosowania oraz wyjaśnienia teorii i algorytmów konkretnych, wykorzystanych w niniejszej pracy technik geostatystycznych. Bardziej dokładnie omówione zostaną te zagadnienia, które w cytowanych wyżej pracach były pominięte, albo potraktowane skrótowo, a mają istotne znaczenie dla zrozumienia uzyskanych przez autora wyników. Niezbędne jest przy tym używanie zwartej i jednoznacznej notacji matematycznej. We wszystkich takich przypadkach używano konwencji wprowadzonej w serii podręczników wydawanych przez Oxford University Press pod wspólną nazwą "Applied Geostatistics Series" (Deutsch, Journel 1992, 1998, Goovaerts 1997).

### 2.2. Miary ciągłości lub zmienności przestrzennej

Niech  $z(\mathbf{u}_\alpha)$ ,  $\alpha = 1, 2, \dots, n$  określa zbiór  $n$  wartości pomiarów dowolnej cechy  $z$  (wielkości) dokonanych w obrębie badanego obszaru, gdzie  $\mathbf{u}_\alpha$  oznacza wektor współrzędnych konkretnej obserwacji  $\alpha$ . W większości przypadków nie są to wartości kompletnie losowe, co oznacza, że pomiary wykonane bliżej są zazwyczaj do siebie bardziej podobne od tych, które dzieli większa odległość. Podobieństwo to można określić ilościowo, porównując wyniki pomiarów dla stanowisk odległych od siebie o coraz większe odległości. Symbolicznie określane to jest jako porównanie dowolnej danej  $z$  określonej w lokalizacji  $\mathbf{u}_\alpha$ , czyli  $z(\mathbf{u}_\alpha)$ , z dowolną inną odległą o wektor  $\mathbf{h}$ ,

czyli  $z(\mathbf{u}_\alpha + \mathbf{h})$ . W wypadku każdego typu próbkowania poza regularnym, wartość wektora  $\mathbf{h}$  oznacza w rzeczywistości pewien przedział odległości<sup>7</sup>. Na przykład, porównujemy pomiary odległe od siebie 0-1 km, 1-2 km, 2-3 km i tak dalej<sup>8</sup>. Najprościej można tego dokonać używając wykresu kartezjańskiego XY. Zazwyczaj za pomocą tego wykresu przedstawia się relacje między dwoma cechami (parametrami) zmierzonymi w tej samej lokalizacji i/lub w tym samym czasie. Tym razem służy on do porównania wartości tej samej cechy zmierzonej w dwóch różnych lokalizacjach i w związku z tym na osi X odkłada się wartości  $z(\mathbf{u}_\alpha)$ , a na Y –  $z(\mathbf{u}_\alpha + \mathbf{h})$ . Wykres taki nazywany jest „*h-scattergram*”, co można przetłumaczyć na „wykres rozrzutu z przesunięciem  $\mathbf{h}$ ” (Goovaerts 1997, Pannatier 1996, Zawadzki 2005). W geostatystyce przyjęto określenie „ogona” (ang. *tail value*) dla wartości będącej początkiem wektora  $\mathbf{h}$ , czyli  $z(\mathbf{u}_\alpha)$ , podczas gdy wartość stanowiąca jego koniec, czyli  $z(\mathbf{u}_\alpha + \mathbf{h})$ , nazywana jest „głową” (ang. *head value*). Jako przykład zaprezentowano wykresy rozrzutu z przesunięciem dla pierwszych sześciu odstępów, o szerokości 2,5 km każdy, zbioru danych MSDO z lipca roku 1977 (ryc. 3). Pokazują one wyraźnie, wyczuwaną intuicyjnie, właściwość spadku podobieństwa wyników pomiarów wraz z odległością. Na kolejnych wykresach chmura punktów staje się coraz „szersza” – są one bardziej oddalone od przekątnej symbolizującej idealną zależność wprost proporcjonalną. Tradycyjnie do wyrażenia siły relacji między dwoma zmiennymi używa się współczynnika korelacji liniowej Pearsona, oznaczanego symbolem  $\rho$ . Jest on standaryzowaną (niezależną od skali pomiarowych) formą kowariancji obu zmiennych. W tym wypadku można go wyrazić w sposób następujący (Goovaerts 1997):

$$\rho(\mathbf{h}) = \frac{C(\mathbf{h})}{\sqrt{\sigma_{-\mathbf{h}}^2 \cdot \sigma_{+\mathbf{h}}^2}} \quad \in [-1, +1] \quad [1]$$

gdzie:  $C(\mathbf{h})$  oznacza autokowariancję par pomiarów oddalonych od siebie o  $\mathbf{h}$  (wzór [2]), natomiast  $\sigma_{-\mathbf{h}}^2$  i  $\sigma_{+\mathbf{h}}^2$  – odpowiednio wariancję podzbioru danych ogona i głowy (wzory [3] i [4]):

$$C(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [z(\mathbf{u}_\alpha) \cdot z(\mathbf{u}_\alpha + \mathbf{h}) - m_{-\mathbf{h}} \cdot m_{+\mathbf{h}}] \quad [2]$$

<sup>7</sup> Przy analizie kierunkowej (anizotropowej) również kierunek wektora ma charakter przedziałowy.

<sup>8</sup> Żeby wyraźnie zaznaczyć charakter wektora  $\mathbf{h}$ , który nie określa jednej konkretnej odległości między konkretną parą lokalizacji, ale ogólną różnicę położenia w przestrzeni dowolnej pary danych, w geostatystyce nazwa się go *lag*, a nie *distance*. Najbardziej trafnym polskim odpowiednikiem tego pojęcia jest „odstęp”, ewentualnie „przesunięcie”. W niniejszej rozprawie stosowane będzie najczęściej określenie pierwsze.

$$\sigma_{\cdot\mathbf{h}}^2 = \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [z(\mathbf{u}_{\alpha}) - m_{\cdot\mathbf{h}}]^2 \quad [3]$$

$$\sigma_{+\mathbf{h}}^2 = \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [z(\mathbf{u}_{\alpha} + \mathbf{h}) - m_{+\mathbf{h}}]^2 \quad [4]$$

$N(\mathbf{h})$  oznacza liczbę par danych w obrębie określonej klasy odległości i kierunku, a symbolami  $m_{\cdot\mathbf{h}}$  i  $m_{+\mathbf{h}}$  we wzorze autokowariancji [2] i wzorach wariancji podzbiorów ogona [3] i głowy [4] określono odpowiednie średnie arytmetyczne obu tych podzbiorów:

$$m_{\cdot\mathbf{h}} = \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} z(\mathbf{u}_{\alpha}) \quad [5]$$

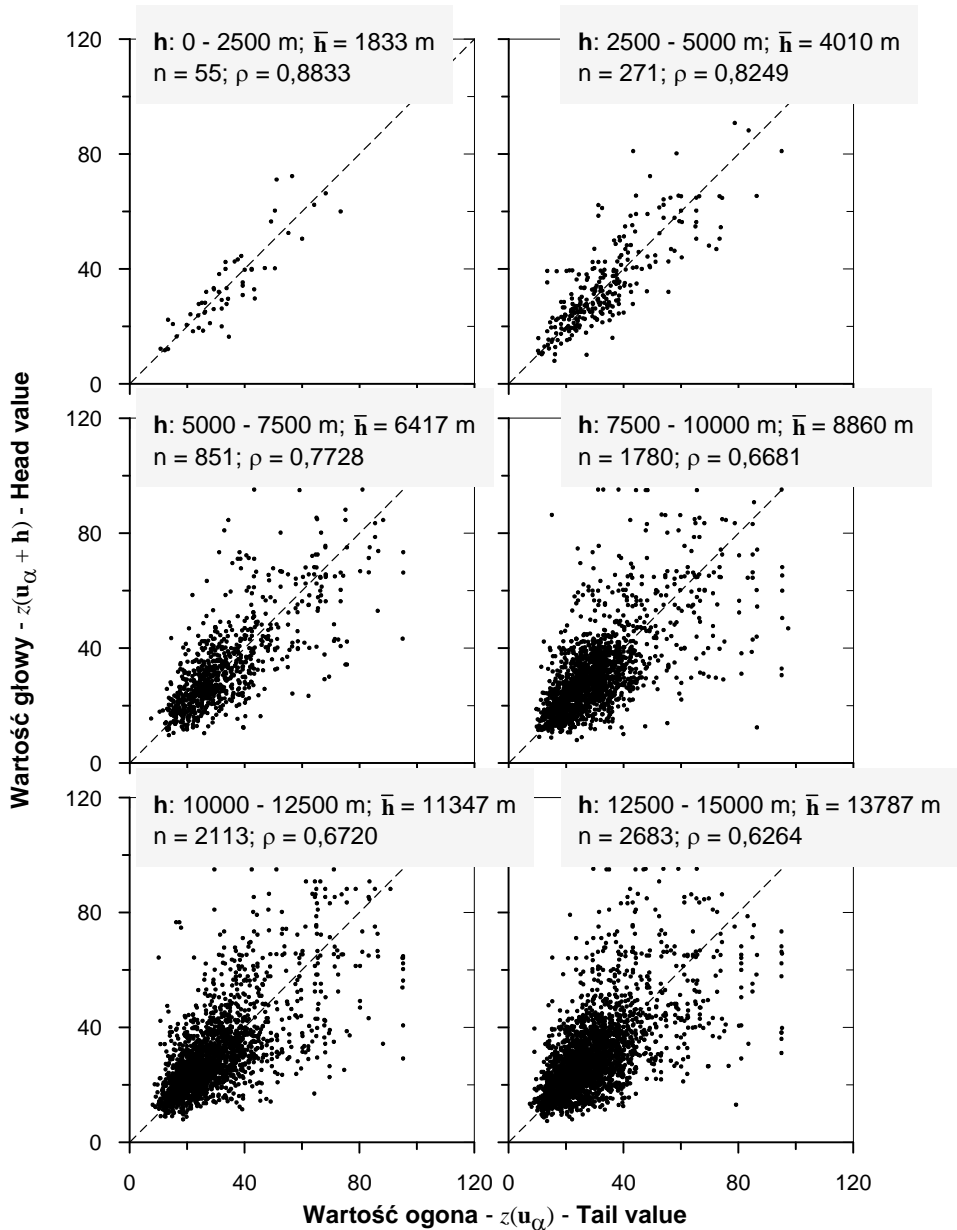
$$m_{+\mathbf{h}} = \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} z(\mathbf{u}_{\alpha} + \mathbf{h}) \quad [6]$$

Na wykresie (ryc. 3) współczynnik korelacji (autokorelacji) maleje od 0,8833 dla par punktów należących do pierwszego odstępu (0-2,5 km) do 0,6264 dla szóstego odstępu (pary odległe o 12,5-15,0 km). Obliczone, dla kolejnych rosnących odstępów, współczynniki autokorelacji, przedstawione w relacji do średniej odległości par należących do danego odstępu, tworzą tak zwany korelogram eksperymentalny. Rycina 4 jest przykładem takiego wykresu dla danych MSDO z lipca 1977 roku. Wynika z niego, że oprócz kilku niewielkich wahań, spadek autokorelacji następuje konsekwentnie do odległości około 150 km. Na takim dystansie nie ma już statystycznie żadnego podobieństwa wyników pomiarów – współczynnik korelacji oscyluje wokół zera.

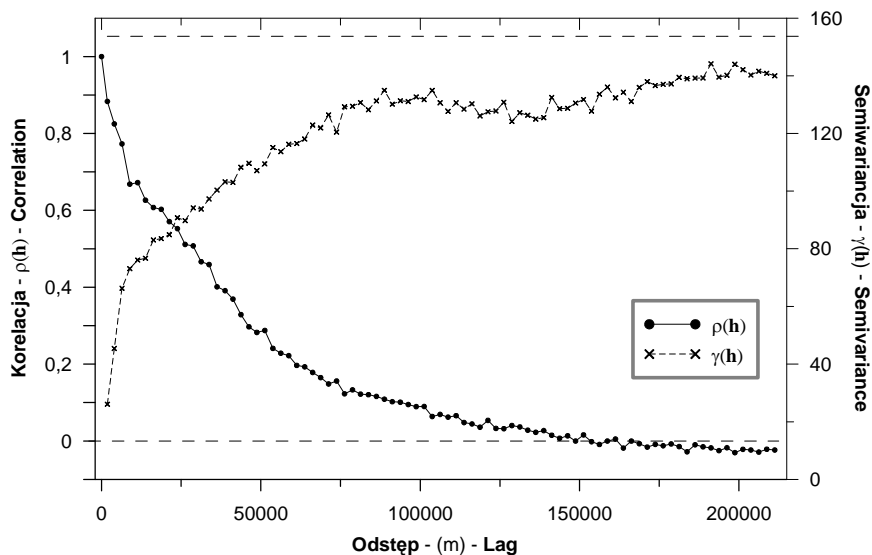
Alternatywnym, i silniej ugruntowanym w literaturze geostatystycznej, sposobem przedstawiania relacji przestrzennych jest kategoria niepodobieństwa (zamiast podobieństwa) pomiędzy obserwacjami, jako funkcji dzielącej je odległości. Miarą średniego niepodobieństwa jest semiwariancja, zdefiniowana jako połowa średniej kwadratów różnic wartości cechy w lokalizacjach odległych o wektor  $\mathbf{h}$  (Gringarten, Deutsch 2001, Goovaerts 1997, Zawadzki 2005):

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [z(\mathbf{u}_{\alpha}) - z(\mathbf{u}_{\alpha} + \mathbf{h})]^2 \quad [7]$$

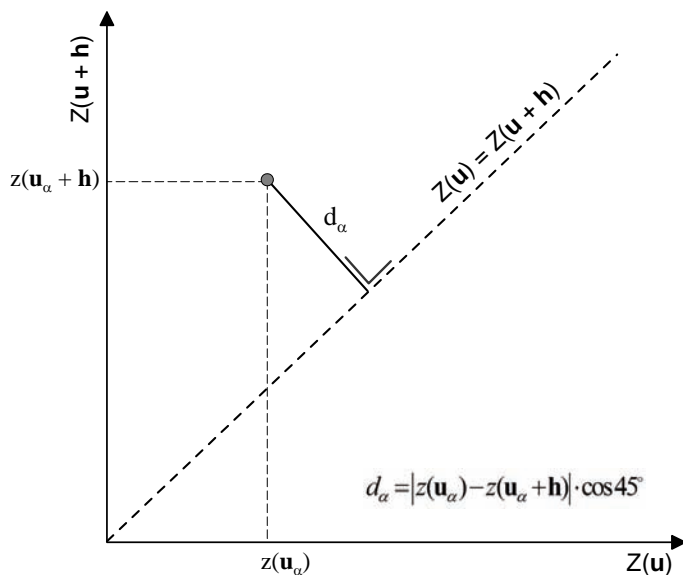
Geometryczną interpretację semiwariancji stanowi odległość dzieląca każdą naniesioną na wykres rozrzutu z przesunięciem parę pomiarów oddalonych o wektor  $\mathbf{h}$  od przekątnej wykresu (pierwszego bisektora) symbolizującej idealną relację, wprost proporcjonalną między porównywanymi cechami (Goovaerts 1997, ryc. 5).



**Ryc. 3.** Wykresy rozrzutu z przesunięciem (ang. *h-scattergram*) pierwszych 6 odstępów, o szerokości 2500 m każdy, dla danych MSDO z lipca roku 1977. Przy wykresach podano zakres odległości, średni odstęp dla wszystkich par danych znajdujących się w danym przedziale, liczbę par i współczynnik korelacji liniowej pomiędzy wynikami pomiarów. Należy zwrócić uwagę na konsekwentny spadek korelacji MSDO wraz ze wzrostem odstepu między stanowiskami



Ryc. 4. Korelogram i semiwariogram danych MSDO z lipca 1977 obliczony dla 85 przedziałów odległości po 2500 m. Na wykresie zaznaczono zerowy poziom autokorelacji i wariancję próby wynoszącą 153,72, a oznaczanej w geostatystyce jako  $C(0)$  – autokowariancja dla odstępu równego 0



Ryc. 5. Interpretacja geometryczna wartości semiwariogramu  $\gamma(h)$ , jako średniej wszystkich podniesionych do kwadratu ortogonalnych odległości  $d_\alpha$  od przekątnej wykresu (pierwszego bisektora) rozrzutu z przesunięciem (Goovaerts 1997)



W geostatystyce wynik pomiaru  $z(\mathbf{u}_\alpha)$  jest interpretowany jako jedna z możliwych realizacji funkcji losowej  $FL(\mathbf{u}_\alpha)$ , która jest w pełni charakteryzowana przez jej rozkład prawdopodobieństwa  $F(\mathbf{u}_\alpha; z) = \text{Prob}\{Z(\mathbf{u}_\alpha) \leq z\}$ . Dla stacjonarnej  $FL^9$  istnieje funkcyjna relacja pomiędzy semiwariancją a poprzednio zdefiniowanymi ([1] i [2]) autokorelacją i autokowariancją (Goovaerts 1997):

$$\begin{aligned} 2\gamma(\mathbf{h}) &= \text{Var}\{Z(\mathbf{u} + \mathbf{h}) - Z(\mathbf{u})\} \\ \gamma(\mathbf{h}) &= C(0) - C(\mathbf{h}), \quad \forall \mathbf{u} \end{aligned} \quad [8]$$

oraz,

$$\rho(\mathbf{h}) = 1 - \frac{\gamma(\mathbf{h})}{C(0)} \quad [9]$$

gdzie  $C(0)$  oznacza wariancję próby.

Na wykresie (ryc. 4) semiwariogram danych MSDO z lipca 1977 roku rośnie, a jego obraz jest do pewnego stopnia lustrzanym odbiciem przebiegu korelogramu. Są jednak wyraźne różnice. Wykres osiąga maksimum (staje się płaski) na dystansie około 80 km, a dalsze zmiany mają charakter raczej oscylacji niż konsekwentnego wzrostu. Odległość, na której zachodzi stabilizacja semiwariancji nazywana jest zasięgiem (ang. *range*), a jej wartość – semiwariancją progową (ang. *sill*). Wzrost niepodobieństwa reprezentowany przez semiwariogram nie zachodzi jednak w stałym tempie. Wykres wykazuje co najmniej dwa załamania: jedno na dystansie około 9 km, drugie, słabiej widoczne, przy odstępnie około 55 km. Określa się je jako zasięgi cząstkowe (ang. *partial ranges*), a odpowiadające im wartości semiwariancji jako progi cząstkowe (ang. *partial sills*). Świadczą one, że analizowana cecha (parametr) jest efektem działania kilku zjawisk operujących w różnych skalach przestrzennych (Goovaerts 1997).

Wraz ze wzrostem odstepu  $|\mathbf{h}|$  korelacja między jakimikolwiek dwoma zmiennymi losowymi  $Z(\mathbf{u})$  i  $Z(\mathbf{u} + \mathbf{h})$  zazwyczaj dąży do zera:

$$C(h) \rightarrow 0 \quad \text{dla} \quad |\mathbf{h}| \rightarrow \infty \quad [10]$$

Biorąc pod uwagę zależność [8], wariancja progowa (ang. *sill*) semiwariogramu ograniczonego dąży do wariancji  $C(0)$ :

---

<sup>9</sup> Model funkcji losowej spełnia założenie stacjonarności jeśli: (1) wartość oczekiwana  $E\{Z(\mathbf{u})\}$  istnieje i nie zależy od lokalizacji w obrębie analizowanego obszaru, (2) dla każdej pary zmiennych losowych  $\{Z(\mathbf{u}), Z(\mathbf{u} + \mathbf{h})\}$  istnieje ich kowariancja, która jest zależna jedynie od wektora odstepu  $\mathbf{h}$ .

$$\gamma(\mathbf{h}) \rightarrow C(0) \text{ dla } |\mathbf{h}| \rightarrow \infty \quad [11]$$

Nieciągłość na początku semiwariogramu (to jest dla odstępów równego zero) nazywana jest efektem nuggetowym<sup>10</sup> (Goovaerts 1997, Gringarten, Deutsch 2001, Chilés, Delfiner 1999). Jego źródłem są błędy pomiarowe i/lub zmienność przestrzenna w skali mniejszej niż najkrótszy odstęp pomiarów, uwzględniając w tym także występującą w obrębie próbek<sup>11</sup>.

### 2.3. Struktura przestrzenna w klasach natężenia analizowanej cechy

Funkcja autokowariancji i semiwariogram to charakterystyki ciągłości przestrzennej (lub zmienności) dla całego zakresu wartości cechy. Struktura ciągłości przestrzennej może jednak różnić się, zależnie czy pod uwagę bierzemy rozkład punktów danych o niskich, średnich czy wysokich wartościach (Deutsch, Journel 1998, Goovaerts 1997). W wielu sytuacjach spotykanych w środowisku przyrodniczym lub społeczno-gospodarczym, losowo występujące wysokie wyniki pomiarów są otoczone większymi obszarami o średnich lub niskich wartościach, które zmieniają się w sposób ciągły i stopniowy. Czy wartości ekstremalne są w przestrzeni rozproszone czy skupione, jaki jest ich zasięg ma duże znaczenie dla wyjaśniania zjawiska oraz jakości estymacji.

Określenie prawidłowości rozkładu przestrzennego wartości cechy z występujących powyżej lub poniżej ustalonego poziomu progowego  $z_k$  wymaga uprzedniego przekodowania każdego wyniku pomiaru  $z(\mathbf{u}_\alpha)$  do formy binarnej zgodnej z poniższą regułą:

$$i(\mathbf{u}_\alpha; z_k) = \begin{cases} 1 & \text{jeżeli } z(\mathbf{u}_\alpha) \leq z_k \\ 0 & \text{poza tym} \end{cases} \quad [12]$$

Dane kodowane  $i(\mathbf{u}_\alpha; z_k)$  mogą dalej być analizowane przy użyciu każdej z wymienionych poprzednio miar struktury przestrzennej. Po odpowiednim zmodyfikowaniu wzoru [2] oblicza się na jego podstawie eksperymentalną autokowariancję kodów:

<sup>10</sup> Źródłem tej terminologii były analizy zasobów złóż złota w RPA, dokonywane w latach pięćdziesiątych XX wieku przez D. Krige'a (1951, 1952). Głównym powodem występowania nieciągłości było tam losowe występowanie samorodków złota (ang. *nugget*). W związku z tym w polskiej literaturze geostatystycznej używa się często terminu „efekt samorodka” (Zawadzki 2005, Namysłowska-Wilczyńska 2006).

<sup>11</sup> W tym przypadku zmienność w obrębie powierzchni zbiorczej deszczomierza, czyli w skali około 16 cm.

$$\begin{aligned}
C_I(\mathbf{h}; z_k) &= \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} i(\mathbf{u}_\alpha; z_k) \cdot i(\mathbf{u}_\alpha + \mathbf{h}; z_k) - F_{-\mathbf{h}}(z_k) \cdot F_{+\mathbf{h}}(z_k) \\
&= F(\mathbf{h}; z_k) - F_{-\mathbf{h}}(z_k) \cdot F_{+\mathbf{h}}(z_k)
\end{aligned} \tag{13}$$

gdzie:

$$F_{-\mathbf{h}}(z_k) = \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} i(\mathbf{u}_\alpha; z_k) \quad F_{+\mathbf{h}}(z_k) = \frac{1}{N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} i(\mathbf{u}_\alpha + \mathbf{h}; z_k)$$

$F_{-\mathbf{h}}(z_k)$  i  $F_{+\mathbf{h}}(z_k)$  oznaczają proporcje (ułamek) wartości ogona i głowy nie przekraczających poziomu wartości progowej  $z_k$ . Autokowariancja kodów określa, jak często dwie wartości tej samej cechy oddalone od siebie o wektor  $\mathbf{h}$  są jednocześnie nie większe od wartości progowej  $z_k$ .

Autokowariancja kodów podzielona przez pierwiastek z iloczynu wariancji podzbiorów ogona i głowy przybiera postać standaryzowaną – eksperymentalną autokorelację kodów:

$$\rho_I(\mathbf{h}; z_k) = \frac{C_I(\mathbf{h}; z_k)}{\sqrt{\sigma_{-\mathbf{h}}^2(z_k) \cdot \sigma_{+\mathbf{h}}^2(z_k)}} \in [-1, +1] \tag{14}$$

gdzie:  $\sigma_{-\mathbf{h}}^2(z_k) = F_{-\mathbf{h}}(z_k)[1 - F_{-\mathbf{h}}(z_k)]$  oznacza wariancję wartości kodów ogona, a  $\sigma_{+\mathbf{h}}^2(z_k) = F_{+\mathbf{h}}(z_k)[1 - F_{+\mathbf{h}}(z_k)]$  – wariancję wartości kodów głowy.

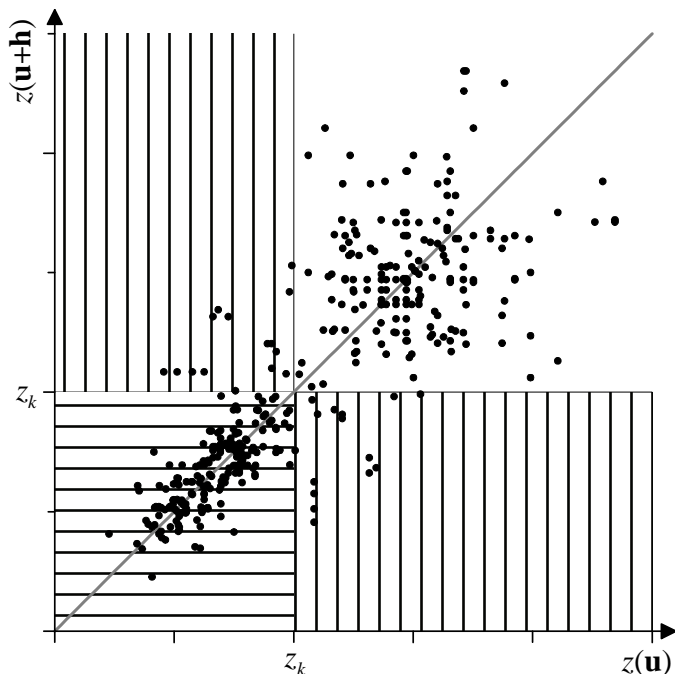
Analogicznie eksperymentalny semiwariogram kodów obliczany jest na podstawie wzoru [7] po podstawieniu zamiast wartości zmierzonej  $z(\mathbf{u}_\alpha)$  jej binarnej transformacji  $i(\mathbf{u}_\alpha; z_k)$ :

$$\gamma_I(\mathbf{h}; z_k) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [i(\mathbf{u}_\alpha; z_k) - i(\mathbf{u}_\alpha + \mathbf{h}; z_k)]^2 \tag{15}$$

Wariogram kodów ( $2\gamma(\mathbf{h}; z_k)$ ) określa, jak często dwie wartości analizowanej cechy oddalone o wektor  $\mathbf{h}$  znajdują się po przeciwnych stronach wartości progowej  $z_k$ <sup>12</sup>. Innymi słowy,  $2\gamma(\mathbf{h}; z_k)$  odzwierciedla częstość przejść między dwoma klasami wartości cechy jako funkcję odległości ( $\mathbf{h}$ ). Im jest większy, tym mniejszą ciągłość przestrzenną wykazują niskie lub wysokie wartości.

Kowariancja i semiwariancja danych kodowanych może być przedstawiona graficznie jako proporcja punktów (par danych) znajdujących się w określonych częściach wykresu rozrzutu z przesunięciem (ryc. 6, Goovaerts 1997). Autokowariancja kodów  $C_I(\mathbf{h}; z_k)$  odzwierciedla udział par danych ( $z(\mathbf{u}_\alpha), z(\mathbf{u}_\alpha + \mathbf{h})$ ), które jednocześnie nie przekraczają wartości progowej  $z_k$

<sup>12</sup> Semiwariogram wartości kodowanych (patrz wzór 15) jest tylko wtedy niezerowy, kiedy  $i(\mathbf{u}_\alpha; z_k)$  równa się 0, a  $i(\mathbf{u}_\alpha + \mathbf{h}; z_k)$  równa się 1, lub odwrotnie.



**Ryc. 6.** Autokowariancję i semiwariogram danych kodowanych można interpretować jako proporcję punktów (par danych), które występują w określonych częściach wykresu rozrzutu z przesunięciem: autokowariancja – obszar zasrafiowany poziomo, semiwariogram – obszar zasrafiowany pionowo (Goovaerts 1997). Do wykonania ryciny wykorzystano dane maksymalnej rocznej sumy dobowej opadu z 18 lipca 1970 roku. Na wykresie znajdują się wartości dla 392 par stanowisk odległych od siebie o 1,5-4,5 km. Wartość progową  $z_k$  stanowi suma dobową równą 100 mm

(ryc. 6: obszar zasrafiowany poziomo). W obliczonym wariogramie kodów  $2\gamma(\mathbf{h}; z_k)$  udział mają jedynie te pary  $z(\mathbf{u}_\alpha)$  i  $z(\mathbf{u}_\alpha + \mathbf{h})$ , które znajdują się po przeciwnych stronach ustalonej wartości progowej  $z_k$ . Stanowi go zatem ułamek całego zbioru par dla danego odstępów, zlokalizowany na rycinie 6 w obszarze zasrafiowanym pionowo.

#### 2.4. Przykład analizy danych kodowanych

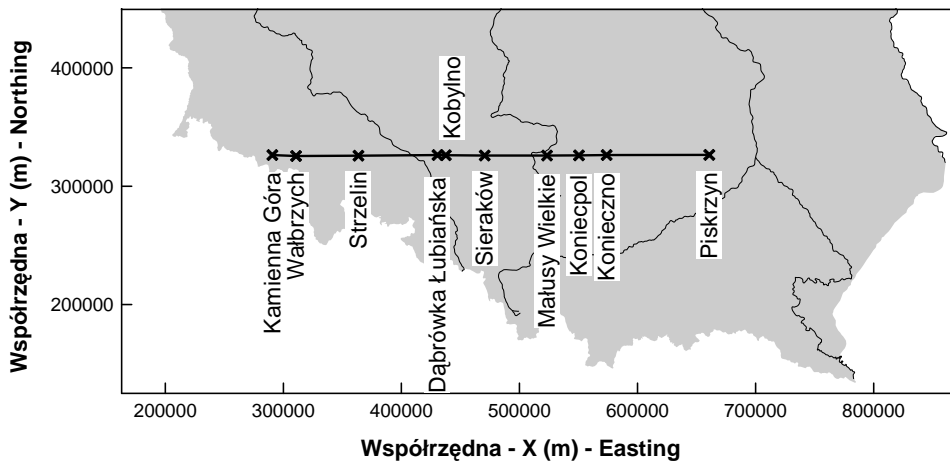
W celu prostego i obrazowego zilustrowania stosowanej w niniejszej rozprawie, raczej mało znanej i wykorzystywanej, metodyki analizy danych kodowanych posłużono się specjalnie do tego celu przygotowanym, jednowymiarowym przykładem (ryc. 7 i 8). Z całego zbioru maksymalnych opadów dobowych zarejestrowanych na terytorium Polski w maju 1980 roku

wybrano 10 stanowisk leżących w przybliżeniu na jednej rzędnej Y w układzie współrzędnych 1992 ( $N = 326\,000 \pm 500$  m). Profil o długości 370 km rozpoczyna się na zachodzie stanowiskiem Kamienna Góra w Sudetach Środkowych, a kończy na wschodzie w Piskorzynie na Wyżynie Sandomierskiej. Największe wysokości terenu są na obu końcach profilu: Kamienna Góra – 420, Wałbrzych – 490 oraz Małusy Wielkie – 280 i Piskorzyn – 300 m n.p.m. Najniżej usytuowane są stanowiska na Równinie Wrocławskiej (Strzelin, 165 m n.p.m.) oraz Równinie Opolskiej (Dąbrówka Łubiańska, Kobylno, 165 i 175 m n.p.m.). Omawiane dane zostały wybrane głównie ze względu na relatywnie duże zróżnicowanie zmierzonych w maju 1980 maksymalnych opadów dobowych (ryc. 8). Najniższe opady zarejestrowano w dwóch stanowiskach leżących na zachodnim krańcu profilu (5,5 i 4,2 mm). W kolejnych punktach pomiarowych usytuowanych dalej w kierunku wschodnim zanotowano coraz wyższe sumy dobowe, z maksimum w Małusach Wielkich – 52,5 mm. Dalej ku wschodowi opady konsekwentnie malały do wartości 11,0 mm w Piskorzynie. Na wybranym profilu duże jest także zróżnicowanie odległości sąsiadujących stanowisk – od 7 km pomiędzy Kobylnem a Dąbrówką Łubiańską do 87 km pomiędzy Piskorzynem a Koniecznem.

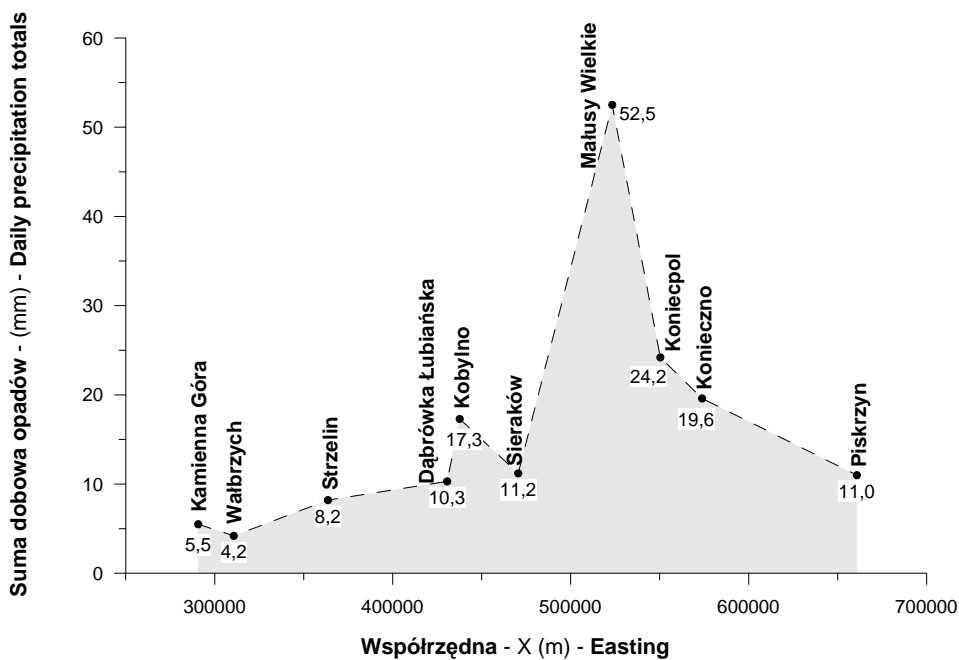
Procedura postępowania jest następująca:

**1) konstrukcja i dyskredytyzacja globalnej skumulowanej funkcji rozkładu** (ang. *cumulative distribution function* = *cdf*, ryc. 9). Na podstawie dostępnych danych pomiarowych (próby) obliczana jest i wykreślana krzywa *cdf*. W przedstawianym poniżej przykładzie posłużono się wszystkimi wynikami pomiarów maksymalnych sum dobowych opadów na terenie Polski w maju 1980. Następnie wybiera się wartości progowe służące do dyskredytyzacji *cdf*. Powinno ich być jak najmniej (ze względu na czasochłonność analizy), ale dostatecznie dużo, żeby uchwycić najbardziej charakterystyczne cechy rozkładu. Szczególną uwagę zwracać zazwyczaj trzeba na skrajne części krzywej, obrazujące częstość występowania wartości ekstremalnych. Z drugiej strony, wybranie bardzo skrajnych wartości progowych, na przykład 0,01 lub 0,99, przy małym zbiorze danych pomiarowych pociąga za sobą ryzyko dużych problemów z określeniem wiarygodnego modelu struktury przestrzennej. W prezentowanym jednowymiarowym przykładzie dla uproszczenia wybrano jedynie 6 wartości progowych odpowiadających 20, 40, 60, 80, 90 i 95 percentylowi rozkładu MSDO na obszarze Polski w maju 1980 roku (ryc. 9);

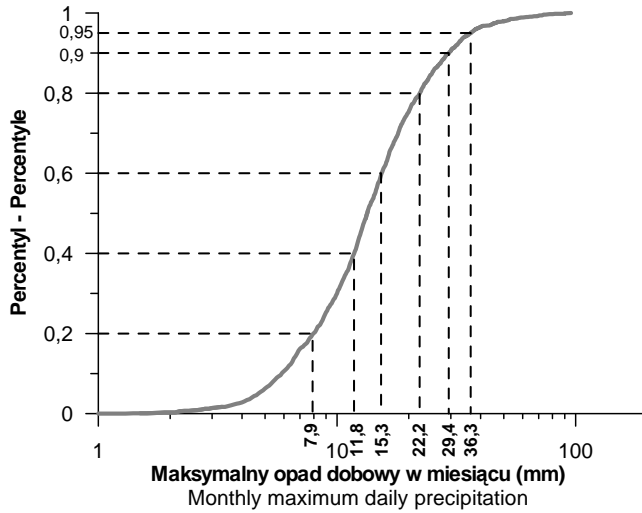
**2) utworzenie dla każdej lokalizacji danych pomiarowych wektora danych binarnych** (ryc. 10 i 11). Oryginalne, ciągłe dane pomiarowe przetwarza się do formy binarnej zgodnie ze wzorem [12]. Wartość MSDO zmierzona w Kamiennej Górze (5,5 mm) jest mniejsza od najniższego proggu (20 percentyl = 7,8 mm), dlatego po przekodowaniu uzyskujemy wektor składający się tylko z jedynek (ryc. 11). W Kobylnie zanotowano 17,3 mm, a zatem



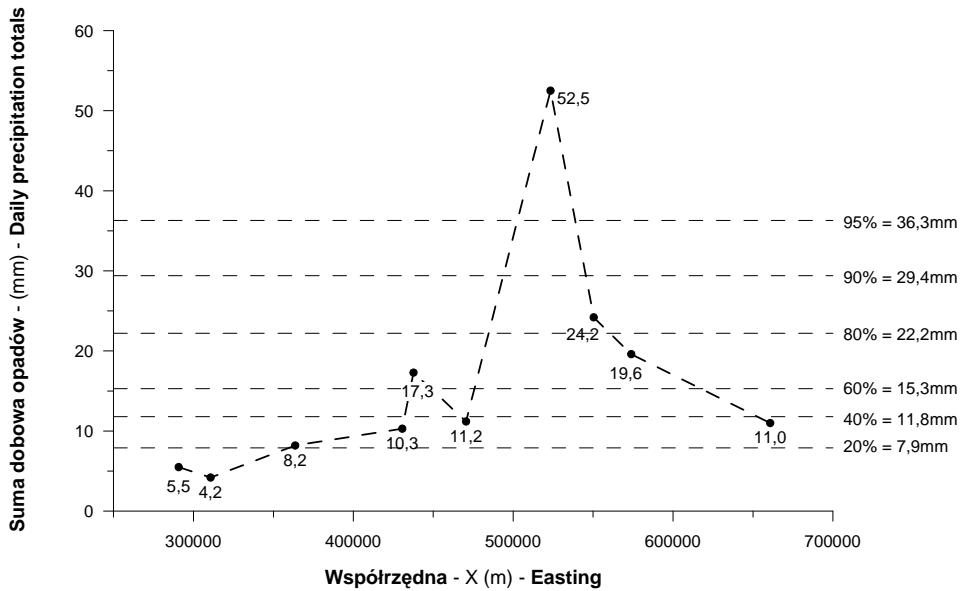
Ryc. 7. Profil posterunków opadowych (dane jednowymiarowe) funkcjonujących w maju 1980 roku wykorzystany do ilustracji metodyki geostatystycznej analizy danych kodowanych



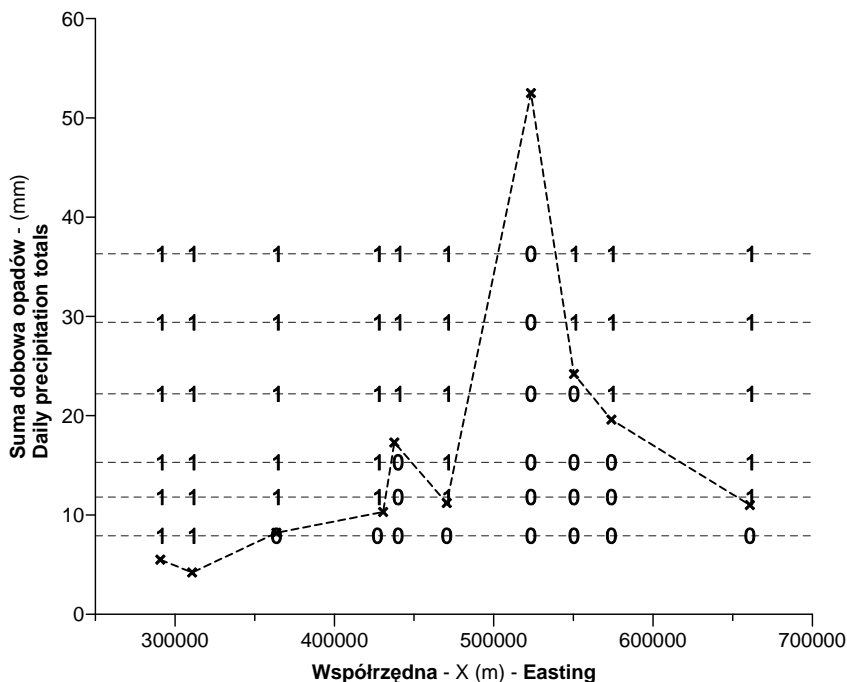
Ryc. 8. Maksymalne sumy dobowe opadów zarejestrowane w maju 1980 roku na posterunkach usytuowanych na profilu (ryc. 7)



Ryc. 9. Skumulowany rozkład prawdopodobieństwa (*cdf*) maksymalnych dobowych sum opadów zarejestrowanych w posterunkach opadowych na terenie Polski w maju 1980 roku. Na wykresie zaznaczono wysokości sum dobowych o prawdopodobieństwie przewyższenia 0,2, 0,4, 0,6, 0,8, 0,9 i 0,95 (percentyle 20, 40, ..., 95%)



Ryc. 10. Wartości progowe maksymalnych sum dobowych opadów w maju 1980 roku o prawdopodobieństwie 0,2, 0,4, 0,6, 0,8, 0,9 i 0,95 (ryc. 9, percentyle 20, 40, 60, 80, 90 i 95%) naniesione na dane profilowe (ryc. 7 i 8)

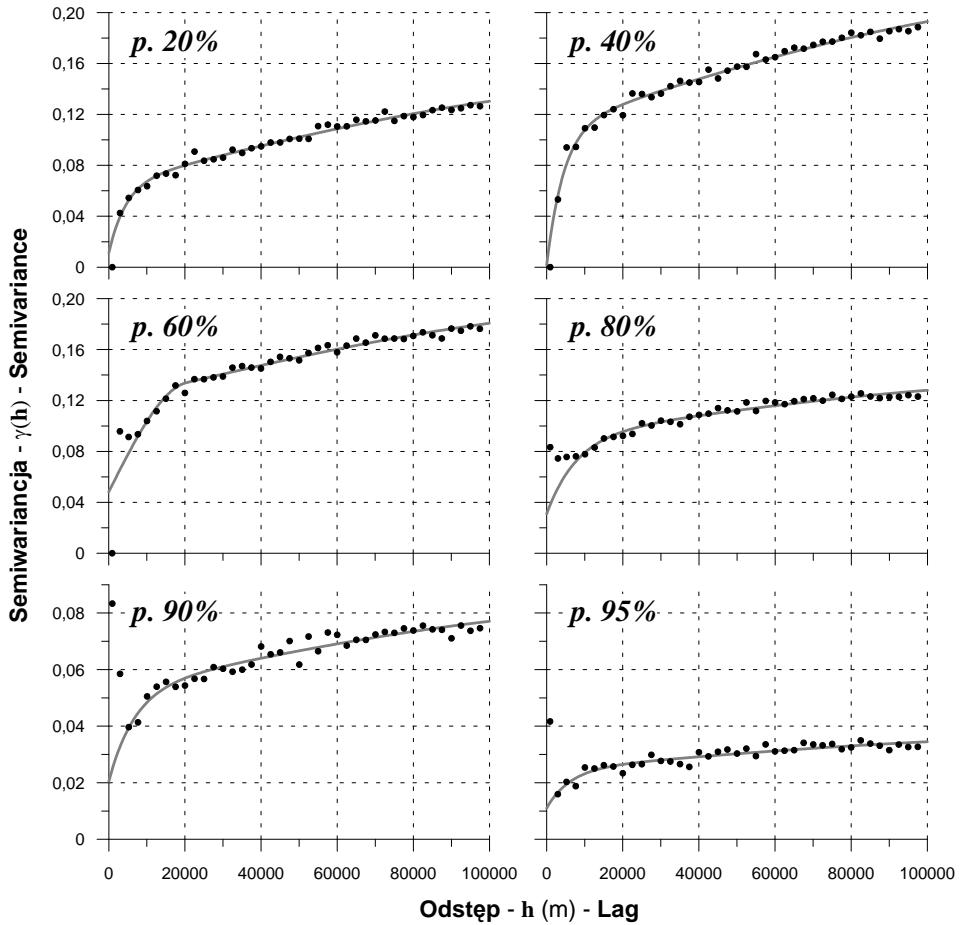


Ryc. 11. Maksymalne sumy dobowe opadów zarejestrowane w maju 1980 roku na analizowanym profilu (ryc. 7 i 8) przekodowane na wektory danych binarnych zgodnie ze wzorem [12] w zależności od przekroczenia wartości progowych wyznaczonych z globalnej krzywej skumulowanego rozkładu prawdopodobieństwa (percentyle 20, 40, 60, 80, 90 i 95%, ryc. 9)

więcej od pierwszych trzech wartości progowych (20% = 7,9 mm, 40% = 11,8 mm i 60% = 15,3 mm) i dlatego wektor wartości binarnych jest następujący: 0,0,0,1,1,1. Dla lokalizacji stacji Małusy Wielkie wektor ten składa się z samych zer, bo zmierzona tam w maju 1980 roku MSDO wynosząca 52,5 mm była większa od wszystkich wybranych wartości progowych.

**3) obliczanie i modelowanie miar struktury przestrzennej wartości kodowanych** (ryc. 12). Dla kolejnych progów obliczane są empiryczne miary struktury przestrzennej, najczęściej semiwariancje danych kodowanych. Do dyskretnych wartości semiwariancji dopasowywane są następnie funkcje matematyczne, zgodnie z regułami liniowego modelu regionalizacji (patrz podrozdz. III.2.5.2). Jeśli modele te mają być później używane do estymacji w węzłach siatki interpolacyjnej, warunkowych funkcji skumulowanego rozkładu (ang. *conditional cumulative distribution function - ccdf*) metodą krigingu danych kodowanych (patrz rozdz. XI.2), należy na etapie modelowania przestrzegać określonych zasad. Mają one na celu ograniczenie występowania





Ryc. 12. Semiwariogramy empiryczne i ich modele dla wartości kodowanych (percentyle 20, 40, 60, 80, 90 i 95% – ryc. 9) maksymalnych sum dobowych opadów na terenie Polski w maju 1980 roku

w estymowanych *ccdf* błędów relacji porządkowych<sup>13</sup> (ang. *order relations errors*, Deutsch, Journal 1998, Goovaerts 1997). W szczególności modele dla poszczególnych wartości progowych powinny być tworzone jako kombinacja tych samych modeli elementarnych, a ich parametry winny zmieniać się między kolejnymi progami w sposób stopniowy. W prezentowanym jednowymiarowym przykładzie (ryc. 7 i 8) semiwariancje empiryczne obliczono z kodowanych danych całego zbioru MSDO na obszarze Polski w maju 1980

<sup>13</sup> W dowolnej lokalizacji  $\mathbf{u}$  każde estymowane *posteriori* prawdopodobieństwo  $[F(\mathbf{u}; z_k | (n))]^*$  musi należeć do przedziału  $[0,1]$ , a seria  $K$  takich szacunków musi być niemalejącą funkcją wielkości wartości progowej  $z_k$ .

roku (ryc. 12). Do ich modelowania (por. rozdz. III.2.5.2) użyto za każdym razem identycznego zestawu trzech modeli elementarnych: nuggetowego, wykładniczego o zasięgu kilku do kilkunastu kilometrów i sferycznego o arbitralnie przyjętym zasięgu 150 km. Zastosowane funkcje były bardzo dobrze dopasowane do danych empirycznych oprócz pierwszych 1-2 odstępów niektórych progów (np. od 0,6 do 0,95), gdzie ze względu na małą liczbę pomiarów wartości semiwariogramu były chaotyczne. Uzyskane modele (ryc. 12) wyraźnie pokazują odmienność struktury przestrzennej dla MSDO o różnej wysokości. Względny udział<sup>14</sup> semiwariancji nuggetowej był najmniejszy przy niskich opadach (percentyl 40 i 20 - odpowiednio 0,007 i 0,07), a osiągał maksimum przy najwyższych progach (percentyl 90 i 95 - odpowiednio 0,234 i 0,1). Zasięg pierwszej struktury był największy przy percentylu 80 (23,5 km), nieco mniejszy przy opadach najwyższych (20,5 oraz 18,5 km), a najmniejszy (13,5 km) przy pierwszym i drugim proggu. Opady najniższe (20 i 40 percentyl) charakteryzowały się największym gradientem spadku podobieństwa wraz z odległością, najwyższe (95 percentyl) - najmniejszym.

## 2.5. Modelowanie struktury przestrzennej danych

W niniejszej pracy modelowanie struktury przestrzennej było zasadniczym czynnikiem decydującym o jakości wszystkich najważniejszych jego wyników<sup>15</sup>. Z drugiej strony, liczba modeli, jakie trzeba było opracować, zmuszała do szukania takiego rozwiązania, które umożliwiłoby uzyskanie dobrych wyników w rozsądnym czasie. Dlatego, problematykę tę przedstawiono w niniejszym podrozdziale bardzo szeroko.

### 2.5.1. Wprowadzenie do problematyki modelowania struktury przestrzennej

Wszystkie geostatystyczne algorytmy estymacji i symulacji przestrzennej wymagają podania w matematycznej (parametrycznej) formie modelu struktury przestrzennej analizowanej cechy (Goovaerts 1997, Deutsch, Journel 1998, Isaaks, Srivastava 1989). Nie mogą to być po prostu wartości takich empirycznych miar struktury przestrzennej, jak kowariogram, korelogram, czy semiwariogram. Przyczyn jest kilka.

---

<sup>14</sup> Porównania dokonuje się po standaryzacji polegającej na podzieleniu każdej wartości semiwariogramu empirycznego i modelowego przez wariancję danych kodowanych analizowanego zbioru.

<sup>15</sup> Wyniki modelowania zamieszczono na dołączonym dysku DVD (patrz załącznik XII.2).

W macierzach krigingowych dla każdego oczka siatki interpolacyjnej (patrz dodatek X.2) potrzebne są wartości kowariancji pomiędzy danymi pomiarowymi znajdującymi się w sąsiedztwie szukania (ang. *data covariances*) oraz kowariancji danych do lokalizacji estymowanej nieznannej (ang. *data-to-unknown covariances*). Oznacza to, że muszą one być możliwe do określenia dla dowolnej kombinacji odległości i kierunku. Wszystkie zaś wymienione wyżej miary empiryczne autokorelacji przestrzennej są dyskretne, czyli nieciągłe, obliczane bowiem jako średnie przedziałowe (odstępów odległości i sektory kierunków). W estymacjach potrzebne są też często wartości kowariancji dla odległości i kierunków, dla których nie ma danych empirycznych. Nie bywa jednak stosowane, wydawałoby się najprostsze, rozwiązanie, to jest interpolacja i ekstrapolacja kowariancji empirycznych. Składają się na to z kolei dwie przyczyny.

Po pierwsze, empiryczne miary struktury przestrzennej stanowią statystyki obliczane z próby i jako takie są tylko przybliżeniem relacji przestrzennych w całej populacji. Przy założeniu, że dysponuje się danymi niezawierającymi błędów, ich wiarygodność jest uzależniona z jednej strony od wielkości i reprezentatywności próby, a z drugiej – od skali zmienności analizowanej cechy (Webster, Oliver 2001). Zawsze dochodzi jednak trzeci element, jakim są nieuniknione błędy, zarówno pomiaru cechy, jak i określenia lokalizacji stanowiska, gdzie pomiar został dokonany. Zazwyczaj analizowane są próby stanowiące milionowe, miliardowe lub nawet mniejsze części całej populacji<sup>16</sup>, a zalecenie ich losowości często nie jest w pełni spełnione. Wszystkie wymienione czynniki, a także czułość stosowanych zazwyczaj miar struktury przestrzennej na naturalnie występujące wartości anomalne (por. dodatek X.5) powoduje, że ich wykresy są często chaotyczne i nie odzwierciedlają wiarygodnie stosunków istniejących w całej zbiorowości. Matematyczny model może przynajmniej zredukować znaczenie pierwszej z wyżej wymienionych wad, wygładzając chaotyczne fluktuacje danych empirycznych.

Druga przyczyna jest natury „wewnętrznej” i wiąże się z matematyczną formą estymatora krigingowego. Jak wspomniano poprzednio (por. rozdz. III.2.2), estymowane lub symulowane wartości są w geostatystyce traktowane jako zmienne losowe będące liniową kombinacją innych znanych zmiennych losowych. Wariancja zaś jakiegokolwiek liniowej kombinacji  $Y$  zmiennych losowych  $Z(\mathbf{u}_\alpha)$ ,  $\mathbf{u}_\alpha \in A$  jest wówczas liniową kombinacją wartości kowariancji owych zmiennych i musi być nieujemna (ang. *non-negative*):

---

<sup>16</sup> W niniejszym opracowaniu (por. rozdz. V) korzystano z wyników pomiarów w średnio 2485 lokalizacjach na powierzchni 319 114 km<sup>2</sup> ( $319,114 \times 10^9$  m<sup>2</sup>). Ponieważ powierzchnia zbiorcza standardowego deszczomierza czy pluwiografu sięga 200 cm<sup>2</sup>, to średnia sumaryczna powierzchnia próby MSDO, którą dysponowano wynosiła 49,6 m<sup>2</sup>. W stosunku do całej populacji stanowi to  $1,554 \times 10^{-10}$ , a zatem dziesięciomiliardową jej część.

$$\begin{aligned} \text{Var}\{Y\} &= \text{Var}\left\{\sum_{\alpha=1}^n \lambda_{\alpha} Z(\mathbf{u}_{\alpha})\right\} \\ &= \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C(\mathbf{u}_{\alpha} - \mathbf{u}_{\beta}) \geq 0 \end{aligned} \quad [16]$$

dla jakiejkolwiek z wybranych  $n$  lokalizacji  $\mathbf{u}_{\alpha} \in A$  i dla jakiejkolwiek wagi  $\lambda_{\alpha}$ . Spełnienie tego warunku jest możliwe tylko przy zastosowaniu takich funkcji kowariancji  $C(\mathbf{h})$ , nieparametrycznych czy parametrycznych, które są pozytywnie połowicznie określone (ang. *positive semidefinite*). Stosowanie interpolowanych/ekstrapolowanych wartości empirycznych miar struktury przestrzennej nigdy nie gwarantuje, że obliczenia estymacji/symulacji dadzą jakikolwiek wynik. Gwarancję taką można mieć jedynie przy zastosowaniu modelu matematycznego o takiej postaci, który jest z góry pozytywnie połowicznie określony (Deutsch, Journel 1998, Goovaerts 1997, Zawadzki 2005). Modele takie określa się jako dozwolone (ang. *permissible*).

### 2.5.2. Proste i złożone, dopuszczalne funkcje używane przy modelowaniu struktury przestrzennej

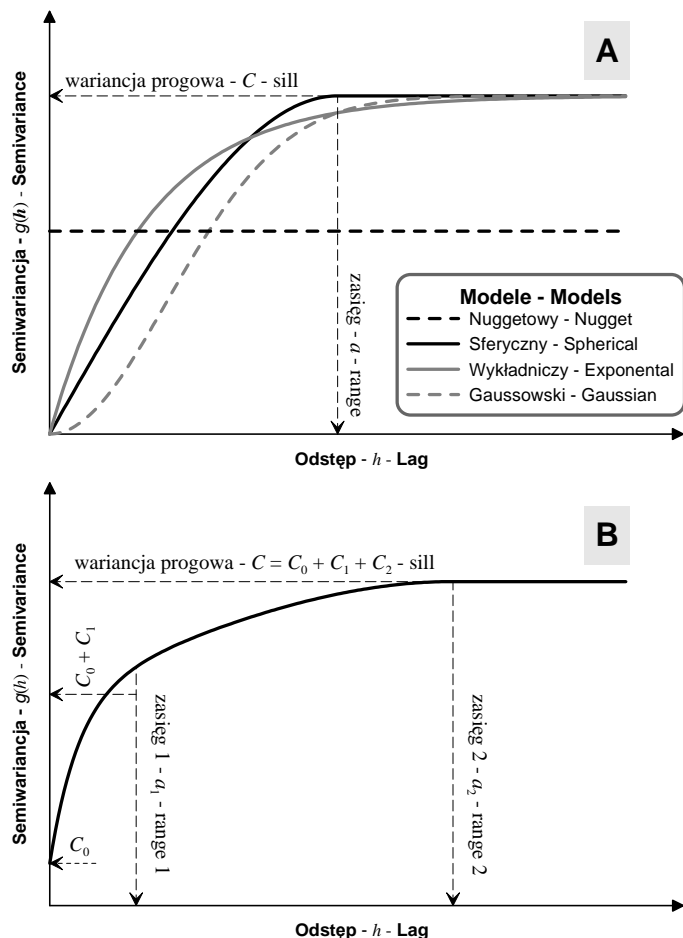
W literaturze geostatystycznej (Bleinès i in. 2007, Chilès, Delfiner 1999, Cressie 1993, Olea 1999, Webster, Oliver 2001) podawane jest w sumie kilkanaście funkcji spełniających warunek pozytywnej połowicznej określoności. Nie wszystkie jednak mogą być stosowane bez ograniczeń – na przykład dotyczących liczby wymiarów przestrzeni danych. W praktyce wykorzystywane jest jedynie kilka z nich, które zapewniają w większości przypadków bardzo dobre lub dobre odwzorowanie struktur przestrzennych spotykanych w środowisku. Decyzja o wykonywaniu w ramach całego projektu badawczego estymacji pola prawdopodobieństwa MSDO metodą krigingu wartości kodowanych (ang. *Indicator Kriging*, IK) i jego symulacji metodą *p-pola* (ang. *p-field*), przy wykorzystaniu programu IKSIM (Ying 2000), miała również konsekwencje dotyczące modelowania struktury przestrzennej (patrz dodatek X.2). We wspomnianym bowiem programie komputerowym dopuszczalne są jedynie cztery wymienione i opisane niżej oraz przedstawione na rycinie 13A dozwolone modele struktury przestrzennej danych.

- Model nuggetowy<sup>17</sup> (ang. *nugget effect model*):

$$g(h) = \begin{cases} 0, & \text{jeżeli } h = 0 \\ C_0, & \text{poza tym} \end{cases} \quad [17]$$

---

<sup>17</sup> Jak wspomniano, w polskiej literaturze (Zawadzki 2005, Namysłowska-Wilczyńska 2006) stosowany jest termin model „efektu samorodka”.



**Ryc. 13.** Podstawowe modele semiwariogramów wykorzystywane do opisu struktury przestrzennej maksymalnych sum dobowych opadów oraz estymacji i symulacji ich pola prawdopodobieństwa (A). Przykład złożonego (zagnieżdżonego) modelu semiwariogramu składającego się z trzech modeli podstawowych: nuggetowego, wykładniczego i sferycznego (B). Na wykresach zaznaczono podstawowe parametry modeli: wariancję progową ( $C$ , sill), wariancję składowych ( $C_x$ , partial sills), zasięg ( $a$ , range), zasięgi składowych ( $a_x$ , partial ranges)

- Model sferyczny (ang. *spherical*) o zasięgu  $a$ :

$$g(h) = C \cdot Sph\left(\frac{h}{a}\right) = \begin{cases} C \cdot \left[ 1,5 \cdot \frac{h}{a} - 0,5 \cdot \left(\frac{h}{a}\right)^3 \right], & \text{jeżeli } h \leq a \\ C, & \text{poza tym} \end{cases} \quad [18]$$

- Model wykładniczy (ang. *exponential*) o „praktycznym”<sup>18</sup> zasięgu  $a$ :

$$g(h) = C \cdot \text{Exp}\left(\frac{h}{a}\right) = C \cdot \left[1 - \exp\left(\frac{-3h}{a}\right)\right] \quad [19]$$

- Model gaussowski (ang. *gaussian*) o „praktycznym” zasięgu  $a$ :

$$g(h) = C \cdot \left[1 - \exp\left(-\frac{(3h)^2}{a^2}\right)\right] \quad [20]$$

gdzie:  $C_0$  oznacza wariancję nuggetową,  $C$  - wariancję progową (ang. *sill*),  $a$  - zasięg lub zasięg „praktyczny” (ang. *range*, *practical range*), zaś  $h$  - odległość (odstęp, ang. *lag*).

Zastosowanie modelu nuggetowego ([17], ryc. 13A) oznacza w praktyce stwierdzenie braku autokorelacji zjawiska. Wariancja procesu jest stała dla każdej odległości większej od zera. Ponieważ większość parametrów środowiskowych jest w jakiejś skali przestrzennej ciągła, konieczność użycia tego modelu oznacza zazwyczaj, że odstęp próbkowania był większy niż zasięg autokorelacji. Inną możliwością stwarza istnienie zjawiska w 100% losowego, nieciągłego, którego charakterystyki zmieniają gwałtownie przy przejściu z jednego punktu do drugiego. Obrazem symulacji bezwarunkowej z użyciem tego modelu jest, zgodnie z terminologią stosowaną w analizie serii czasowych, „biały szum”, podobny do obserwowanego na ekranie telewizora (lampy katodowej), do którego nie dociera żaden sygnał<sup>19</sup>.

Model sferyczny ([18], ryc. 13A) może być stosowany w 1, 2 i 3 wymiarach, i jest jednym z najczęściej stosowanych w geostatystyce do charakterystyki struktury przestrzennej. Daje on reprezentację cech ciągłych, które mają podobną rozciągłość, a ich zmienność ma charakter przeplatających się nieregularnych płatów z wysokimi i niskimi wartościami. Średnia średnica owych płatów jest reprezentowana przez zasięg modelu. Model sferyczny ma w początkowym odcinku charakter funkcji liniowej o nachyleniu  $3C/2a$ .

Podobnie często wykorzystywany jest model ujemnie wykładniczy ([19], ryc. 13A). Funkcja ta osiąga wariancję progową asymptotycznie i dlatego nie ma skończonego zasięgu. Zamiast niego podawany jest tak zwany zasięg „praktyczny” lub „efektywny”, zdefiniowany jako odległość, na jakiej model osiąga 95% wartości wariancji progowej. Model wykładniczy ma również liniowy charakter w fazie początkowej, ale o większym nachyleniu niż w przypadku sferycznego:  $C/a$ . Funkcja ta odgrywa bardzo ważną rolę teoretyczną. Stanowi bowiem istotę losowości w ujęciu przestrzennym. Jest to semiwariogram procesów autoregresyjnych pierwszego rzędu i procesów

<sup>18</sup> Pojęcie zasięgu praktycznego jest wyjaśnione w dalszej części tekstu.

<sup>19</sup> Model nuggetowy opisuje również pewne mikrostruktury regularne typu mozaiki (por. Chilès, Delfiner 1999, ryc. 15, s. 53).

Markowa. Można się spodziewać semiwariogramu wykładniczego wówczas, kiedy głównym źródłem zmienności cechy jest występowanie odmiennych typów systemów, a granice między typami występują losowo zgodnie z procesem Poissona. Przykładem może być sytuacja, kiedy zmienność przestrzenna pewnej cechy gleb, na przykład odczynu, jest uwarunkowana głównie różnicami pomiędzy typami gleb. Inaczej mówiąc, jest to semiwariogram cech ciągłych, których struktury mają zasięg losowy.

Ostatni z wykorzystywanych, dopuszczalnych, modeli – gaussowski ([20], ryc. 13A) – również osiąga poziom wariancji progowej asymptotycznie, dlatego, jak w poprzednim przypadku, zasięg praktyczny definiowany jest jako odległość, dla której wartość semiwariancji wyliczona z modelu osiąga 95% wariancji progowej. Model gaussowski odróżnia się od dwóch wymienionych poprzednio przede wszystkim parabolicznym kształtem w początkowym odcinku. Oznacza to, że „dotyka” on osi z zerowym nachyleniem. Stanowi to granicę zmienności losowej, przy której w rzeczywistości istnieje wartość stała i podwójnie różniczkowalna (Wackernagel 2003). Model ten ze względu na swój deterministyczny charakter daje w wielu przypadkach nierealistyczne wyniki estymacji i zazwyczaj nie może być stosowany samodzielnie. Jego użycie jest czasami uzasadnione przy analizie przestrzennej i prognozowaniu parametrów o bardzo regularnej i łagodnej zmienności przestrzennej, na przykład poziomu wód gruntowych w obszarach o mało zróżnicowanej rzeźbie i jednolitej budowie geologicznej.

W wielu sytuacjach, w celu dokładnego odwzorowania kształtu semiwariogramu empirycznego konieczne jest połączenie dwóch lub większej liczby modeli podstawowych  $g(\mathbf{h})$ . Problem tkwi w tym, że nie wszystkie kombinacje dopuszczalnych modeli dają w efekcie funkcję dopuszczalną, to znaczy z nieujemną wariancją. Najprostszym sposobem utworzenia modelu dopuszczalnego jest stworzenie najpierw funkcji losowej. Semiwariogram takiej funkcji jest z definicji dopuszczalny. Praktycznie rzecz biorąc, konieczne jest spełnienie dwóch warunków tak zwanego liniowego modelu regionalizacji<sup>20</sup> (ang. *linear regionalisation model*):

1) wszystkie użyte w modelu złożonym podstawowe funkcje  $g_l(\mathbf{h})$  muszą być dopuszczalne,

2) wariancja progowa  $b^l$  każdego podstawowego modelu semiwariogramu musi być dodatnia [21], a wówczas:

$$\begin{aligned} \gamma(\mathbf{h}) &= \frac{1}{2} E \left\{ [Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2 \right\} \\ &= \sum_{l=0}^L b^l g_l(\mathbf{h}) \quad \text{z} \quad b^l = (a^l)^2 \geq 0 \end{aligned} \quad [21]$$

<sup>20</sup> Oczywiście, pojęcie to nie ma nic wspólnego z regionalizacją w sensie nauk geograficznych.

Model złożony  $\gamma(\mathbf{h})$  w takiej sytuacji jest wyrażony jako pozytywna liniowa kombinacja podstawowych modeli semiwariogramów  $g_i(\mathbf{h})$ . W literaturze przedmiotu (Deutsch, Journel 1998, Goovaerts 1997) popularna jest jego także alternatywna nazwa: *nested model*, czyli model zagnieżdżony. Rycina 13B prezentuje przykład takiego modelu, składającego się z trzech modeli podstawowych: nuggetowego, wykładniczego i sferycznego. Oprócz kilku wyjątków wszystkie z około 4800 opracowanych w niniejszej pracy modeli struktury przestrzennej miały charakter złożony, zawierając minimum dwa elementy (w tym nuggetowy), a maksimum – pięć. **Złożona struktura przestrzenna analizowanej cechy (w tym wypadku MSDO), świadczy, że jest ona efektem działania kilku procesów operujących w różnych skalach.**

### 2.5.3. Specyfika matematycznego modelowania struktury przestrzennej

Modelowanie matematyczne struktury przestrzennej nie jest, jakby się wydawało, nawet przy aktualnym poziomie mocy obliczeniowej komputerów osobistych, zagadnieniem łatwym. Wręcz przeciwnie, automatyczne algorytmy zdają egzamin jedynie przy izotropowych, i to raczej najprostszych, przypadkach. Próby stworzenia takich uniwersalnych procedur podejmowane są od ponad 30 lat bez większego powodzenia (Webster, Oliver 2001). W modułach modelowania struktury przestrzennej w większości dostępnych programów geostatystycznych dokonuje się tego na drodze manualnej lub półautomatycznej, przy czym i w tym drugim przypadku operator ma całkowitą kontrolę nad przebiegiem obliczeń, łącznie z możliwością całkowitego wyłączenia „automatyki” i/albo „ręcznej” modyfikacji wyników. Tam, gdzie stosuje się wariant w 100% manualny, operator obserwując na ekranie wykres semiwariogramu empirycznego (albo innej miary struktury przestrzennej), wybiera podstawowe składowe modelu: liczbę, typ i następstwo struktur elementarnych, a następnie metodą prób i błędów optymalizuje ich parametry: zasięg, wariancje cząstkowe, kierunek i stopień anizotropii. W procedurach półautomatycznych drugi wymieniony wyżej etap, to jest optymalizacja parametrów, odbywa się w mniejszym lub większym stopniu automatycznie. Jeśli operator nie jest zadowolony z końcowego wyniku, może dokonać jego modyfikacji, cały czas obserwując efekt swoich działań na ekranowym wykresie. Ten element procedury geostatystycznej jest bardzo często krytykowany ze względu na subiektywizm. Bywa jednak, że może to doprowadzić do radykalnego polepszenia wyników estymacji lub symulacji. Związane jest to bowiem z możliwością uwzględnienia wiedzy *a priori*, wiedzy eksperta, o naturze zmienności przestrzennej zjawiska,



które wygenerowało obserwowany rozkład analizowanej cechy (Goovaerts 1997). Dane pomiarowe ze względu na małą próbę, czy błędy pomiarowe, mogą tych relacji nie wykazywać. Kiedy na przykład zadaniem jest estymacja skażenia gleb wokół emitora, jakim może być komin elektrowni ciepłej, wiemy, że musi ona wykazywać anizotropię o kierunku i rozmiarach uzależnionych od lokalnego reżimu wiatrów. Wiedzę tę możemy w trakcie modelowania wykorzystać, „wymuszając” uwzględnienie anizotropii.

Oprócz problemów merytorycznych wadą ręcznego i półautomatycznego modelowania struktury przestrzennej jest także jego czasochłonność, co było szczególnie istotne w kontekście niniejszej pracy. Opracowanie jednego złożonego, anizotropowego modelu może trwać wiele minut. W trakcie prac ze zbiorami danych miesięcznych i rocznych MSDO wykonano ich blisko 4800.

#### 2.5.4. Kryterium jakości dopasowania modelu

Zarówno przy ręcznym, jak i automatycznym modelowaniu struktury przestrzennej potrzebne jest obiektywne kryterium jakości dopasowania modelu do danych empirycznych. W pierwszym przypadku ma ono charakter pomocniczy. Operator może, ale nie musi, opierając się na obliczanych po każdej wprowadzonej przez niego zmianie w parametrach modelu, wartościach owego kryterium, dążyć do optymalizacji wyniku. Optymalizacja automatyczna musi być dokonywana w odniesieniu do precyzyjnie zdefiniowanego jej celu, którym zazwyczaj jest minimalizacja albo maksymalizacja wartości jakiejś funkcji. Konieczne jest również podanie warunku zakończenia obliczeń – zazwyczaj dokonywanych metodą kolejnych przybliżeń (iteracyjnie). Najczęściej w geostatystyce stosowanym kryterium dopasowania modelu do danych empirycznych jest Ważona Suma Kwadratów (ang. *Weighted Sum of Squares*, WSS, Cressie 1985, 1991, Jian i in. 1996, Pardo-Igúzquiza 1999) różnic pomiędzy eksperymentalnymi  $\hat{\gamma}(\mathbf{h}_k)$  a modelowanymi  $\gamma(\mathbf{h}_k)$  wartościami semiwariogramu:

$$WSS = \sum_{k=1}^K \omega(\mathbf{h}_k) \cdot [\hat{\gamma}(\mathbf{h}_k) - \gamma(\mathbf{h}_k)]^2 \quad [22]$$

Waga  $\omega(\mathbf{h}_k)$  przypisana do każdego odstępów  $\mathbf{h}_k$  jest zazwyczaj proporcjonalna do liczby  $N(\mathbf{h}_k)$  par danych, które są uwzględniane w obliczeniu wartości semiwariogramu empirycznego  $\hat{\gamma}(\mathbf{h}_k)$ . Podstawą takiego rozwią-

zania jest założenie, że wiarygodność semiwariogramu empirycznego wzrasta wraz z wielkością próby. Nie jest to jednakże jedyny wariant wagi w kryterium WSS. Inny, równie często stosowany, przywiązuje większe znaczenie do semiwariogramów obliczanych dla pierwszych odstępów<sup>21</sup>, poprzez podzielenie liczby par danych przez podniesioną do kwadratu wartość mode-

lu:  $\frac{N(\mathbf{h}_k)}{[\gamma(\mathbf{h}_k)]^2}$ . W innych podejściach wiarygodność statystyczną wartości

semiwariogramu empirycznego ocenia się nie poprzez liczbę par danych, która posłużyła do jego obliczenia, ale poprzez zróżnicowanie indywidualnych wartości różnic obliczonych dla każdej pary: małe zróżnicowanie – duża wiarygodność. Waga WSS jest wówczas odwrotnie proporcjonalna do odchylenia standardowego indywidualnych wartości różnic. Na koniec omawiania tego zagadnienia należy wspomnieć, że metoda WSS, jakkolwiek najbardziej popularna, nie jest jedynym rozwiązaniem problemu automatycznego dopasowania parametrów modeli struktury przestrzennej. Stosowane są również algorytmy tzw. maksymalnej wiarygodności (ang. *maximum likelihood, ML*) lub ograniczonej ML (ang. *restricted ML, REML*), gdzie model tworzony jest bezpośrednio na podstawie surowych wartości różnic (Dietrich, Osborne 1991, Pardo-Igúzquiza 1997, 1998, Zimmerman 1989). Ponieważ jednak bazują one na założeniu rozkładu normalnego, ich oszacowania są często obciążone. Obliczenia wykonywane metodami ML i REML są także stosunkowo wolne przy dużych próbach (Olea 1999).

### 2.5.5. Optymalizacja modelu struktury przestrzennej

Z półautomatycznym i ręcznym konstruowaniem złożonego (zagnieżdżonego) modelu struktury przestrzennej wiąże się jeszcze jeden istotny problem – niepewności co do wyboru optymalnej liczby i kombinacji funkcji podstawowych (elementarnych). Z jednej strony model powinien być jak najlepiej dopasowany do danych eksperymentalnych, z drugiej zaś wiadomo, że ich niewielkie fluktuacje mogą być zupełnie przypadkowe. Zagadnienie to można rozpatrywać w dwóch kontekstach. Pierwszy z nich ma charakter optymalizacyjny drugi – praktyczny.

Z praktycznego punktu widzenia najlepszym modelem jest nie ten najlepiej dopasowany do danych obserwacyjnych, ale dający najbardziej dokładną prognozę. Budując model rzadko dysponuje się taką liczbą danych,

---

<sup>21</sup> Jakość estymacji zależy głównie od poprawnego określenia „kształtu” modelu u jego początku, natomiast semiwariancje empiryczne pierwszych odstępów obliczane są zazwyczaj na podstawie znacznie mniejszej ilości par punktów niż dalsze.

która umożliwia dokonanie niezależnej jego walidacji – porównania rzeczywistych wartości analizowanej cechy z prognozowanymi na jego podstawie. Dlatego, zazwyczaj w geostatystyce stosuje się uproszczoną metodę testowania jakości modelu, zwaną kroswalidacją (Davis 1987, Goovaerts 1997, Webster, Oliver 2001). Polega ona na wykonywaniu sekwencyjnie  $n$  estymacji dla każdej lokalizacji z posiadanego zbioru danych, z wyłączeniem każdorazowo w trakcie obliczeń zmierzonej w tym miejscu wartości cechy. Wykonany szacunek jest oparty na pozostałych w puli danych ( $n - 1$ ) i modelu opracowanym na podstawie wszystkich  $n$  danych. W efekcie, dla każdego z  $n$  punktów pomiarowych dokonuje się porównania rzeczywiście zmierzonych wartości analizowanej cechy z prognozą. Używając różnych syntetycznych miar jakości estymacji, jak na przykład średni błąd ( $ME$ ), pierwiastek średniego błędu kwadratowego ( $RMSE$ ), czy współczynnik korelacji ( $r$ ), wykonuje się porównania alternatywnych modeli i wybiera najlepszy. Należy jednakże podkreślić, że kroswalidacja nie jest techniką w pełni obiektywną i dającą całkowicie wiarygodne wyniki. Jej słabość polega na używaniu na wszystkich etapach obliczeń (budowy modelu, estymacji/walidacji) tych samych danych, a oceny błędów dotyczą tylko lokalizacji pomiarowych, a nie najbardziej interesujących miejsc, w których pomiarów nie wykonano.

Optymalizacyjne podejście do zagadnienia budowy złożonego modelu semiwariancji opiera się idei, że musi istnieć równowaga pomiędzy prostotą modelu, a więc także łatwością wykonywanych na jego podstawie estymacji, a jakością jego dopasowania do danych eksperymentalnych. Polepszenie modelu poprzez minimalizację kryterium WSS może się dokonywać praktycznie w nieskończoność, jeśli zwiększać się będzie liczba funkcji elementarnych. Jednakże, kolejne komplikowanie jego postaci skutkuje coraz mniejszym przyrostem jakości dopasowania, dlatego ważne byłoby ustalenie kryterium, które umożliwiłoby w obiektywny i powtarzalny sposób zachowanie proporcji między dążeniem do prostoty modelu a wiernością odwzorowania wyników pomiarów. Webster i McBratney (1989) zaproponowali zastosowanie do tego celu kryterium informacyjnego Akaike (ang. *Akaike Information Criterion, AIC*, [23]):

$$\widehat{AIC} = \left\{ n \ln \left( \frac{2\pi}{n} \right) + n + 2 \right\} + n \ln R + 2p \quad [23]$$

gdzie:  $n$  jest liczbą punktów na wariogramie,  $p$  ilością parametrów modelu, a  $R$  stanowi średnią podniesionych do kwadratu różnic pomiędzy wartościami eksperymentalnymi a modelem. Do dalszego etapu procedury geostatystycznej wybiera się ten model, dla którego  $\widehat{AIC}$  jest najmniejsze. Fragment wzoru [23] znajdujący się w nawiasie jest stały dla każdego konkretnego semiwariogramu, dlatego można go uprościć do postaci [24]:

$$\widehat{AIC} = n \ln R + 2p \quad [24]$$

Minimalizacja kwadratów odchyłeń (WSS) zmniejsza wartość  $R$ , lecz jeśli dalsze jego obniżanie dokonuje się jedynie poprzez zwiększanie  $p$ , to w pewnym momencie spadek  $\widehat{AIC}$  zostaje zatrzymany.

Zastosowane oprogramowanie (patrz dalej), ale przede wszystkim nie-spotykana w typowych opracowaniach geostatystycznych liczba modeli, które trzeba było przygotować, uniemożliwiały rutynowe wykonywanie oceny ich optymalności. Przy czym, rzadko występowała taka potrzeba, szczególnie w odniesieniu do semiwariogramów danych znormalizowanych (patrz dalej w tym podrozdziale). W większości przypadków układ punktów był bowiem bardzo regularny, a odstępy pomiędzy kolejnymi załamaniami krzywej na tyle duże, że nie było wątpliwości, które z podstawowych funkcji należy użyć. Więcej problemów było z niektórymi semiwariogramami danych kodowanych, co szczegółowo opisano dalej w poniższym podrozdziale. Trzeba wyraźnie podkreślić, że jakkolwiek w ocenie autora znaczenie optymalizacji przy modelowaniu struktury przestrzennej w kontekście analizowanych w niniejszej pracy zbiorów danych nie jest duże, to jednak ten problem istnieje i nie został rozwiązany. Będzie to tematem osobnego opracowania.

Przedstawiony powyżej obraz modelowania struktury przestrzennej był paradygmatem geostatystyki przez ostatnie 40 lat. W każdym dostępnym oprogramowaniu realizującym funkcje geostatystycznej estymacji, symulacji bądź optymalizacji wymagane jest podanie przez operatora modelu struktury przestrzennej w postaci parametrycznej. Prawdopodobnie ulegnie to w najbliższym czasie zmianie. Pojawiła się bowiem idea pominięcia tradycyjnego sposobu budowy tego modelu – nieparametryczna alternatywa umożliwiająca automatyzację i obiektywizację procedury (Yao, Journal 1998). Sprowadza się to do transformacji eksperymentalnych map korelacji (lub kroskorelacji) do map gęstości spektralnej przy użyciu szybkiej transformaty Fouriera (FFT). Owe mapy gęstości spektralnej są następnie wygładzane przy zastosowaniu ograniczeń dodatniości i sumowania do jedności. Przeprowadzana następnie transformacja zwrotna przez odwrotność FFT daje w efekcie dozwolone, pozytywnie połowicznie określone mapy korelacji. Dzięki tej metodzie „prawidłowe” i praktycznie użyteczne mapy korelacji uzyskuje się automatycznie, bez konieczności analitycznego tworzenia modelu struktury. Potrzebne do obliczeń estymacji/symulacji wartości kowariancji dla dowolnej odległości i dowolnego kierunku uzyskiwane są z tych map przez interpolację lub ekstrapolację. Trudność weryfikacji w przestrzeni danych surowych warunku połowicznie pozytywnej określoności jest usunięta, ponieważ w domenie częstości jedynym wymogiem jest pozytywność uzyskanej funkcji gęstości i jej sumowanie do jedności.

### 2.5.6. Modelowanie struktury przestrzennej MSDO w programie ISATIS

Modelowanie struktury przestrzennej całych<sup>22</sup> analizowanych zbiorów miesięcznych i rocznych MSDO wykonywano w programie ISATIS (Bleinès i in. 2007), metodą półautomatyczną. Zdecydowano się na przeprowadzenie jedynie analizy izotropowej, co bardzo ją ułatwiło i przyspieszyło (por. dodatek X.4). Automatyczna procedura dopasowania wariancji cząstkowych w programie ISATIS zmierza do minimalizacji odległości pomiędzy wartością semiwariogramu empirycznego dla danego odstępów a odpowiadającą mu wartością modelu (za pomocą omówionego poprzednio kryterium WSS). Minimalizacja ta jest przeprowadzana przy uwzględnieniu jednej z czterech możliwych kombinacji wag:

- 1) każda wartość semiwariogramu niezależnie od odstępów i kierunku jest traktowana identycznie (bez wag);
- 2) waga dla każdego odstępów danego kierunku jest proporcjonalna do liczby par punktów wszystkich odstępów tego kierunku;
- 3) waga każdego odstępów danego kierunku jest wprost proporcjonalna do liczby par i odwrotnie proporcjonalna do średniej odległości tego odstępów;
- 4) waga każdego odstępów danego kierunku jest odwrotnie proporcjonalna do liczby odstępów na tym kierunku.

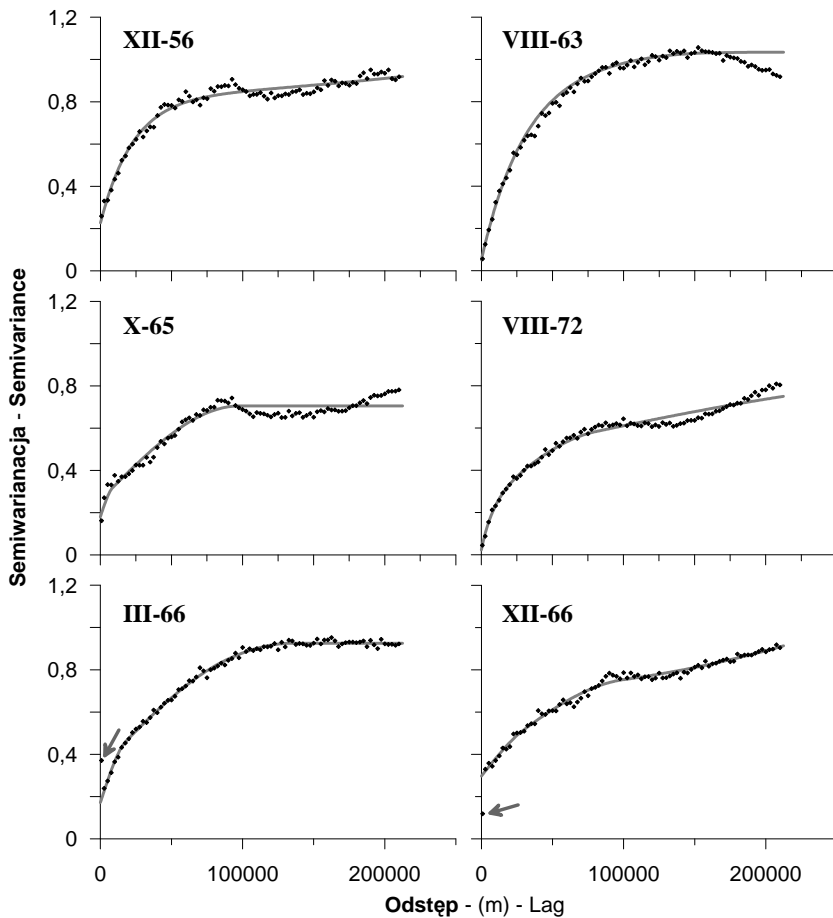
Szczegóły algorytmu optymalizacji wariancji cząstkowych poszczególnych składowych modelu przedstawione są w cytowanej publikacji Bleinèsa i in. (2007). W trakcie całego przebiegu prac modelowania semiwariogramów, zarówno danych znormalizowanych jak i kodowanych, używano trzeciego wariantu ważenia.

### 2.5.7. Modelowanie struktury przestrzennej danych znormalizowanych

Tworzenie modeli danych znormalizowanych było stosunkowo proste. Zazwyczaj po kilku próbach z wyborem typu struktury i jej zasięgu uzyskiwano bardzo zadowalający efekt. Nie było żadnej potrzeby ingerencji w automatyczną część całej procedury. Czasami jednak nie można było, używając dopuszczalnych typów struktur, uzyskać optymalnego dopasowania. Rozbieżności zazwyczaj jednak nie były duże i dotyczyły wyłącznie miesięcznych danych MSDO. Najczęściej spotykane sytuacje przedstawiono na rycinie 14.

---

<sup>22</sup> Oprócz tego wykonano również analizę struktury przestrzennej 208 indywidualnych przypadków rocznych MSDO (roczne MSDO, które miały miejsce tego samego dnia). Uzyskane wyniki zostaną opublikowane osobno.



Ryc. 14. Przykłady problemów przy modelowaniu struktury przestrzennej danych znormalizowanych

Pierwsza z nich (na ryc. 14 przykład z grudnia 1956), kiedy wartości semiwariogramu wykazywały chaotyczne fluktuacje na krótkich dystansach, zdarzała się najrzadziej i w zasadzie ograniczyła do kilku miesięcy w roku 1956 i na początku 1957. Był to okres z najmniejszą liczbą stanowisk pluwiometrycznych i jednocześnie z najszybszym tempem ich przyrostu. Oznacza to, że wielu obserwatorów dopiero uczyło się wykonywać pomiary i chyba to jest przyczyną opisywanych anomalii.

Częściej, bo kilkanaście razy, zanotowano sytuacje ilustrowane przykładami z sierpnia roku 1963, października 1965 i znów sierpnia, tym razem roku 1972 (ryc. 14). Polegały one na występowaniu długodystansowych, raczej regularnych fluktuacji, czasem wyraźnie cyklicznych, zachodzących po

osiągnięciu przez semiwariogram poziomu plateau, które można było interpretować jako semiwariancję progową. Świadczyłoby to o występowaniu stosunkowo mało znaczących (o niewielkiej amplitudzie) długodystansowych struktur opadów (o skali 50, 100 i więcej kilometrów), na które „nałożone” były silnie zmienne opady „lokalne”. Jeśli posiadane *a priori* informacje wskazują na autentyczną cykliczność zjawiska przestrzennego, tego typu struktura empiryczna jest przedstawiana za pomocą odpowiednich modeli zwanych „hole effect”, których elementem jest funkcja *cosinus* (Deutsch, Journal 1998, Olea 1999, Webster, Oliver 2001). W odniesieniu do analizowanych w niniejszej pracy zbiorów danych takich informacji nie było, a stosowane do estymacji i symulacji oprogramowanie nie dopuszczało użycia modeli typu „hole effect”. Dlatego, opisane wyżej fluktuacje były w modelowaniu pomijane. Ich znaczenie związane jest tylko z pytaniem o przypuszczalną genezę, ponieważ stosowanie modeli, w których ich obecności nie uwzględniono, nie wpływa w istotny sposób na jakość uzyskiwanych estymacji i symulacji pola prawdopodobieństwa MSDO.

Dwa ostatnie przedstawione na rycinie 14 przykłady (z marca i grudnia 1966) ilustrują siedem przypadków, kiedy wartości semiwariancji dla tak zwanego „zerowego” odstępu znacznie odbiegają<sup>23</sup> od tendencji zmian autokorelacji wykazywanych przez semiwariancję obliczone dla kolejnych, dalszych odstępów. Odstęp „zerowy”, zwany także „połówkowym”, obejmuje wszystkie pary punktów pomiarowych, które są od siebie odległe nie więcej niż 1250 m. W rozdziale V.2 omówiono szeroko problem dokładności określenia położenia stanowisk pluwiometrycznych, z których pochodzą analizowane dane MSDO. Lokalizacja ich była znana z dokładnością do 1 minuty kątowej, co oznacza że w „połówkowym” przedziale znalazły się zarówno stanowiska odległe dokładnie o 1 minutę długości geograficznej, ale tylko położone w północnej Polsce oraz tak zwane „duplikaty” (których z kolei było więcej w Polsce południowej). Te drugie miały takie same współrzędne geograficzne (por. rozdz. V.2). Po przeliczeniu ich do układu współrzędnych płaskich GUGIK 92/19, w celu uniknięcia obecności w zbiorach danych punktów o tej samej lokalizacji, zmieniano losowo wartości współrzędnych prostokątnych tak, aby mieściły się w obszarze znajdującym się w promieniu 0,5 km od lokalizacji „wyjściowej”. Par punktów w odstępie „połówkowym” było od ponad siedemdziesięciu w roku 1956 do dziewięciu w roku 1980. Oznacza to znacznie mniejszą wiarygodność statystyczną wyliczonej z nich wartości semiwariancji w porównaniu z odstępami następnymi, gdzie tych par było od kilkuset do kilku tysięcy. Wystąpienie pojedynczej anomalnej wartości opadu mogło w takiej sytuacji znacząco wpłynąć na wynik obliczenia. W dodatku X.5 omówiono procedurę masko-

---

<sup>23</sup> Zaznaczone na rycinie 14, dla zwrócenia uwagi, strzałkami.

wania w trakcie obliczeń semiwariancji danych tzw. „odstających”<sup>24</sup>. Jest ona jednakże obciążona pewną dozą subiektywizmu, co oznacza możliwość pominięcia niektórych przypadków. Podsumowując ów długi i wielowątkowy wywód można napisać, że przedstawione dwa ostatnie przykłady (ryc. 14) są związane z niedokładnościami określenia lokalizacji stanowisk pomiarowych, które były szczególnie istotne przy ich niewielkiej od siebie odległości, oraz ze specyfiką procedury maskowania danych anomalnych. Również w tym wypadku różnice między modelem a danymi empirycznymi nie są istotne. Odstająca wartość semiwariancji odstępu „połówkowego” była ignorowana, a wówczas model idealnie „pasował” do danych kolejnych odstępów. W sumie, wszystkie omówione wyżej „trudne” przypadki stanowiły w trakcie modelowania struktury przestrzennej zbiorów znormalizowanych danych MSDO tylko około 10% całości.

### 2.5.8. Modelowanie struktury przestrzennej danych kodowanych

Procedurę modelowania semiwariogramów danych kodowanych wykonywano analogicznie. Ze względu na 13 wartości progowych (percentyle 1, 5, 10, 20, ..., 90, 95, 99) dla każdego z 325 zbiorów danych MSDO = 4225 modeli był to jeden z najbardziej pracochłonnych i czasochłonnych etapów niniejszej pracy.

Z jednej strony procedura ta była prostsza niż omówiona wyżej. Mniejsza liczba uwzględnianych odstępów i krótszy w związku z tym zasięg semiwariogramu powodował, że zazwyczaj dobrze dopasowany model składał się jedynie z trzech struktur, w tym nuggetowej. Dalej, semiwariogramy empiryczne dla „sąsiadujących” ze sobą wartości progowych, szczególnie w środkowym przedziale zakresu rozkładu, zazwyczaj różniły się od siebie niewiele, co również ułatwiało szybkie opracowanie optymalnego modelu.

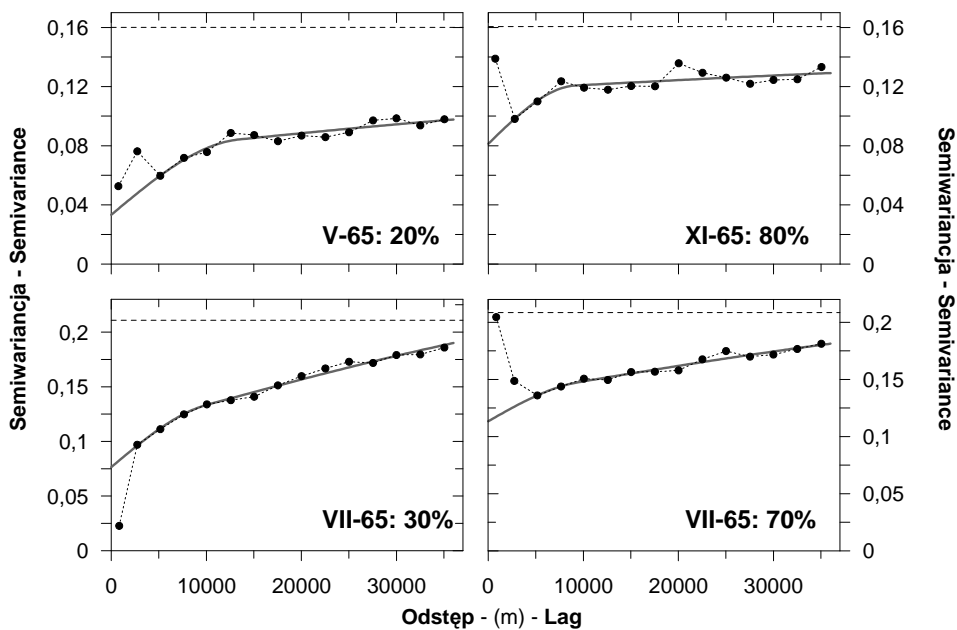
W trakcie tworzenia modeli struktury przestrzennej danych kodowanych napotymano jednak pewne problemy. Można je pogrupować w trzy klasy. Przykłady takich przypadków przedstawiono na rycinach 15-17.

Pierwszy, często spotykany problem, występujący przy wszystkich wartościach progowych oraz wyraźnie częściej w miesiącach zimowych (I - III), stanowią anomalne układy wartości semiwariancji dla pierwszego lub dwóch pierwszych odstępów (ryc. 15, tab. 1). Odbiegają one wyraźnie od konsekwentnego przebiegu zmian struktury przestrzennej widocznego dla kolejnych, dalszych odstępów. Przyczynę tego zjawiska omawiano już poprzednio w niniejszym podrozdziale. Wynika ono z jednej strony z relatywnie

---

<sup>24</sup> Dane te (patrz podrozdz. V.2) to albo naturalnie występujące anomalie - głównie opady orograficzne, albo błędne pomiary.





Ryc. 15. Przykłady semiwariogramów empirycznych danych kodowanych z odstającymi wartościami pierwszych odstępów i ich modele. Linią przerywaną zaznaczono poziom wariancji danych

Tabela 1. Przybliżony procentowy udział semiwariogramów kodowanych z anomalnymi wartościami dla pierwszych odstępów w zależności od miesiąca i wartości progowej (percentyla, P)

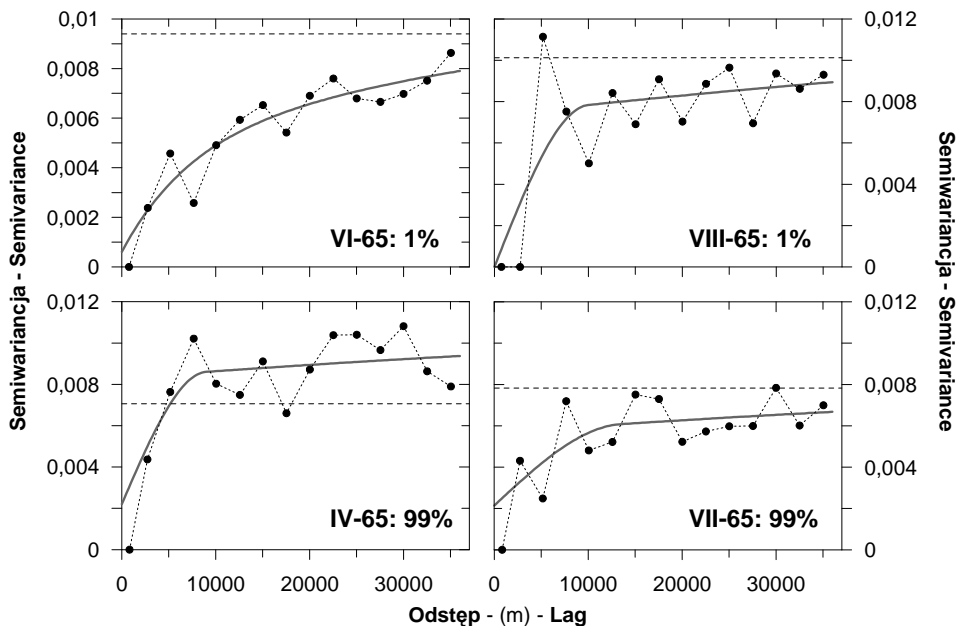
P	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	R. / Y.	$\bar{x}$
1	12	16	8	12	16	28	20	16	24	20	12	12	0	16,3
5	28	44	32	24	36	32	40	32	24	28	36	40	28	33,0
10	52	36	68	48	36	28	44	20	36	36	44	48	40	41,3
20	56	48	64	36	28	28	48	44	48	32	36	48	36	43,0
30	48	52	68	48	48	36	20	36	36	40	44	40	52	43,0
40	64	44	40	48	44	40	36	28	40	36	40	32	36	41,0
50	68	52	36	36	72	48	52	52	36	40	44	60	32	49,7
60	64	36	56	56	48	52	36	36	36	44	60	40	40	47,0
70	56	64	52	56	40	52	44	36	56	48	60	52	40	51,3
80	44	72	52	72	48	60	44	72	52	48	48	52	40	55,3
90	48	40	44	28	32	44	44	48	52	56	36	36	36	42,3
95	12	20	20	20	20	20	36	32	24	32	12	24	12	22,7
99	12	0	4	8	8	8	4	0	8	12	4	4	8	6,0
$\bar{x}$	43,4	40,3	41,8	37,8	36,6	36,6	36,0	34,8	36,3	36,3	36,6	37,5	37,5	

R - roczne MSDO,  $\bar{x}$  - wartość średnia. Kolorem oznaczono przedziały wartości: 0 - 5, 6 - 10, a następnie co 10 od 11 do 80.

małej liczby par punktów, z której obliczane były wartości semiwariancji dla pierwszych odstępów, oraz z niskiej precyzji lokalizacji stanowisk. Mała liczba danych i niepewność ich klasyfikacji do poszczególnych odstępów powodowała niższą reprezentatywność obliczonej wartości. Kodowanie binarne redukuje co prawda czułość semiwariancji na występowanie ekstremalnych przypadków, ale z drugiej strony przy małej liczbie danych niewielkie jest także prawdopodobieństwo istnienia takich samych proporcji wartości powyżej i poniżej progów, jak w całej populacji. Są to zatem anomalie w pełni „uzasadnione” i stosunkowo łatwe do skorygowania. Nie mają również większego wpływu na jakość modelu. Procedura półautomatycznego modelowania w programie ISATIS umożliwia bowiem wykluczenie z obliczeń optymalizacji wartości poszczególnych odstępów. Selekcji dokonuje się na podstawie kryterium odległości lub liczby par danych.

Drugi typ problemów przy modelowaniu semiwariogramów danych kodowanych występował przy skrajnych progach, najczęściej przy 1 i 5 percentylu (ryc. 16 i tab. 2). Nie wykazywał on tym razem zmienności sezonowej, a polegał na chaotycznych fluktuacjach wartości semiwariancji w całym zakresie analizowanych odległości. Widoczne jest istnienie struktury przestrzennej, to jest generalnie semiwariancja rośnie wraz z przyrostem odległości pomiędzy porównywanymi danymi, ale trudno zidentyfikować typ struktury i jej zasięg. Jest to oczywiście efekt braku stabilności dwupunktowej statystyki w sytuacji małej ilości danych poniżej/powyżej progów i ich raczej rozproszonego rozmieszczenia. Takie układy są zatem częstsze w miesiącach z większą ilością opadów typu konwekcyjnego o stosunkowo małym zasięgu przestrzennym. W wyborze typu struktur i ich zasięgu w trakcie modelowania takich przypadków kierowano się zarówno widocznym na wykresie kształtem semiwariogramu empirycznego, jak i charakterem modelu dla sąsiedniej wartości progowej. Wychodzono bowiem z założenia, że w większości przypadków zmiany struktury przestrzennej pomiędzy poszczególnymi progami dokonują się w sposób „płynny”.

Trzeci, najbardziej kłopotliwy typ semiwariogramów danych kodowanych związany był głównie z najwyższymi wartościami progowymi – 95 i 99 percentylem (ryc. 17, tab. 3). Więcej było też takich przypadków pomiędzy październikiem a kwietniem. Wykazywały one brak struktury przestrzennej albo wręcz zmniejszanie się semiwariancji wraz ze wzrostem odległości, zwłaszcza dla kilku pierwszych odstępów (przykład z sierpnia 1965 dla 99 percentyla). Średnie wartości semiwariancji dla wszystkich odstępów w całym uwzględnianym zakresie odległości pomiędzy stanowiskami były często wyższe od wariancji próby. Tego typu semiwariogramy empiryczne przy braku jakichkolwiek informacji *a priori* o charakterze zmienności przestrzennej analizowanej cechy należałoby modelować używając pojedynczej struktury – nuggetowej. Oznaczałoby to, że najwyższe sumy dobowe dla

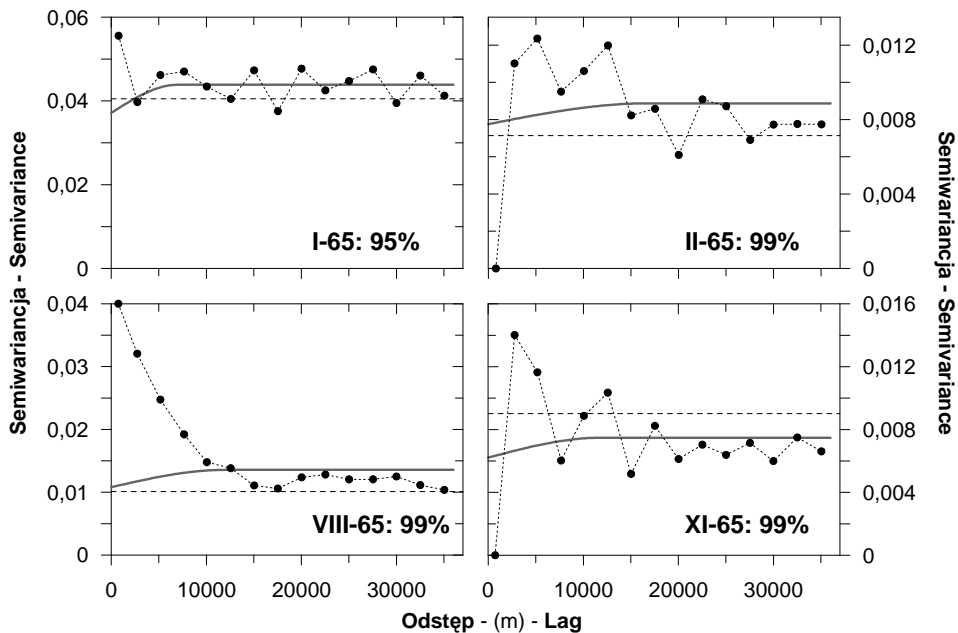


Ryc. 16. Przykłady chaotycznych semiwariogramów empirycznych danych kodowanych i dopasowane do nich modele. Linia przerywaną zaznaczono poziom wariancji danych

Tabela 2. Przybliżony procentowy udział semiwariogramów kodowanych z chaotycznymi fluktuacjami wartości dla całego analizowanego zakresu odległości w zależności od miesiąca i wartości progowej (percentyla, P)

P	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	R. / Y.	$\bar{x}$
1	56	60	64	76	84	80	68	80	64	60	80	64	76	69,7
5	56	32	44	52	48	48	20	52	56	56	56	48	44	47,3
10	28	32	16	40	24	20	16	8	28	28	32	16	28	24,0
20	8	20	0	4	8	12	8	8	4	8	8	12	4	8,3
30	4	8	0	8	0	4	0	4	4	12	4	8	0	4,7
40	4	0	0	8	0	4	0	4	0	8	0	0	0	2,3
50	4	0	0	0	0	16	0	4	4	4	0	4	0	3,0
60	4	0	0	8	0	4	0	4	4	4	0	0	0	2,3
70	12	0	0	8	4	4	0	0	0	4	4	4	0	3,3
80	20	4	8	8	4	12	8	8	8	4	12	8	4	8,7
90	32	32	20	48	28	24	28	32	20	40	40	28	40	27,7
95	44	40	36	44	44	40	60	44	48	40	32	36	52	39,7
99	40	36	32	40	36	20	48	48	52	32	40	16	28	33,3
$\bar{x}$	24,0	20,3	16,9	26,5	21,5	22,2	19,7	22,8	22,5	23,1	15,1	18,8	18,8	

R - roczne MSDO,  $\bar{x}$  - wartość średnia. Kolorem oznaczono przedziały wartości: 0 - 5, 6 - 10, a następnie co 10 od 11 do 80.



Ryc. 17. Przykłady semiwariogramów empirycznych danych kodowanych niewykazujących wyraźnej struktury oraz ich modele. Linią przerywaną zaznaczono poziom wariancji danych

Tabela 3. Przybliżony procentowy udział semiwariogramów kodowanych nie wykazujących wyraźnej struktury przestrzennej w zależności od miesiąca i wartości progowej (percentyla, P)

P	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	R. / Y.	$\bar{x}$
1	36	24	28	24	8	8	20	16	24	28	12	36	20	22,0
5	12	20	16	4	0	4	8	4	4	8	0	8	24	7,3
10	4	12	0	0	4	0	0	4	8	4	4	16	4	4,7
20	0	0	0	0	0	0	0	0	0	0	0	4	0	0,3
30	0	0	0	0	0	0	0	0	0	0	0	4	0	0,3
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0
70	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0
80	4	0	0	4	4	0	0	0	0	0	0	4	4	1,3
90	8	4	16	12	8	0	8	8	8	4	12	32	8	10,0
95	36	32	40	48	28	24	8	20	20	40	44	44	16	32,0
99	48	60	60	56	56	64	40	52	40	60	60	80	60	56,3
$\bar{x}$	11,4	11,7	12,3	11,4	8,3	7,7	6,5	8,0	8,0	11,1	10,2	17,5	10,5	

R - roczne MSDO,  $\bar{x}$  - wartość średnia. Kolorem oznaczono przedziały wartości: 0 - 5, 6 - 10, a następnie co 10 od 11 do 80.

danego zbioru danych MSDO występowały całkowicie losowo. Praktyka modelowania w opisywanych wyżej przypadkach polegała na całkowitej rezygnacji z automatycznej optymalizacji parametrów. Używane do estymacji i symulacji oprogramowanie (Ying 2000) nie dopuszcza wykorzystania modelu nuggetowego jako jedynego do stosowania samodzielnie (pojedynczo) do opisu struktury przestrzennej. Dlatego, za każdym razem konieczne było „sztuczne” dodanie składowej „regularnej” – zazwyczaj sferycznej – o bardzo niewielkiej wariancji cząstkowej. Takie postępowanie umożliwiało omińnięcie numerycznych ograniczeń stosowanych algorytmów, bez znaczącego zniekształcenia wyników estymacji i symulacji. Dodając „sztuczną” składową należało także podjąć decyzję o jej zasięgu. Brano wówczas pod uwagę zarówno szczegóły kształtu modelowanego chaotycznego semiwariogramu, jak i zasięg modeli opracowanych dla bardziej regularnych semiwariogramów sąsiednich, niższych wartości progowych. Należy jednak brać pod uwagę, szczególnie przy interpretacji zmienności uwarunkowanej względnym różnicowaniem wysokości sumy opadów, że struktura przestrzenna MSDO interpretowana na podstawie modeli danych kodowanych jest przy ich górnej granicy obciążona dużą dozą subiektywizmu.

Ocena liczby semiwariogramów danych kodowanych, które można by zakwalifikować do trzech opisanych wyżej grup jest nieco utrudniona (tab. 1-3). Wynika to przede wszystkim z płynnej, subiektywnej oceny stopnia regularności/chaotyczności semiwariogramu (odróżnienie grupy 2). Również jednoznaczne rozdzielanie przypadków należących do grupy drugiej i trzeciej nie było w wielu przypadkach możliwe, dlatego podane niżej odsetki należy traktować jako przybliżone. Do grupy pierwszej (anomalne wartości dla pierwszych odstępów) zaliczono 34,9% przypadków modelowanych semiwariogramów danych kodowanych, do grupy drugiej (semiwariogramy chaotyczne) – 19,5%, a do trzeciej (brak struktury) – 9,5%.

### 2.5.9. Nieciągłość i asynchroniczność danych MSDO a ich struktura przestrzenna

Przy omawianiu szczegółów estymacji metodą krigingu danych kodowanych (dodatek X.2) wspomniano o problemie naruszania relacji porządkowych, polegającym głównie na braku konsekwentnego następstwa szacowanych prawdopodobieństw dla kolejnych wartości progowych odczytywanych ze skumulowanych funkcji rozkładu (*cdf*). W celu zmniejszenia możliwości zaistnienia takich wyników zaleca się stosowanie przy modelowaniu zawsze tej samej kombinacji modeli podstawowych, a ich parametry powinny przy kolejnych wartościach progowych zmieniać się stopniowo. Takie postępo-

wanie ma pełne uzasadnienie w wypadku struktury przestrzennej synchronicznych zmiennych (cech) ciągłych. W kolejnych klasach wielkości/natężenia takich cech zmienia się ona niewiele i raczej w sposób „płynny”. Sumy dobowe opadów, a tym bardziej ich podzbiór wykorzystywany w niniejszej pracy, takiej natury nie mają. Z jednej strony opady są nieciągłe przestrzennie, z drugiej – dane MSDO dla poszczególnych miesięcy i lat pochodzą z różnych terminów, czyli są niesynchroniczne. Zróżnicowana jest także ich geneza – powstają w efekcie działania kilku zjawisk operujących w innych skalach przestrzennych i czasowych. Dlatego, nie można ignorować faktu, często spotykanego w trakcie niniejszej pracy, że semiwariogramy empiryczne dla kolejnych wartości progowych zmieniały się czasami dość znacznie, raczej skokowo niż „płynnie”.

## IV

---

### Charakterystyka pluwiotermiczna lat 1956-1980

Okres 1956-1980 jest z punktu widzenia reżimu opadów w XX wieku dość specyficzny. Według bazy danych TYN CY 1.1 ([http://www.cru.uea.ac.uk/~timm/cty/obs/TYN\\_CY\\_1\\_1.html](http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1.html)), w jego trakcie na terytorium Polski zanotowano aż 5 lat z najwyższymi sumami opadów w stuleciu (1970, 1966, 1974, 1967 i 1960, ryc. 18). Kożuchowski (1985) podaje nieco inne wartości niż cytowane powyżej. W jego zestawieniu, obejmującym stulecie 1881 do 1980, w latach 1956-1980 wstąpiły dwie najwyższe sumy roczne (1970 i 1974) oraz czwarta z kolei (1966). Były to, niezależnie od źródła informacji, opady wyraźnie przekraczające 700 mm. Kożuchowski (1985) ocenia, że dekada lat 1965-1974 była najwilgotniejsza od roku 1881. Według tego samego autora, na drugim biegunie pod tym względem był okres od roku 1950 do 1959, a więc obejmujący cztery pierwsze lata analizowanego wielolecia<sup>25</sup>. W wartościach względnych zatem, ujmując całe 25 lat, różnice średnich sum rocznych opadów w stosunku do całego stulecia nie były jednak zbyt duże. Opierając się na danych zgromadzonych w bazie TYN CY 1.1 stwierdzono jedynie 2,7% przeciętnego wzrostu sum rocznych (613,5-597,5 mm). Nieco mniejszą różnicę, wynoszącą +2,2%, obliczono z danych Kożuchowskiego (1985), porównując średnie roczne sumy opadów analizowanego wielolecia z latami 1881-1980 (645,2-631,5 mm). Jak można się było spodziewać, opierając się na przedstawionym powyżej opisie, istota pluwiometrycznej odmienności lat 1956-1980 w odniesieniu do pozostałej części XX stulecia związana jest ze stopniem zróżnicowania opadów. Testowanie istotności różnic średnich sum rocznych dało wynik negatyw-

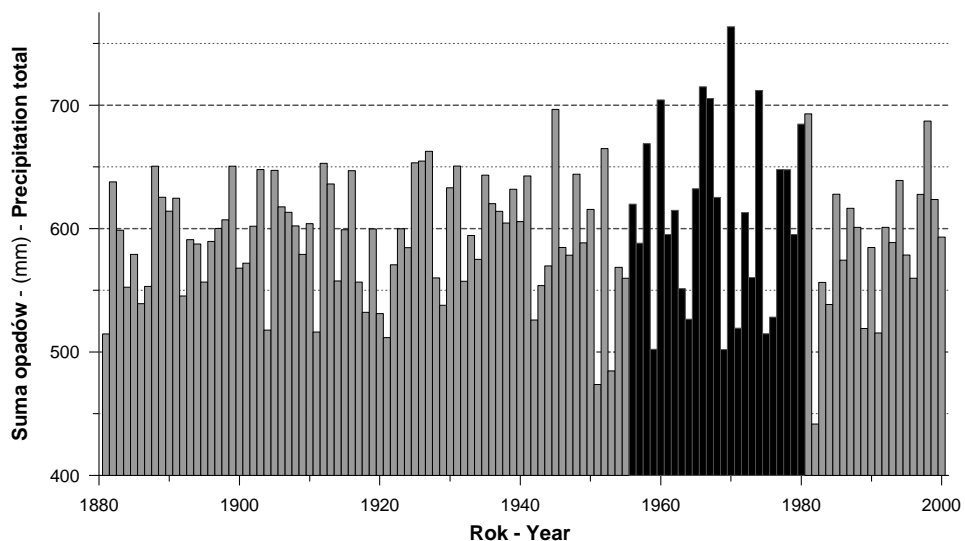
---

<sup>25</sup> Najniższe sumy roczne w tej dekadzie wystąpiły jednakże w latach 1951 (473,3 mm) i 1953 (484,7 mm). Najsuchszym rokiem w całym stuleciu był 1982 z roczną sumą wynoszącą 441,6 mm.

ny, podczas gdy każdy z zastosowanych testów jednorodności wariancji wykazywał jej brak na poziomie  $p \approx 0,01$ .

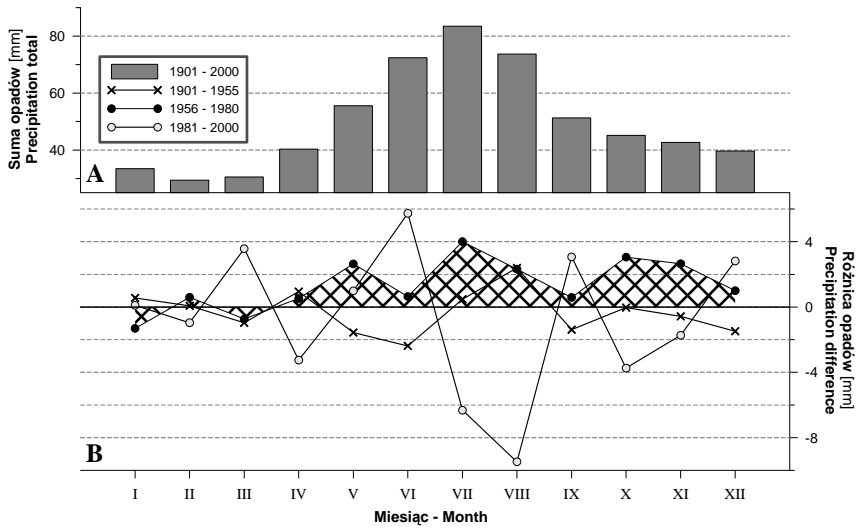
Zwrócono również uwagę na rozkład sezonowy opadów (ryc. 19). Porównanie pod tym względem analizowanego wielolecia z danymi dla całego XX stulecia wykazuje występowanie konsekwentnie wyższych średnich sum miesięcznych od maja do grudnia. Nie są to jednak w żadnym wypadku różnice istotne statystycznie. Na uwagę zasługuje jednak zdecydowanie większa zmienność sum miesięcznych w lutym ( $p = 0,004$ ) oraz maju ( $p = 0,049$ ).

Pod względem termicznym lata 1956-1980 również nie odróżniają się wyraźnie od charakterystyk typowych do całego stulecia (ryc. 20 i 21). Brak było przede wszystkim w tym okresie lat bardzo gorących, o średnich temperaturach rocznych równych lub większych od  $9^{\circ}\text{C}$ , szczególnie częstych w ostatnim dwudziestolecu. Z tego właśnie powodu, zwłaszcza średnie miesięczne okresu zimowo-wiosennego były w latach 1956-1980 nieco niższe niż w całym stuleciu. Mogło to w konsekwencji doprowadzić do nieco wyższej częstości występowania opadów stałych. Ponownie jednak w wypadku opadów różnice średnich miesięcznych mieszczą się w zakresie zmienności losowej. Tylko wariancja temperatur marca była w wieloleciu 1956-1980 ( $p = 0,001$ ) znacznie większa niż w pozostałej części XX wieku.

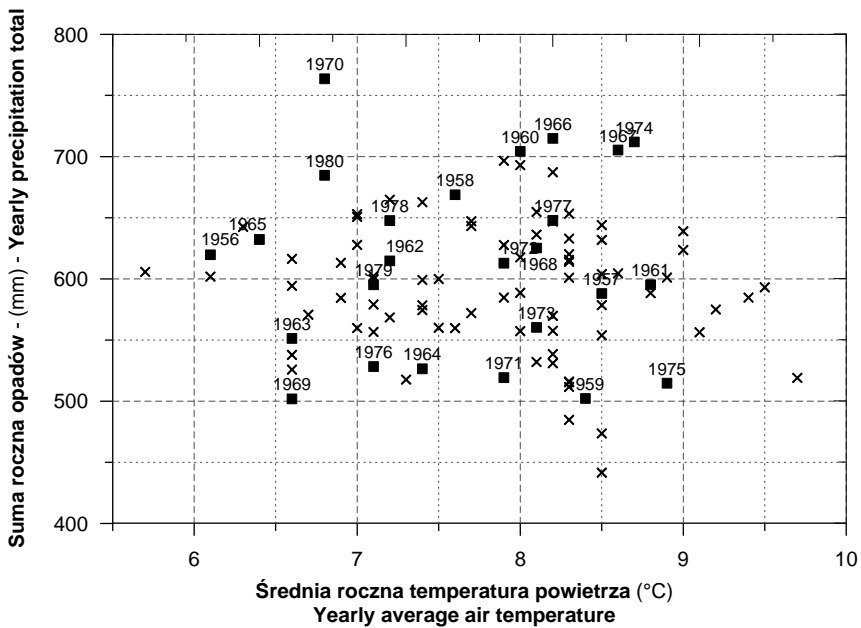


Ryc. 18. Sumy roczne opadów atmosferycznych na terenie Polski w latach 1881-2000. Dane dla stulecia 1901-2000 pochodzą z bazy danych TYN CY 1.1 ([http://www.cru.uea.ac.uk/~timm/cty/obs/TYN\\_CY\\_1\\_1.html](http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1.html)), natomiast dla lat 1881-1990 pozyskano z Kozuchowskiego (1985) i przeliczono za pomocą regresji liniowej

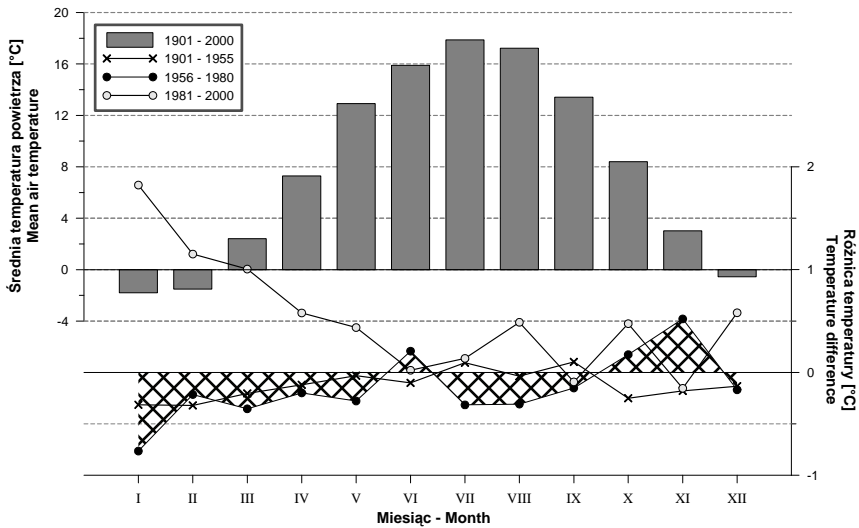




Ryc. 19. Średnie sumy miesięczne opadów w stuleciu 1901-2000 (A) oraz średnie od nich odchylenia (B) w trzech okresach wieloletnich: 1901-1955, 1956-1980 (dla wyróżnienia zaznaczono nie szrafem) oraz 1981-2000 (źródło: baza danych TYN CY 1.1 ([http://www.cru.uea.ac.uk/~timm/cty/obs/TYN\\_CY\\_1\\_1.html](http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1.html)))



Ryc. 20. Relacje między średnimi rocznymi temperaturami powietrza i sumami rocznymi opadów na obszarze Polski w stuleciu 1901-2000 (źródło: baza danych TYN CY 1.1 ([http://www.cru.uea.ac.uk/~timm/cty/obs/TYN\\_CY\\_1\\_1.html](http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1.html))). Analizowane wieloletnie wyróżniono typem symbolu



Ryc. 21. Średnie miesięczne temperatury powietrza w stuleciu 1901-2000 (A) oraz średnie od nich odchylenia (B) w trzech okresach wieloletnich: 1901-1955, 1956-1980 (dla wyróżnienia zaznaczone szrafem) oraz 1981-2000 (źródło: baza danych TYN CY 1.1 ([http://www.cru.uea.ac.uk/~timm/cty/obs/TYN\\_CY\\_1\\_1.html](http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1.html)))

Z punktu widzenia tematu niniejszej rozprawy, o reprezentatywności analizowanego wielolecia niewątpliwie decyduje przede wszystkim reżim opadów dobowych. Niestety, w momencie przygotowywania pracy do druku nie dysponowano odpowiednią liczbą dostatecznie długich serii pomiarów sum dobowych, aby taką analizę przeprowadzić<sup>26</sup>.

<sup>26</sup> Dysponowano seriami sum dobowych opadów z lat 1951–2006 dla kilku stacji usytuowanych w północno-wschodniej Polsce. Dane te nie zawierały informacji o typie opadów. Dla Kętrzyna i Białegostoku wykonano ocenę istotności różnic w ilości dni z opadem i bez opadu porównując analizowane wielolecie (lata 1956–1980) z całym okresem lat 1951 – 2006, oraz z przedziałem od roku 1981 do 2006. We wszystkich porównywanych układach różnice były statystycznie nieistotne ( $p > 0,05$ ). Również test Kołmogorowa-Smirnowa zastosowany do oceny różnic rozkładów wielkości sum dobowych opadów wykazał ich nieistotność. Nie świadczy to jednakże o braku jakichkolwiek różnic. Mogą one bowiem mieć charakter sezonowy i regionalnie zmienny.

---

# Dane pomiarowe maksymalnych sum opadów dobowych

## 1. Źródła i charakter danych

Podstawowym materiałem źródłowym, który wykorzystano w niniejszej pracy są publikowane w rocznikach „Opady Atmosferyczne” PIHM/IMGW<sup>27</sup> dane maksymalnych sum dobowych opadów rejestrowanych na stacjach meteorologicznych i posterunkach opadowych w poszczególnych miesiącach wielolecia 1956-1980, uzupełnione datą wystąpienia maksymalnej rocznej sumy dobowej opadów. Obejmują one lata 1954-1981, zatem maksymalnym standardowym okresem wieloletnim, który może być opracowany na ich podstawie jest dwudziestopięciolecie. Oprócz danych z obszaru Polski udało się uzyskać z **Deutscher Wetterdienst (DWD, Niemiecka Służba Meteorologiczna)** wartości dobowych sum opadów dla wszystkich punktów pomiarowych istniejących w latach 1956-1980, a leżących na terytorium byłego NRD, w pasie około 30 km od granicy Polski (patrz załącznik XII.1). Korzystając z tych materiałów, samodzielnie przygotowano analogiczny zbiór maksymalnych miesięcznych i rocznych sum opadów dobowych. W sumie dysponowano bazą danych, na którą składało się 747 486 wartości maksymalnych opadów dobowych w poszczególnych miesiącach (średnio ok. 2492) i 61 940 dat/sum maksymalnych rocznych opadów dobowych (średnio ok. 2478). Lokalizacji, dla których istniały informacje o dacie wystąpienia rocznych MSDO było w istocie tylko 61 558. Różnica między obiema wyżej podanymi liczbami wynika z tego, że w tym samym punkcie notowano w roku dwukrotnie lub nawet w wyjątkowych sytuacjach trzykrotnie maksymalny opad o tej samej – w ramach dokładności pomiaru oczywiście – wysokości. Obecnie sieć pomiarowa opadów na

---

<sup>27</sup> Państwowy Instytut Hydrologiczno-Meteorologiczny zmienił swoją nazwę na obowiązującą obecnie – Instytut Meteorologii i Gospodarki Wodnej – w 1973 roku.

terytorium Polski jest o ponad połowę mniejsza (Stachy 1999). Biorąc pod uwagę nizinny w przewodzie charakter naszego kraju, zagęszczenie stanowisk pluwiometrycznych było w latach 1956-1980 duże, a uzyskany w tym okresie materiał pomiarowy jest bardzo, z naukowego punktu widzenia, cenny.

Według wiedzy autora, dane te nigdy do tej pory nie były w całości analizowane i opracowane. Dlatego, w kolejnych rozdziałach sporządzono ich wszechstronną charakterystykę, która może być użyteczna nie tylko ze względu na główne zagadnienia poruszane w niniejszej rozprawie.

## 2. Błędy i niepewność danych

Błędy w analizie przestrzennej są związane zarówno z wynikami pomiarów rozpatrywanej cechy, jak i z określeniem lokalizacji stanowisk, gdzie owe pomiary wykonano (Gabrosek, Cressie 2002). W niniejszym opracowaniu mogły one pochodzić zarówno ze źródła danych (roczników opadowych), jak i zostać wprowadzone na etapie ich digitalizacji. W celu zminimalizowania możliwości zaistnienia drugiej z wymienionych wyżej ewentualności, po dokonaniu archiwizacji elektronicznej dane MSDO były dwukrotnie sprawdzane – zarówno przez osobę dokonującą archiwizacji, jak i przez kogoś drugiego. W wielu wypadkach wątpliwości dotyczyły złej jakości druku Roczników bądź ich znacznego w ciągu wielu lat zużycia. Sugerowano się wówczas wartościami zarejestrowanymi w sąsiadujących stanowiskach. Identyfikacja błędnych wpisów w oryginalnej analogowej bazie była znacznie trudniejsza. Oczywiście, najłatwiej było wyselekcjonować wartości nierealne, na przykład współrzędne stanowisk pomiarowych wskazujące na lokalizacje usytuowane setki kilometrów od obszaru Polski. Wątpliwe wyniki pomiarów MSDO identyfikowano w trakcie analizy wariograficznej (autokorelacji) opisanej w podrozdziałach III.2.5.7 i III.2.5.8 oraz w dodatku X.5. Mogły być one jednak efektem zarówno błędów pomiarowych, jak i autentycznych anomalnych opadów, na przykład orograficznych. Dane te były eliminowane z obliczeń i modelowania struktury przestrzennej, ale uwzględniano je przy estymacjach i symulacjach. Ważnym elementem weryfikacji elektronicznej bazy danych było zestawienie alfabetyczne lokalizacji punktów pomiarowych dla całego wielolecia. Pozwalało to wówczas na odróżnienie, z dużym prawdopodobieństwem, tych przypadków, kiedy stanowisko pomiarowe zostało przeniesione z roku na rok w inne miejsce w tej samej miejscowości (zmiany współrzędnych rzędu 1-3 minut kątowych), od ewidentnie błędnych wpisów. Zdarzały się bowiem sytuacje, kiedy stanowisko istniejące przez wiele lat w tej samej lokalizacji, nagle podawane jest ze współrzędnymi wskazującymi na miejsce odległe

o kilkadziesiąt, czy nawet kilkaset kilometrów. W przypadku dwóch lat – 1960 i 1962 – w rocznikach w ogóle nie podano współrzędnych posterunków opadowych. Niestety, nie było praktycznej możliwości weryfikacji wątpliwości poprzez konfrontację z oryginalną bazą danych przechowywaną w IMGW. Trzeba wyraźnie podkreślić, że mimo bardzo dużej staranności w trakcie przygotowywania bazy danych nie udało się wyeliminować wszystkich błędów. Kilkakrotnie na różnych etapach analizy identyfikowano wartości, które z różnych względów można byłoby zakwestionować. Niestety, opcja powtórzenia obliczeń nie wchodziła już w grę. Można jedynie wskazać, że ze względu na rozmiary wykorzystywanej bazy danych pojedyncze błędne wartości nie powinny znacząco wpływać na uzyskane wyniki.

Istotnym źródłem niepewności uzyskanych w niniejszej pracy wyników analiz przestrzennych jest niska dokładność lokalizacji stanowisk pluwiometrycznych. Lokalizację zarówno polskich, jak i niemieckich stacji pomiarowych znano bowiem z dokładnością do jednej minuty długości i szerokości geograficznej. Zostały one przeliczone na współrzędne płaskie w jednostrefowym układzie GUGIK 1992 (z południkiem osiowym 19°). Błędy algorytmu transformacji oraz związane z odwzorowaniem nie przekraczają na terytorium Polski 0,5 m, a więc są nieistotne w stosunku do poziomu dokładności lokalizacji punktów pomiarowych. Zmiana długości geograficznej o jedną minutę przy 49° szerokości geograficznej oznacza w układzie 1992 odległość około 1313 m, a przy 55° – 1089 m. Przesunięcie o jedną minutę szerokości geograficznej oznacza w tych samych dwóch położeniach różnicę odpowiednio: 1914 i 1865 m. Tego rzędu niepewność lokalizacji (600-900 m) jest do pominięcia, jeśli wykonujemy analizę opierając się na kilkunastu – kilkudziesięciu stanowiskach odległych od siebie o kilkadziesiąt do stu kilkudziesięciu kilometrów. Rejestrujemy wówczas jedynie zmienność regionalną, ignorując czynniki lokalne związane na przykład z topografią terenu<sup>28</sup>. W niniejszej pracy dysponowano jednakże dla każdego terminu zbiorem ponad dwóch tysięcy stanowisk pomiarowych. Według dokonanych obliczeń (patrz podrozdz. V.4), średnia odległość między najbliższymi stanowiskami wynosiła w wieloleciu około 7,3 km. Niepewność położenia mogła zatem maksymalnie wynosić około ¼ tego dystansu. Założono jednak, bo nie było żadnych podstaw do innych ocen, że rozkład błędów określenia położenia stanowisk pomiarowych, przy tak dużej próbie, można traktować jako czysto losowy. Innym, ważnym, a niewyjaśnionym jednoznacznie problemem jest źródło, na podstawie którego pracownicy PIHM/IMGW określali współrzędne położenia tworzonych, a później ewentualnie przenoszonych

---

<sup>28</sup> Na obszarach górskich wspomniany wyżej zakres niepewności lokalizacji może się wiązać z zasadniczymi różnicami w reżimie opadów, związanymi z piętrami wysokościowymi, czy odmienną ekspozycją.

posterunków pomiarowych. Wiadomo, że były one odczytywane z map topograficznych w skali 1 : 100 000, lecz niestety nie wiadomo dokładnie jakich. Rozpoczynając pracę nad analizą MSDO założono, że były to mapy w układzie „1942”. Istnieje jednakże wysokie prawdopodobieństwo, że w latach pięćdziesiątych XX wieku korzystano jedynie z map WIG w układzie „Borowa Góra”. Ponieważ układy odniesienia (elipsoida, punkt przyłożenia) w obu systemach były odmienne, te same współrzędne geograficzne oznaczają w rzeczywistości nieco inne lokalizacje. Różnice w obrębie Polski wynoszą od około 80 do około 190 m. Byłby to jednak błąd systematyczny, bardzo łatwy do skorygowania. Gorzej, jeśli w trakcie rozbudowy sieci, w różnych latach korzystano z różnych map do określenia położenia nowych stanowisk. Jakkolwiek podany wyżej przedział różnic lokalizacji nie jest duży, zwiększałby on jednak znacząco niepewność określenia lokalizacji związaną z minutową rozdzielczością podawanych w Rocznikach Opadowych współrzędnych. Sprawdzono także, na przykładzie kilku posterunków lotniskowych, czy w analizowanym wieloleciu nie było zmiany współrzędnych, które mogłyby sugerować ponowne ich określenie na podstawie innego podkładu topograficznego. Niczego takiego nie stwierdzono.

Analizę struktury przestrzennej (wariograficzną) przeprowadzano dla klas przedziałów odległości stanowisk co 2,5 km. Bardzo duża – i nieznana – liczba porównywanych par wartości MSDO było z pewnością przydzielane do niewłaściwych klas. Dla odległości powyżej 10 km nie miało to w zasadzie żadnego znaczenia, ze względu na liczebność próby przekraczającej zazwyczaj kilka tysięcy. Wyraźny efekt tej właściwości analizowanych danych był dostrzegany jednak dla pierwszych trzech odstępów, gdzie liczba porównywanych par oscylowała od kilku do kilkuset. Objawiało się to często, szczególnie w wypadku semiwariogramów danych kodowanych, chaotycznymi zmianami wartości semiwariancji (patrz ryc. 15). Zagadnienie to zostało szerzej omówione w podrozdziale III.2.5.8.

Innym ważnym problemem związanym z niską precyzją lokalizacji posterunków opadowych było występowanie w każdym analizowanym zbiorze danych miesięcznych i rocznych punktów o takich samych współrzędnych. Były to stanowiska o odmiennych nazwach i różnych rzędnych wysokości terenu, ale położone na tyle blisko siebie, że nie przekraczało to 1 minuty długości i szerokości geograficznej. Dla większości algorytmów estymacji i symulacji jest to niedopuszczalne. Stosuje się wówczas szereg automatycznych procedur powodujących albo eliminację jednego bądź obu zdublowanych stanowisk, albo uśrednianie zmierzonych na nich wartości. W niniejszej pracy zdecydowano się na nieco inne rozwiązanie. Przy analizie zmienności przestrzennej MSDO dane takie były bowiem bardzo cenne. Umożliwiały ocenę owej zmienności na bardzo niewielkich dystansach. Biorąc bowiem pod uwagę niepewność lokalizacji, wynoszącą w różnych miej-

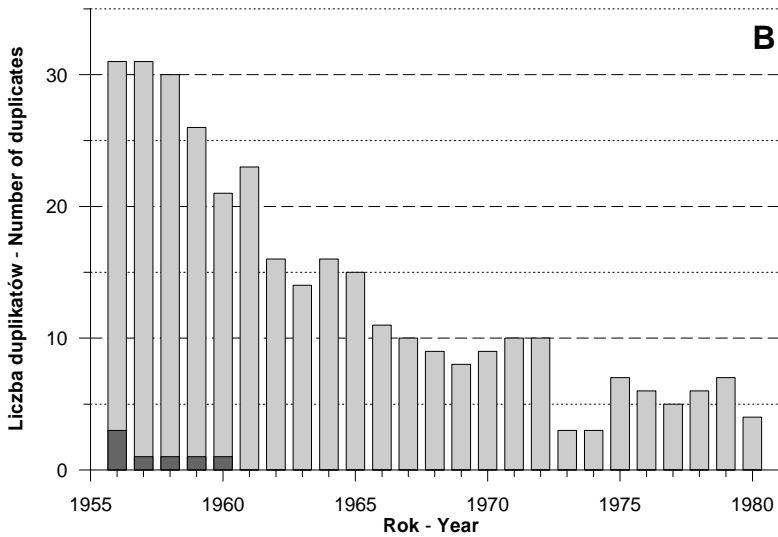
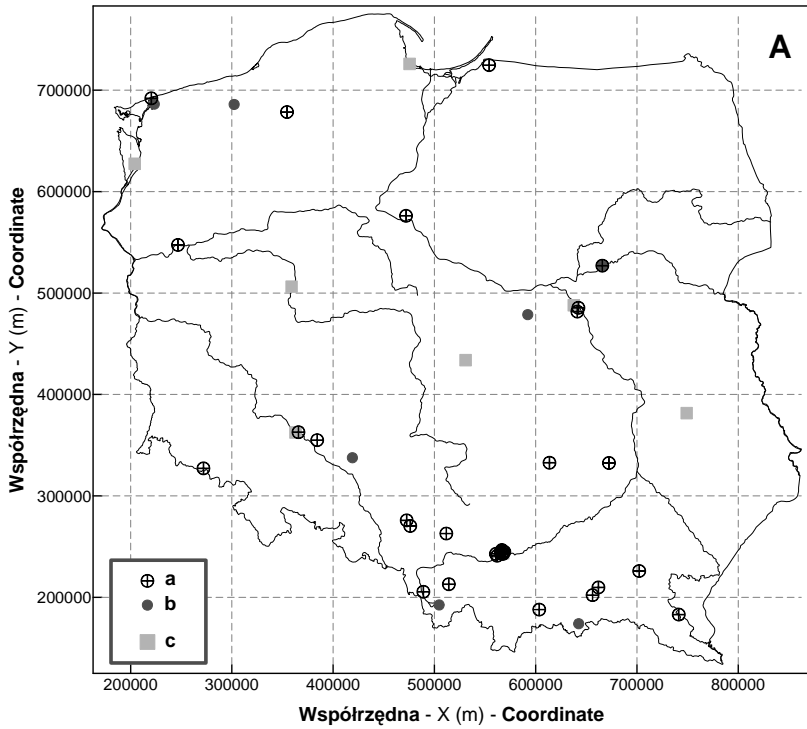
scach Polski  $\pm 0,6-0,9$  km, i przyjęty krok izotropowej analizy autokorelacji przestrzennej równy 2,5 km (patrz podrozdz. VI.3), można było z dużym prawdopodobieństwem przypuszczać, że wszystkie stanowiska o takich samych współrzędnych geograficznych znajdują się w tak zwanym zerowym (połówkowym) odstępnie, wynoszącym 1,25 km. Rzeczywista odległość między nimi, jakkolwiek ważna, ma tutaj znaczenie drugorzędne. Zdecydowano się zatem punktom o tych samych współrzędnych przypisywać losowo nowe położenie w odległości  $\pm 0,5$  km od wyliczonego i całą dalszą procedurę wykonywać na takich danych. Zapewniało to z jednej strony pożądane informacje o zmienności MSDO na krótkich dystansach, a z drugiej, przy przyjętej rozdzielczości siatki estymacji i symulacji wynoszącej  $1 \times 1$  km wystarczający poziom „wygładzania” pomiarów „odstających” ze względu na niepewność określenia ich lokalizacji. Trzeba jednakże dodać, że liczba „zdublikowanych”, ze względu na położenie, stanowisk jest zróżnicowana zarówno geograficznie, jak i czasowo (ryc. 22). Najwięcej lokalizacji pomiarowych, bo około 30, o identycznych współrzędnych było w trzech pierwszych latach analizowanego wielolecia (ryc. 22B). Wtedy również istniały takie przypadki, że identyczne położenie miały aż trzy stanowiska! Przez kolejne 10 lat od roku 1959 trwał dość szybki spadek liczby „duplikatów” – do poziomu około 10. Minimalna liczba takich stanowisk pomiarowych – tylko po 3 – istniała w latach 1973 i 1975. W drugiej połowie lat siedemdziesiątych ich liczba oscylowała między 4 a 7. W latach pięćdziesiątych i pierwszej połowie sześćdziesiątych znajdowały się one przede wszystkim na południu Polski (ryc. 22A), z największym zagęszczeniem w obrębie krakowskiej sieci pomiarowej opadów atmosferycznych. Na obszarze Niżu takich lokalizacji było tylko kilka. W latach siedemdziesiątych ich rozmieszczenie na obszarze kraju było z grubsza równomierne.

### 3. Zmienność liczby stanowisk pomiarowych

W rozpatrywanej serii 25 lat, na podstawie zmian liczebności sieci pomiarowej opadów atmosferycznych na terenie Polski<sup>29</sup> można wyodrębnić kilka okresów (ryc. 23). Najpierw do końca 1961 trwał szybki, a później stopniowo słabnący trend zwiększania liczby stanowisk – od 1867 w lutym 1956 roku do 2432 w październiku roku 1961. Od stycznia 1962 po stosunkowo niewielkim (około 100 punktów), ale raptownym spadku, liczba stanowisk stopniowo rosła, osiągając ponownie poziom około 2400 we wrześniu

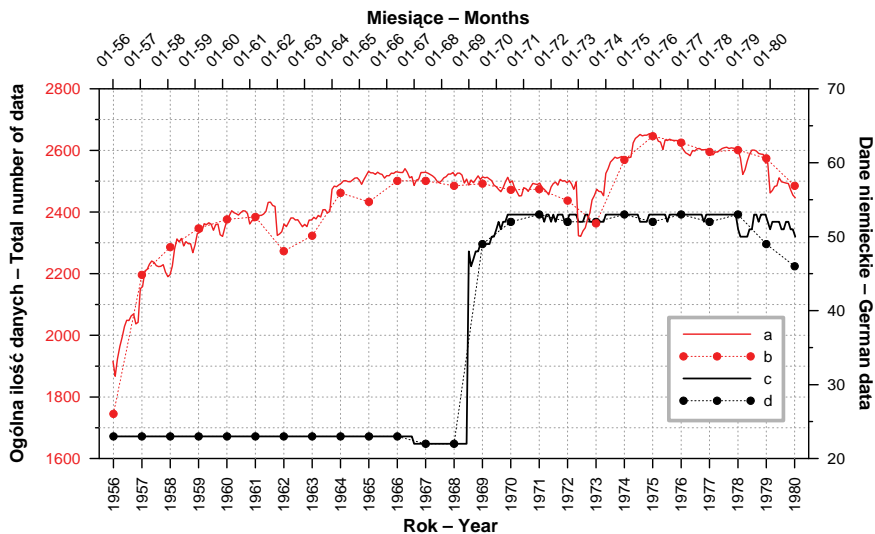
---

<sup>29</sup> Podawane poniżej liczby dotyczą rzeczywistej ilości danych zgromadzonych w rocznikach „Opady Atmosferyczne”, a nie liczby wymienianych lokalizacji. Nie są zatem uwzględniane stanowiska, które istniały, ale dla których w konkretnych terminach nie było wyników (braki danych).



**Ryc. 22.** Rozmieszczenie i liczba stanowisk pomiaru opadów atmosferycznych w wieloletniu 1956-1980 o identycznych współrzędnych lokalizacji: A - rozmieszczenie takich stanowisk w roku 1956 (a) i 1975 (b); symbolem (c) oznaczono położenie głównych miast; B - zmiany liczby w kolejnych latach wieloletnia; ciemniejszym odcieniem zaznaczono liczbę przypadków, gdzie aż trzy stanowiska miały takie same współrzędne





Ryc. 23. Zestawienie liczby pomiarów MSDO w Polsce i Niemczech, które wykorzystano w niniejszym opracowaniu: a – dane miesięczne z Roczników Opadów PIHM/IMGW, b – dane roczne z Roczników Opadów PIHM/IMGW, c – miesięczne dane z przygranicznej strefy na obszarze Niemiec, d – roczne dane z przygranicznej strefy na obszarze Niemiec

roku 1963. Na początku roku 1964 nastąpił skokowy – o około 80 – wzrost i w ciągu następnych blisko 10 lat, do grudnia roku 1972, liczba punktów pomiarowych oscylowała wokół 2500. W styczniu roku 1973, w porównaniu z poprzednim miesiącem, funkcjonowało o 175 stacji mniej. Następnie, przez ponad dwa lata sieć pomiarowa była intensywnie rozbudowywana, osiągając maksymalnie 2654 lokalizacje w sierpniu i wrześniu 1975 roku. Od tego momentu trwała początkowo powolna, a w roku 1980 znacznie przyspieszona tendencja do likwidacji części posterunków.

W polskiej sieci pomiarowej, oprócz dużych zmian liczebności związanych z powstawaniem i likwidacją stanowisk, wyraźnie widoczny jest cykl sezonowy: mniej danych zimą – więcej w pozostałych porach roku, z maksimum latem. Te różnice wynoszą od około 20 do ponad 60 rocznie. W danych pochodzących z Niemiec takich fluktuacji nie było (ryc. 23). W uwzględnianej w pracy, przygranicznej strefie o szerokości 30 km, istniały w latach 1956-1966 23 posterunki opadowe. Przez następne dwa lata było ich o jeden mniej. W roku 1969 i na początku 1970 postępowała rozbudowa sieci. W połowie roku 1970 liczba stanowisk ustabilizowała się na poziomie 53 i tak pozostało do końca analizowanego okresu. Zatem, do roku 1969 dane pochodzące z byłego NRD stanowiły około 1% analizowanego zbioru danych, później ich udział zwiększył się do około 2%.

Komentarza wymagają relacje pomiędzy ilością danych MSDO miesięcznych i rocznych. W dziesiątkach przypadków w rubryce daty maksymalnego rocznego opadu dobowego istniał wpis, mimo że nie było kompletu wartości miesięcznych. Taka sytuacja miała pełne uzasadnienie, kiedy z danych z leżących obok posterunków, posiadających kompletny rejestr, wynikało jednoznacznie, że maksymalny roczny opad dobowy nie mógł wystąpić w miesiącach, dla których istniały luki pomiarowe. Dlatego, w większości analizowanego okresu nie ma większych różnic między ilością danych miesięcznych względem rocznych. Jak wspomniano uprzednio, luki pomiarowe zazwyczaj występowały w miesiącach zimowych, kiedy istniało znikome prawdopodobieństwo wystąpienia rocznego maksymalnego opadu dobowego.

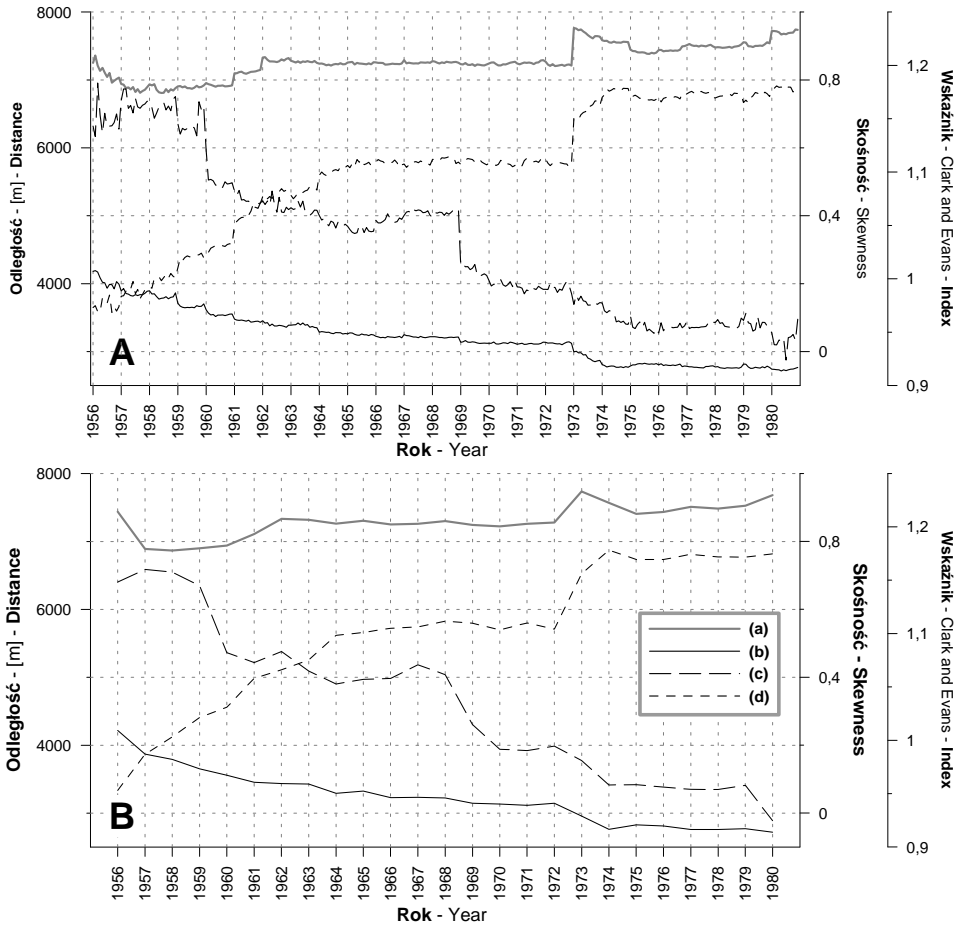
#### 4. Rozmieszczenie przestrzenne stanowisk pomiarowych

Nieco inny obraz daje zmienność wskaźników charakteryzujących rozmieszczenie przestrzenne posterunków opadowych, z których pozyskano dane analizowane w niniejszym opracowaniu (ryc. 24). Pokazuje ona, że zmiany „organizacji przestrzennej” sieci pomiarowej nie są ściśle związane ze zmianami jej liczebności. Jako syntetyczne miary rozkładu przestrzennego punktów pomiarowych użyto statystyki najbliższego sąsiada NNS (ang. *Nearest Neighbor Statistics*)<sup>30</sup>: średnią odległość, odchylenie standardowe i skośność odległości oraz wskaźnik Clarka-Evansa<sup>31</sup> losowości rozkładu przestrzennego. Każdy wymieniony wyżej wskaźnik charakteryzuje nieco inne aspekty zagadnienia. Najbardziej konsekwentny przebieg ma krzywa odchylenia standardowego. Jej wartość, oprócz kilku niewielkich fluktuacji w pierwszym pięcioleciu, przez cały analizowany okres spadała od 4,2 km w lutym 1956 roku do 2,7 km w maju 1980 roku. Najszybszy spadek miał miejsce do roku 1964 (do poziomu 3,26 km). Później, szczególnie w latach 1969-1973 zmiany były minimalne. Kolejny okres większego spadku SD odległości NN trwał w roku 1973 (z 3,14 do 2,77 km). Średnia odległość NN zmieniała się w całym 25-leciu bardzo niewiele: z 6,8 km w czerwcu 1958 roku do blisko 7,8 km w styczniu 1973 roku, czyli około 12%. W tym samym czasie liczebność zbioru zmieniała się przecież o 30% (z 1890 do 2707). Przez 220 miesięcy (73% czasu) wartość średniej odległości mieściła się w przedziale 7,1-7,6 km.

---

<sup>30</sup> Wyniki tych obliczeń są zawarte na załączonym DVD.

<sup>31</sup> Wskaźnik Clarka-Evansa (Cressie 1993) jest to stosunek między rzeczywistą średnią odległością od najbliższego sąsiada a oczekiwaną dla rozkładu losowego. Wartości niższe od jedności wskazują na występowanie skupień punktów, wyższe – na bardziej regularny ich rozkład (np. dla regularnej siatki heksagonalnej wartość wskaźnika wynosi 2,15). Zakres jego zmienności w analizowanym zbiorze danych, to jest od 0,95 do 1,18, wskazuje generalnie losowy charakter rozkładu.

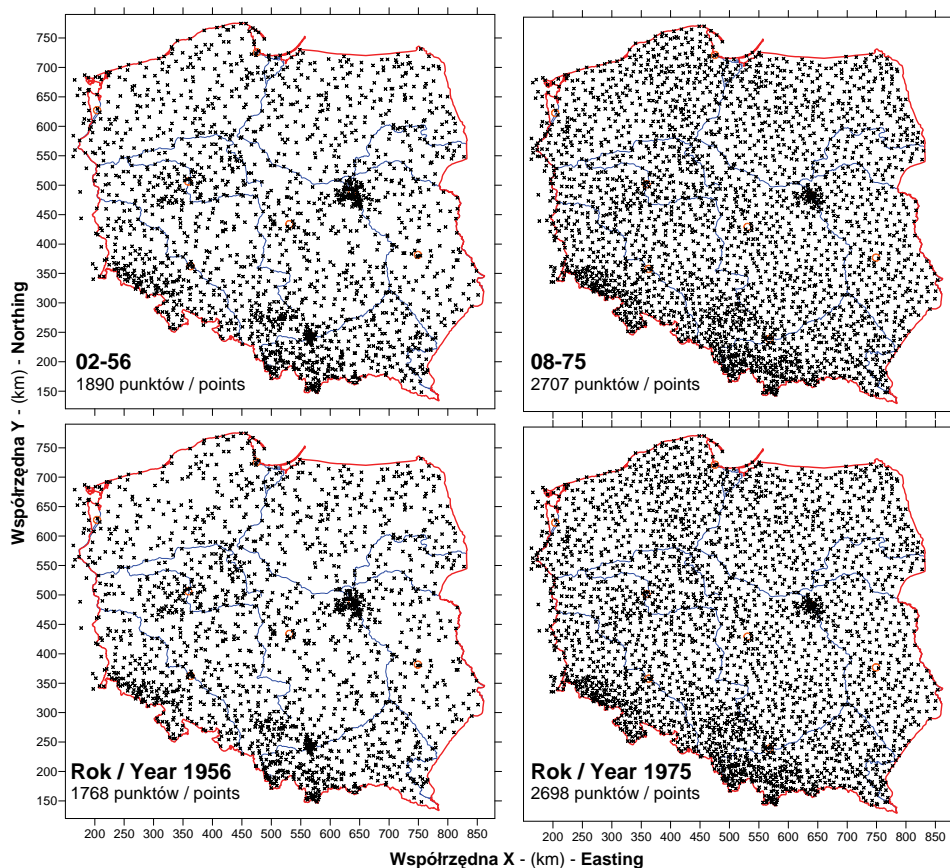


**Ryc. 24.** Statystyki odległości najbliższego sąsiada i losowości rozkładu przestrzennego punktów pomiarowych, dla których analizowano MSDO opadów, w kolejnych miesiącach (A) i latach (B) wielolecia 1956-1980: (a) – średnia odległość do najbliższego sąsiada, (b) – odchylenie standardowe odległości do najbliższego sąsiada, (c) – skośność odległości do najbliższego sąsiada, (d) – wskaźnik Clarka i Evansa losowości rozkładu przestrzennego

Uzupełnieniem obrazu rozkładu przestrzennego punktów pomiarowych, który się rysuje na podstawie analizy średniej odległości do najbliższego sąsiada i jej odchylenia standardowego, jest skośność tego samego parametru. Najwyższe jego wartości (0,65-0,8) i jednocześnie wykazujące największe fluktuacje miały miejsce pod koniec lat pięćdziesiątych. Największy spadek skośności odległości NN nastąpił między grudniem roku 1959 a lutym roku 1960 (z 0,723 do 0,505). Później znacznie wolniejsze obniżanie się

wartości skośności trwały do kwietnia 1965 roku (do poziomu 0,348). Do grudnia 1968 roku wartość tego parametru lekko wzrosła (do około 0,410), by następnie utrzymywać się na stałym poziomie. Potem miało miejsce gwałtowne jego obniżenie – do poziomu 0,252. Przez pozostałe 12 lat wartość skośności odległości NN malała. Nieco szybsze tempo redukcji było widoczne w latach 1973-1975, podczas gdy później utrzymywało się na prawie stałym poziomie.

Analizowane powyżej statystyki ilustrują generalną tendencję do bardziej równomiernego rozmieszczenia punktów pomiarowych w sieci. Jej jednoznacznym przejawem jest spadek odchylenia standardowego i skośności odległości do najbliższego sąsiada (do zera w roku 1980). W połączeniu z niewielkim wzrostem średniej odległości NN świadczy to, że dodatkowe



Ryc. 25. Rozmieszczenie punktów pomiarowych w miesiącach i latach z najmniejszą oraz największą ich liczbą

**Tabela 4.** Statystyki lokalne liczby punktów danych w promieniu 35 km od węzła siatki interpolacyjnej 1 × 1 km oraz odległości od węzła do najbliższego punktu danych dla miesięcy i lat z najmniejszym oraz największym zagęszczeniem sieci pomiarowej

Parametr	Parameter	Liczba punktów - Number of points			
		02-56	08-75	1956	1975
<b>Liczebność</b>	Number of values	319 114 <sup>32</sup>	319 114	319 114	319 114
<b>Minimum</b>	Minimum	1	3	1	3
<b>Maksimum</b>	Maximum	106	95	94	95
<b>Średnia</b>	Mean	21	30	19,6	29,9
<b>Mediana</b>	Median	17	27	16	27
<b>1 kwartył</b>	First quartile	14	23	12	23
<b>3 kwartył</b>	Third quartile	22	33	22	33
<b>Odch. standardowe</b>	Standard deviation	13,7	11,2	12,8	11,2
<b>Skośność</b>	Skew	2,918	1,907	2,700	1,893
<b>Kurtoza</b>	Kurtosis	10,429	4,639	9,025	4,616
		Odległość - Distance (m)			
<b>Minimum</b>	Minimum	0	0	0	0
<b>Maksimum</b>	Maximum	34 589	26 663	34 589	26 663
<b>Średnia</b>	Mean	6 660	4 880	7 066	4 891
<b>Mediana</b>	Median	6 245	4 730	6 506	4 737
<b>1 kwartył</b>	First quartile	4 026	3 173	4 166	3 177
<b>3 kwartył</b>	Third quartile	8 817	6 389	9 341	6 402
<b>Odch. standardowe</b>	Standard deviation	3 533	2 328	3 907	2 337
<b>Skośność</b>	Skew	0,769	0,630	0,855	0,631
<b>Kurtoza</b>	Kurtosis	1,033	1,445	0,954	1,415

punkty pomiarowe lokowane były głównie tam, gdzie ich zagęszczenie do tej pory było najmniejsze, a co za tym idzie – gdzie były najbardziej od siebie oddalone. Ten sam skutek byłby oczywiście widoczny, gdyby jednocześnie (lub zamiast tego) likwidowano lokalne skupienia punktów. Proces ten oczywiście miał miejsce w odniesieniu do obszaru aglomeracji krakowskiej i warszawskiej (ryc. 25).

Zmiany w rozmieszczeniu posterunków opadowych na analizowanym obszarze w 25-leciu 1956-1980 postanowiono zilustrować przez porównanie miesięcy i lat z najmniejszą oraz największą ich liczbą (ryc. 25-32, tab. 4). Są to luty 1956 roku z 1890 punktami i sierpień 1975 z 2707 punktami pomia-

<sup>32</sup> Siatka interpolacyjna obejmowała całe terytorium lądowe Polski, Zalew Szczeciński, część polską Zalewu Wiślanego, Zatokę Pucką, a także strefę szerokości 2 km poza granicami (patrz dodatek X.6).

rowymi. Najmniejszą liczbą – 1768 – danych rocznych MSDO dysponowano dla roku 1956, największą – 2698 – dla roku 1975. Spojrzenie na mapy lokalizacyjne (ryc. 25) pozwala na szybkie uchwycenie najważniejszych różnic. Oprócz wyraźnego wzrostu liczby stanowisk daje się zauważyć:

- obecność w 1956 roku specjalnych sieci miejskich w aglomeracji warszawskiej i krakowskiej; widoczne są również wyróżniające się zagęszczenia stanowisk pomiarowych na zachód od Poznania, w zachodnich Sudetach, na Górnym Śląsku i w Zachodnich Karpatach;

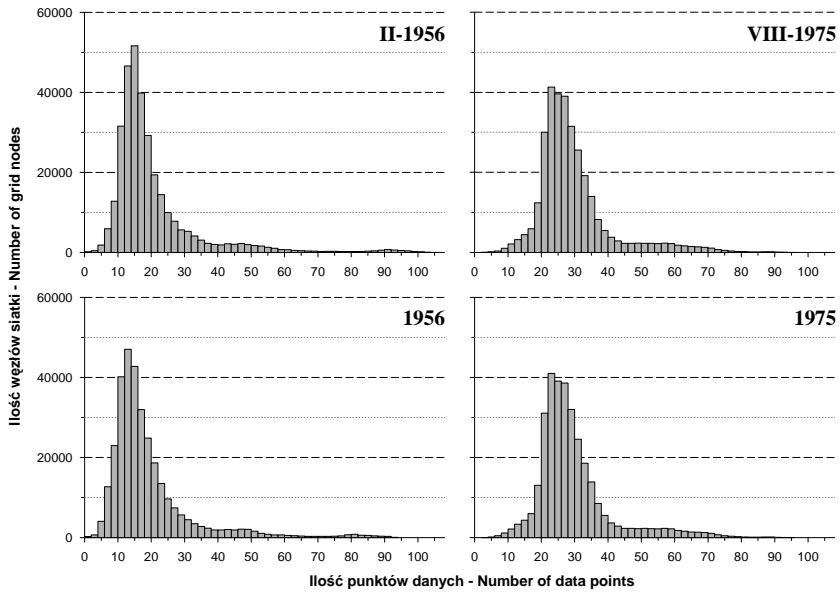
- bardziej równomierne rozmieszczenie w roku 1975 posterunków opadowych; zaznaczało się jedynie wyraźnie wyższe ich zagęszczenie w Sudetach, na Górnym Śląsku i w zachodnich Karpatach; w tym czasie sieć krakowska już nie istniała, a warszawska była znacząco mniejsza.

Bardziej subtelne różnice widoczne są na rycinach 26-32 i w tabeli 4. Posłużono się przy ich konstrukcji wynikami obliczeń statystyk lokalnych w ruchomym oknie dla siatki interpolacyjnej o rozmiarach 1 na 1 km (1 km<sup>2</sup>). Wykorzystano dwie miary gęstości przestrzennej danych: odległości od węzła siatki do najbliższego punktu danych oraz liczby punktów danych znajdujących się w promieniu 35 km od węzła (powierzchnia 3848,5 km<sup>2</sup>). Promień 35 km wybrano dlatego, ponieważ w terminach pomiarów o najmniejszej liczbie danych maksymalna odległość do węzła siatki interpolacyjnej wynosiła około 34 600 m (punkt siatki o współrzędnych w układzie 1992/19 : X = 786 000 i Y = 133 000 m; południowo-wschodni kraniec Polski, w Bieszczadach). Biorąc pod uwagę możliwy maksymalny błąd określenia lokalizacji najbliższego posterunku opadowego (Wetlina), przyjęta długość promienia (35 km) powinna gwarantować, że w sąsiedztwie węzła siatki interpolacyjnej zawsze będzie się znajdować co najmniej jeden punkt pomiarowy.

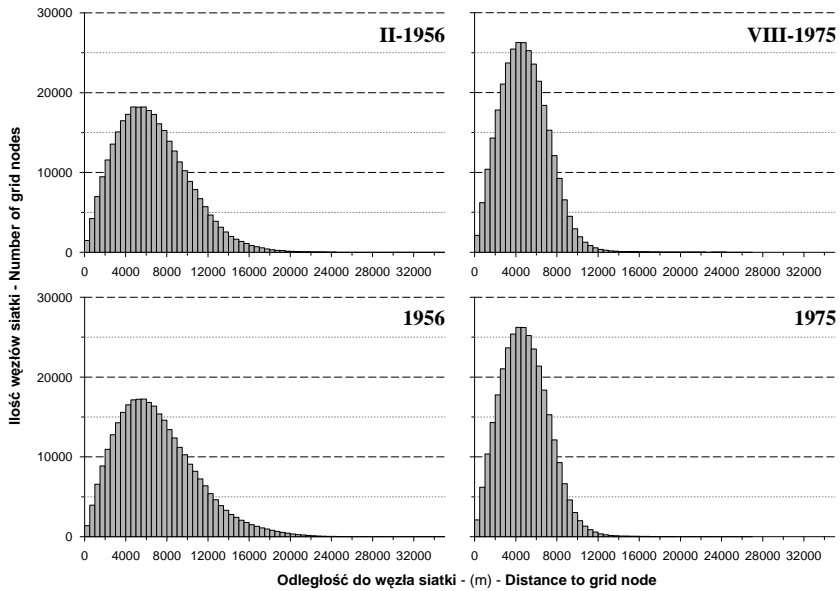
Najbardziej widoczna różnica pomiędzy porównywanymi terminami dotyczy równomierności pokrycia posterunkami pomiarowymi terytorium kraju. Najlepiej zmianę tę ilustrują histogramy (ryc. 26 i 27). W roku 1956 kontrasty były bardzo duże. Wyrażają się one w rozrzucie wartości ekstremalnych obu analizowanych parametrów, ich odchyleniu standardowym, skośności i kurtozie. Zróżnicowanie w roku 1956 mieści się w przedziale od węzła (o współrzędnych X = 571 000 m, Y = 220 000 m), leżącego 20 km na południe od centrum Krakowa i otoczonego w lutym 1956 roku przez 106 punktów pomiarowych, do 126 węzłów na krańcu Bieszczadów z jedynie jednym posterunkiem opadowym w promieniu 35 km. W roku 1975 amplituda wyraźnie zmalała i wynosiła 3-95. Zmieniła się także lokalizacja maksymalnej wartości – stało się nią otoczenie węzła (X = 516 000 m, Y = 210 000 m) usytuowanego w zachodniej części Beskidu Małego<sup>33</sup>. Minimalna gęstość

---

<sup>33</sup> Stosowane w niniejszej pracy nazewnictwo i podział regionalny Polski jest oparte na publikacjach: Kondracki (2000) oraz Kondracki i Richling (1994).



Ryc. 26. Frekwencja liczby punktów danych w promieniu 35 km od węzła siatki interpolacyjnej  $1 \times 1$  km dla miesięcy i lat w wieloletniu 1956-1980 z najmniejszym i największym zagęszczeniem sieci pomiarowej



Ryc. 27. Frekwencja odległości od węzła siatki interpolacyjnej  $1 \times 1$  km do najbliższego punktu danych dla miesięcy i lat w wieloletniu 1956-1980 z najmniejszym oraz największym zagęszczeniem sieci pomiarowej

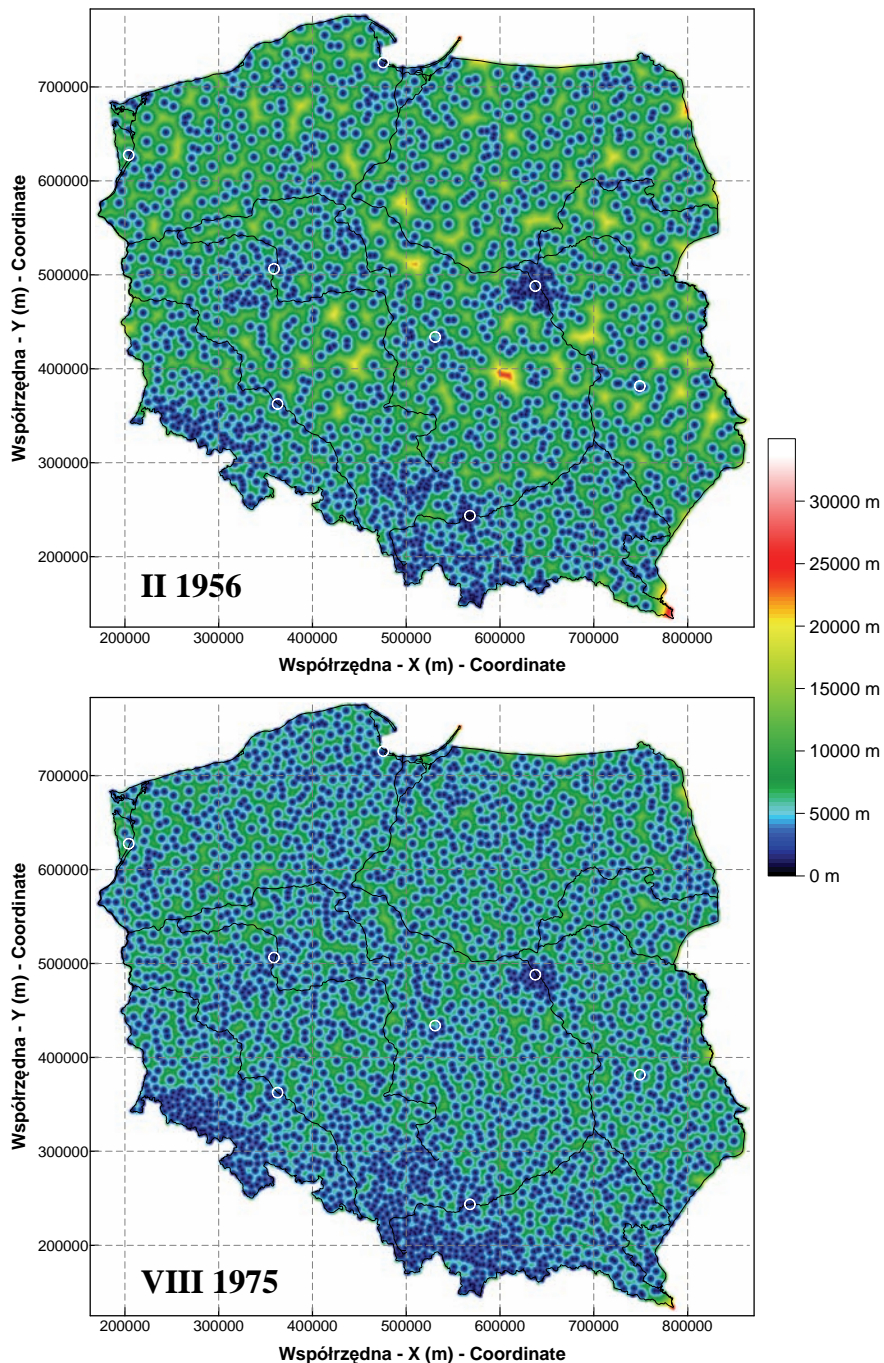
stanowisk – trzy – dotyczyła w roku 1975 tylko 17 węzłów siatki, z czego zaledwie dwa były położone w granicach Polski. W 1956 roku średnio w obrębie 35 km sąsiedztwa znajdowało się około 20 posterunków, w roku 1975 – 30. Przesunięciu uległ zakres międzykwartylowy: w roku 1956 połowa zmienności zawierała się w przedziale 12-22, w 1975 roku – 23-33. W ciągu 20 lat znacząco zmalało odchylenie standardowe, a skośność i kurtoza rozkładu przestrzennego punktów pomiarowych zbliżyły się do charakterystycznych dla rozkładu normalnego. Wraz z ogólnym wzrostem liczby stanowisk w latach 1956-1975 spadła wyraźnie odległość między węzłem siatki interpolacyjnej a najbliższym stanowiskiem pomiarowym; średnia zmalała z blisko 7,0 km do 4,9 km. Radykalnie zmniejszyła się maksymalna odległość (z 34,6 do 26,6 km) i odchylenie standardowe (z ponad 3,5 km do 2,3 km). Efektem mniejszego zróżnicowania było wyraźne zwiększenie kurtozy rozkładu odległości dla danych z roku 1975.


Zamieszczone poniżej mapy (ryc. 28-31) pozwalają precyzyjnie wskazać obszary o niskich i wysokich wartościach porównywanych cech, a przekonującą ilustracją zmian, jakie w 25-leciu 1956-1980 dokonały się w organizacji przestrzennej sieci pomiarów opadów atmosferycznych są mapy różnic (ryc. 32). Oprócz zasadniczych zmian w zagęszczeniu stacji, najbardziej widoczną cechą odróżniającą rok 1956 od 1975 są duże odmienności w obrazie „miesięcznym” i „rocznym”. W roku 1956 trwała najszybsza w całym analizowanym 25-leciu rozbudowa sieci; rok 1975 był pod tym względem raczej stabilny. Obrazów z sierpnia 1975 i całorocznego nie sposób gołym okiem odróżnić.

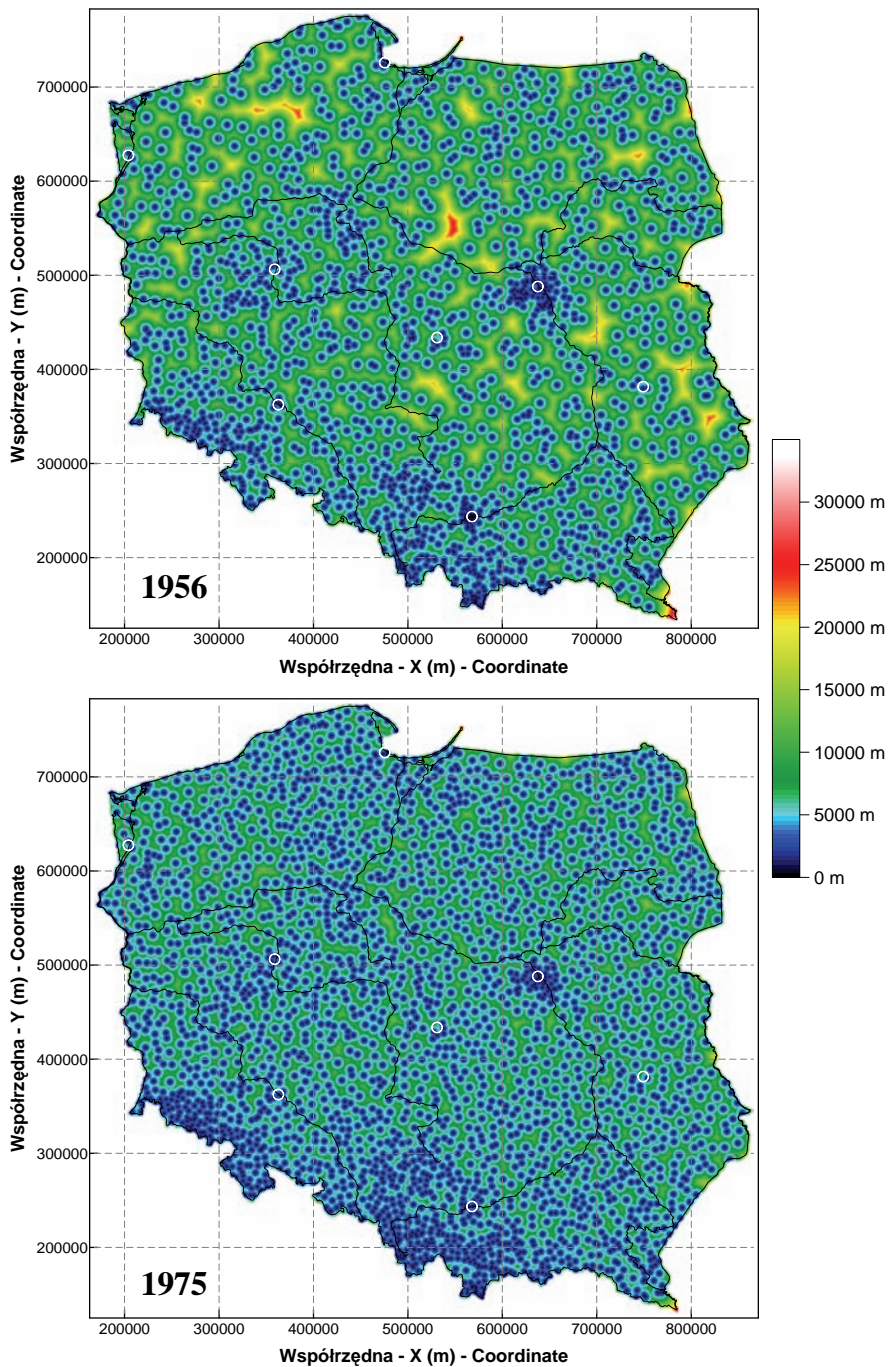
Mapy „odległości” dają kartograficzny obraz właściwości zidentyfikowanej za pomocą syntetycznych wskaźników i histogramów. W roku 1956 widoczne jest duże zróżnicowanie tej cechy, szczególnie w północnej, centralnej i wschodniej części kraju. Ekstremalnie duże odległości od węzłów siatki interpolacyjnej (> 20 km) występują nie tylko w strefach przygranicznych na wschodzie i północy, ale także na kilku powierzchniach w centrum kraju, na Wyżynie Lubelskiej i Pojezierzu Pomorskim. Sieć pluwiometryczna istniejąca w roku 1975 była bardziej jednorodna. Najbardziej wyróżniają się obszary niskich „odległości” w Sudetach i Karpatach Zachodnich oraz skupienie warszawskie.


Bardziej kontrastowy obraz dają mapy liczby stanowisk pomiarowych wokół węzłów siatki interpolacyjnej. Charakterystycznymi elementami rozmieszczenia punktów pomiarów opadów atmosferycznych w roku 1956 były, wspomniane już uprzednio, sieci miejskie w Warszawie i Krakowie, a także duże zagęszczenia (powyżej 40 punktów w promieniu 35 km) w obrębie Sudetów i Niziny Śląskiej oraz Górnego Śląska, części Jury Krakowsko-Częstochowskiej i Karpat Zachodnich. Wspomniane wyżej skupienia „górsko-wyżynne” były rozdzielone w 1956 roku obszarem o stosunkowo małej

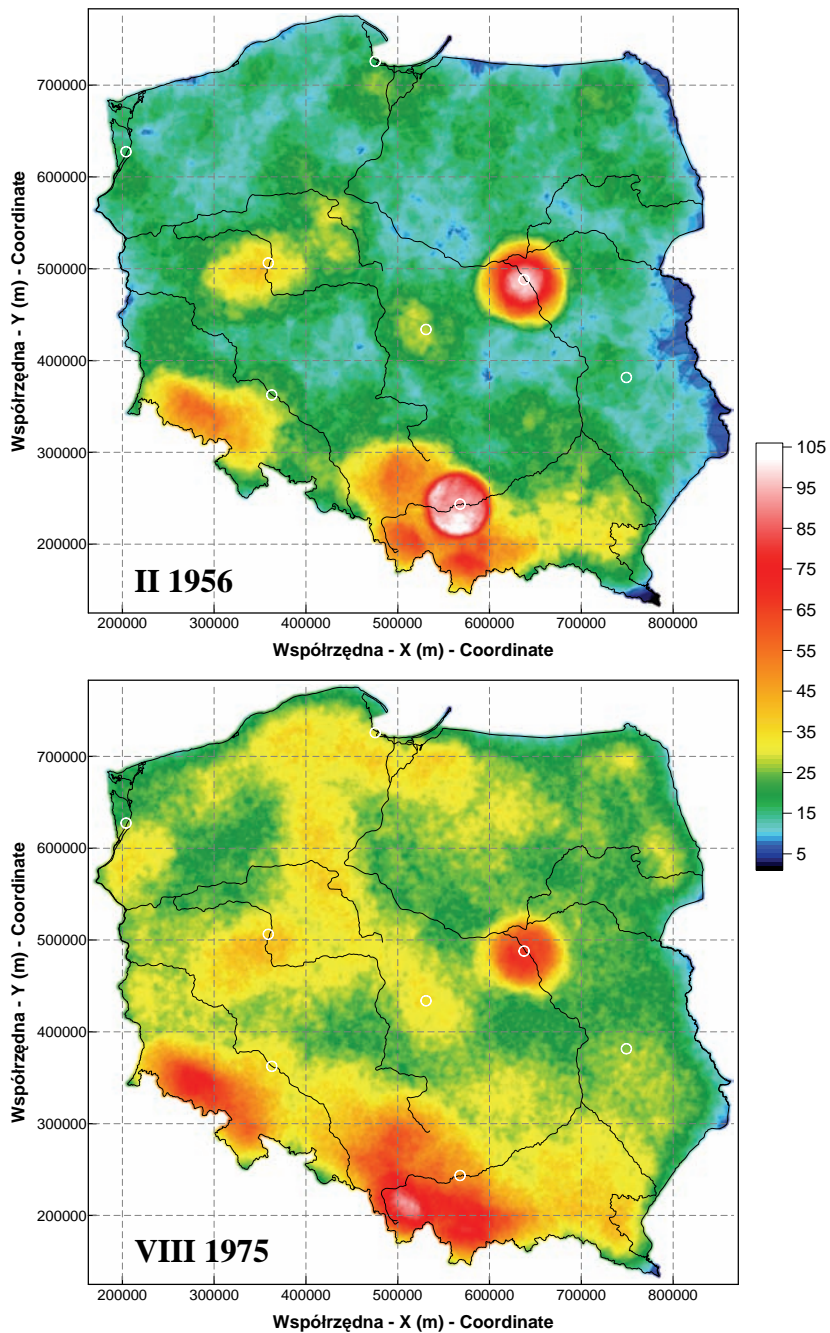





Ryc. 28. Odległość węzła siatki interpolacyjnej ( $1 \times 1$  km) od najbliższego stanowiska pomiarowego w miesiącu z najmniejszą (II-56: 1890) i największą ich liczbą (VIII-75: 2707) 

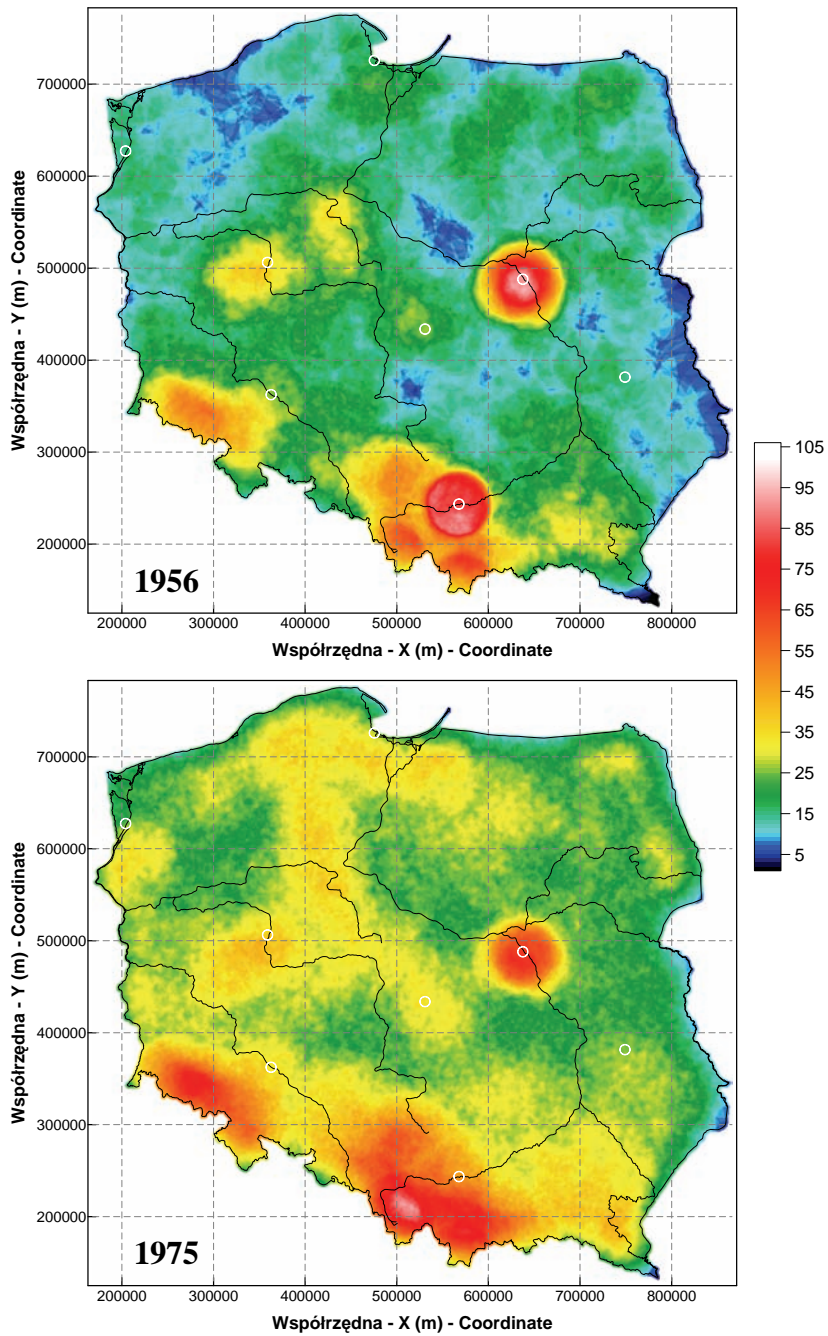


Ryc. 29. Odległość węzła siatki interpolacyjnej ( $1 \times 1$  km) od najbliższego stanowiska pomiarowego w roku z najmniejszą (1956: 1768) i największą ich liczbą (1975: 2698) 

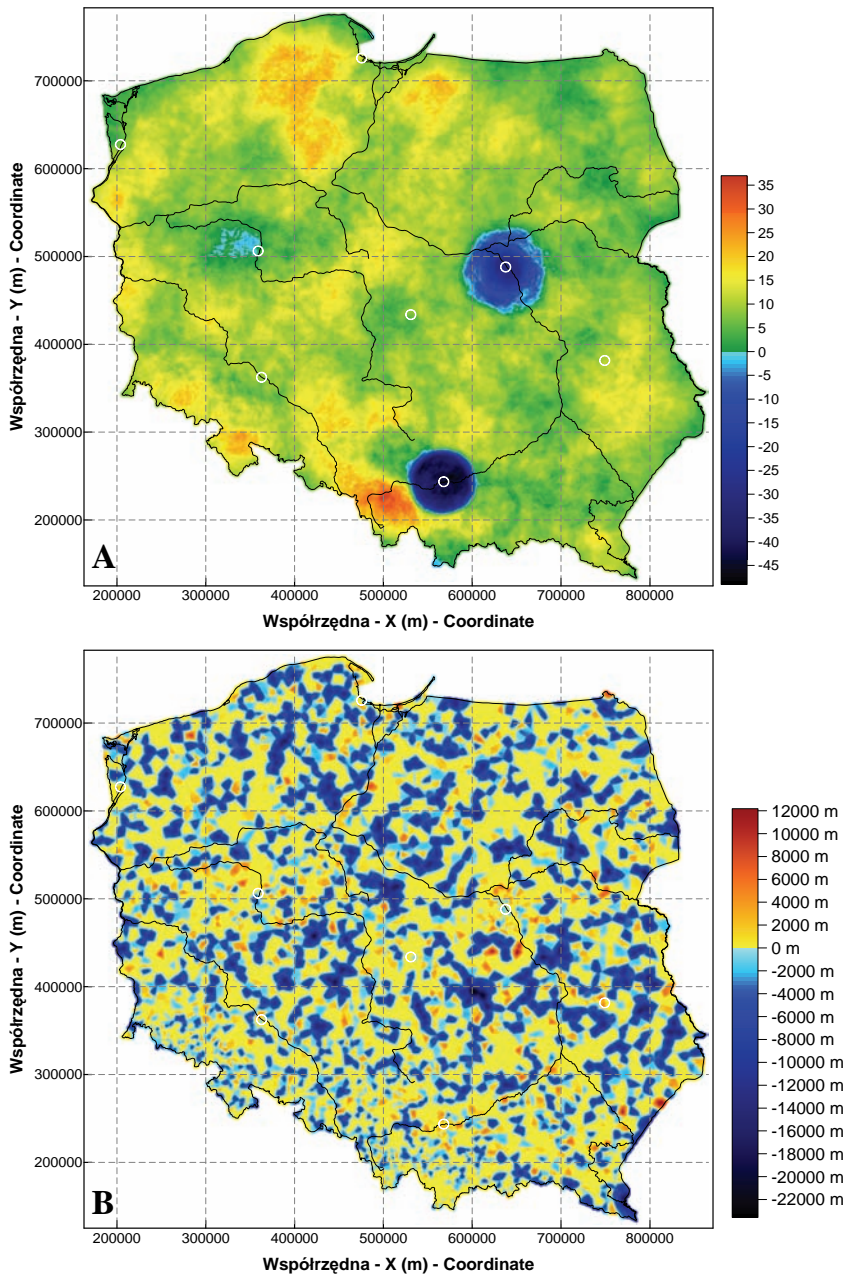


Ryc. 30. Liczba stanowisk pomiarowych w promieniu 35 km od każdego węzła siatki interpolacyjnej ( $1 \times 1$  km) w miesiącu z najmniejszym (II-1956: 1890) i największym ich zagęszczeniem (VIII-1975: 2707) 





Ryc. 31. Liczba stanowisk pomiarowych w promieniu 35 km od każdego węzła siatki interpolacyjnej ( $1 \times 1$  km) w roku z najmniejszym (1956: 1768) i największym ich zagęszczeniem (1975: 2698) 📄



Ryc. 32. Różnica liczby punktów danych w promieniu 35 km od węzła siatki interpolacyjnej  $1 \times 1$  km pomiędzy sierpniem roku 1975 (maksymalna liczba stanowisk pomiaru opadów w wieloletiu 1956-1980) a lutym 1956 (minimalna liczba stanowisk) – A. Różnica odległości węzła siatki interpolacyjnej  $1 \times 1$  km do najbliższego punktu danych pomiędzy sierpniem roku 1975 a lutym 1956 – B

gęstości stacji, usytuowanym na zachód od górnej Odry. Niższe, ale jednak znacznie przekraczające średnią (powyżej 30 punktów) zagęszczenia istniały wówczas także w centralnej Wielkopolsce, na Kujawach, w okolicach Łodzi i na przedgórzu wschodniokarpackim. Najbardziej zwarty i rozległy obszar skrajnie niskiego skupienia stanowisk pluwiometrycznych (< 10 w promieniu 35 km) występował wzdłuż granicy wschodniej kraju, na odcinku rzeki Bug. Obszary o podobnych wartościach analizowanej cechy układały się wąskim i nieco porozrywanym pasem także dalej na północ, wzdłuż wschodniej granicy, a także odcinka granicy z Obwodem Kaliningradzkim. Poza tym, w 1956 roku istniało szereg niewielkich „depresji” powierzchni gęstości rozrzuconych głównie wzdłuż brzegu morskiego i zachodniej granicy. Największą z nich stanowił jednak południowo-wschodni kraniec Polski. Obraz zagęszczenia danych rocznych w 1956 roku miał kilka charakterystycznych cech własnych, nie zaznaczających się wyraźnie na mapie z lutego. Były nimi duże rozproszenie stanowisk pomiaru opadów na północ od środkowej Wisły, nad górną Wartą (Wysoczyzna Bełchatowska) oraz w centralnej i północnej części Pojezierza Pomorskiego (Równina Białogardzka, pojezierza: Bytowskie i Drawskie).

W roku 1975 nie było już tak silnych kontrastów w zagęszczeniu stacji. Wyżynno-górski obszar dużego skupienia stanowisk pomiarowych był bardziej jednolity, choć w dalszym ciągu zaznaczały się poprzednio istniejące maksima. Nowością było pojawienie się elewacji gęstości (> 80) na obszarze na wschód od górnej Wisły (Beskid Mały, Kotlina Żywiecka, Pogórze Śląskie) oraz na pograniczu między Gorcami i Beskidem Żywieckim. Na Niżu w dalszym ciągu dodatnią anomalię stanowiła aglomeracja warszawska, ale maksymalne gęstości nie przekraczały tu już wartości 70. Poza tym, na większości obszaru niżowego w obrębie 35 km sąsiedztwa występuje 15-35 stanowisk pomiarowych. Strefy nieco większego zagęszczenia (około 40) były zlokalizowane w centralnej i południowo-zachodniej Wielkopolsce, Kujawach, Pojezierzu Krajeńskim, Pojezierzu Kaszubskim, Wysoczyźnie Polanowskiej, Pojezierzu Bytowskim oraz na wschód od Wisły, w obrębie Żuław, Wysoczyzny Elbląskiej i Równiny Warmińskiej. Najbardziej rozproszona sieć pomiarowa, jak poprzednio, była w Bieszczadach oraz wzdłuż wschodniej i północnej granicy kraju.

Największe spadki zagęszczenia stacji nastąpiły w obrębie i otoczeniu aglomeracji krakowskiej oraz warszawskiej, jako efekt zlikwidowania specjalnych miejskich sieci pluwiometrycznych (ryc. 32A). Niewielka redukcja nastąpiła również na obszarze na północny-zachód od Poznania i w Tatrach Zachodnich. Na pozostałej części kraju gęstość sieci pomiarowej wzrosła. Największy przyrost nastąpił na terenach wyżynnych, pogórskich i górskich, na południowy zachód od Krakowa (Płaskowyż Rybnicki, kotliny: Oświęcimska i Ostrawska, Pogórze Zachodniobeskidzkie, Beskid Mały).

W południowej części kraju wzrost większy od przeciętnego miał miejsce również w Kotlinie Kłodzkiej, Górach Bardzkich i Żółtych oraz w Kotlinie Jeleniogórskiej i Górach Kaczawskich. Duży zwarty obszar, na którym nastąpiło znaczące zwiększenie gęstości sieci stanowiły: północno-wschodnia część Pojezierza Pomorskiego, szczególnie zaś Wysoczyzna Damnicka i Polanowska, Pojezierze Kaszubskie, Pojezierze Bytowskie, Równina Charzykowska, Dolina Brdy i Wysoczyzna Świecka. Znaczące przyrosty wystąpiły również na wschód od Wisły, a w szczególności w obrębie Pojezierza Iławskiego i Równiny Warmińskiej. Mniejsze obszary dużego wzrostu zagęszczenia stacji pomiarowych zanotowano w zachodniej części Równiny Gorzowskiej i Pojezierza Myśliborskiego, w obrębie Wzniesień Gubińskich i Równiny Opolskiej.

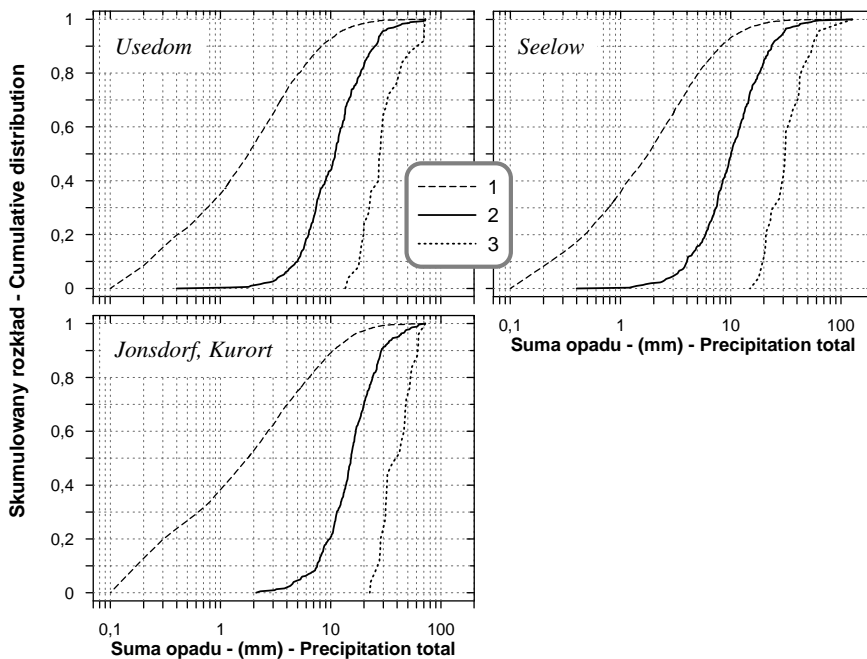
Zmiany odległości węzłów siatki interpolacyjnej do najbliższego stanowiska pomiarowego wykazują rozkład mozaikowy (ryc. 32B). W obrazie całego kraju najbardziej zwraca uwagę różnica pomiędzy południowo-zachodnią jego częścią a jego resztą. Na południowym zachodzie dominują raczej przyrosty odległości, a mozaika przeplatających się obszarów wzrostów i spadków jest bardzo „drobna”. Stanowi to efekt wysokiej już w latach pięćdziesiątych gęstości sieci na obszarach wyżynnych i górskich, która w kolejnych dziesięcioleciach była przede wszystkim optymalizowana pod kątem równomierności pokrycia. W obrębie pozostałej części kraju najbardziej znaczące są stosunkowo zwarte obszary przyrostów odległości na zachód od Poznania i w obrębie aglomeracji warszawskiej. Spadki dominują na obszarze na wschód i zachód od Dolnej Wisły, na południe od Warszawy, na Polesiu i Wyżynie Lubelskiej.

## **5. Maksymalne sumy dobowe opadów a źródłowe serie pomiarowe. Teoretyczne dystrybuanty danych**

Analizowane roczne i miesięczne MSDO stanowią podzbiory baz danych zawierających wszystkie zarejestrowane w poszczególnych stanowiskach pomiarowych sumy dobowe opadów. Relacje, jakie między nimi zachodzą przedstawiono na trzech przykładach zamieszczonych na rycinie 33. Kompletną dwudziestopięcioletnią serią sum dobowych opadów dla analizowanego wielolecia dysponowano jedynie dla kilkunastu stanowisk zlokalizowanych w przygranicznej strefie na terenie byłego NRD (patrz załącznik XII.1). Wybrano z nich trzy, które ukazują zróżnicowanie w układzie południkowym: stacja Usedom na wyspie Uznam na południowy zachód od Świnoujścia ( $H_s = 2$  m n.p.m.), Seelow przy ujściu Warty do Odry ( $H_s = 55$

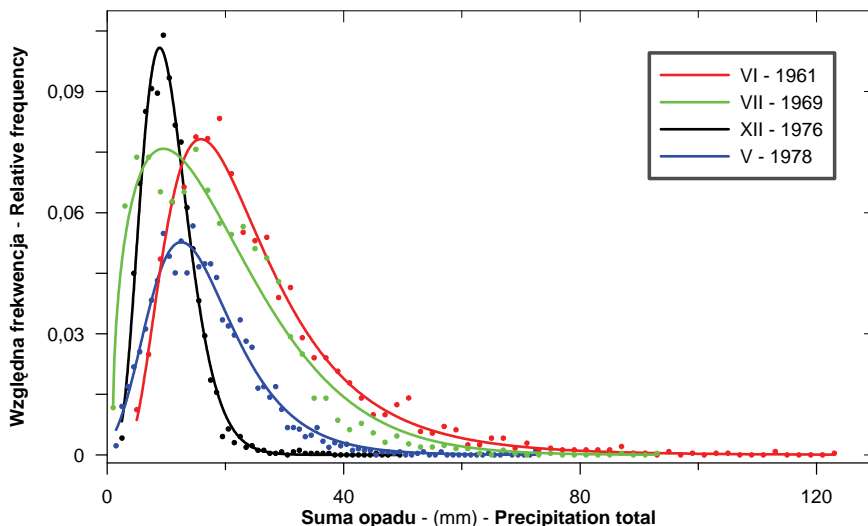
m n.p.m.) i Jonsdorf Kurort przy zbiegu trzech granic w okolicach Bogatyni (Hs = 460 m n.p.m.). Konsekwentne zmiany w tym układzie wykazuje tylko jeden parametr statystyczny sum dobowych opadów – ich odchylenie standardowe (odpowiednio: 4,95, 5,16 i 5,87 mm). Inne parametry swoje minimum albo maksimum mają w danych z Seelow: najniższą częstość dni z opadem (42,0% wobec 46,0% w Usedom i 57,8% w Jonsdorf), najniższą średnią sumę dobową (3,3, 3,4 i 3,8 mm), najwyższą skośność (6,72, 4,60 i 3,50), kurtozę (104,3, 38,1 i 19,8) i najwyższe maksimum (126,8, 72,4 i 72,1 mm). Biorąc pod uwagę częstość dni z opadem, miesięczne MSDO stanowią od 5,69 (Jonsdorf) do 7,83% (Seelow) całego zbioru sum dobowych. Ten sam parametr w odniesieniu do rocznych MSDO waha się od 0,47 do 0,65%.


Selekcja wykorzystywanych w niniejszej pracy danych miesięcznych MSDO wyeliminowała całkowicie wartości poniżej 0,4 mm (Usedom, Seelow) i 2,0 mm (Jonsdorf). Stanowiły one 17,3% (Seelow), 19,8% (Usedom) i 51,7% (Jonsdorf) całego zbioru sum dobowych (ryc. 33). Minimalne wartości rocznych MSDO w analizowanym wieloleciu wynosiły w Usedom 13,6 mm,



Ryc. 33. Porównanie dystrybuant empirycznych wszystkich sum dobowych opadów z 25-lecia 1956-1980 (1) z rozkładami miesięcznych (2) i rocznych (3) MSDO dla trzech stanowisk zlokalizowanych w strefie przygranicznej na obszarze Niemiec (patrz załącznik XII.1)





**Ryc. 34.** Przykłady czterech typów rozkładów statystycznych charakteryzujących zbiory danych MSDO. Punkty przedstawiają wartości empiryczne, zaś linie – najlepiej dopasowane teoretyczne funkcje gęstości prawdopodobieństwa: VI 1961 – logarytmiczno-normalny, VII 1969 – Weibulla, XII 1976 – gamma, V 1978 – wartości ekstremalnych (Gumbela I) 

w Seelow 15,0 mm, a w Jonsdorf Kurort 22,6 mm. Oznacza to, że z całego zbioru dobowych sum opadów „odrzucone” zostało odpowiednio 96,3, 96,9 i 98,1% wartości niższych. Istotne różnice pojawiły się również w charakterze rozkładów statystycznych dwóch pochodnych zbiorów danych. Najlepsze dopasowane dystrybuanty całego zbioru sum dobowych stanowiły bowiem bądź rozkład Weibulla (Usedom, Jonsdorf Kurort), bądź gamma (Seelow)<sup>34</sup>. Po wyselekcjonowaniu maksymalnych miesięcznych i rocznych sum dobowych ich rozkład na wszystkich stanowiskach jest logarytmiczno-normalny.

Dopasowywanie rozkładów empirycznych danych MSDO do najczęściej stosowanych w klimatologii dystrybuant teoretycznych (Pruchnicki 1987) wykazało, że większość analizowanych zbiorów najlepiej opisuje funkcja lo-

<sup>34</sup> Panorska i in. (2007) stwierdzają, na podstawie szczegółowej analizy setek serii dobowych sum opadów z obszaru Ameryki Północnej z okresu ostatnich pięćdziesięciu lat, że w większości przypadków prawdopodobieństwa ekstremalnych sum opadów nie należą w sposób wykładniczy, ale raczej potęgowej. Oznacza to, że rozkłady statystyczne tradycyjnie używane do ich modelowania (to jest wykładniczy, Weibulla, gamma, logarytmiczno-normalny) generalnie niedoszacowują prawdopodobieństwo ekstremów. Stopień tego odchylenia wykazuje zmienność regionalną i sezonową.

garytmiczno-normalna (ryc. 34). Aż 60% przypadków danych miesięcznych i 68% rocznych było bowiem zgodnych z tą funkcją. Dla 66 miesięcy (22%) i 8 lat (32%) najlepszy okazał się rozkład wartości ekstremalnych (funkcja Gumbela I). Poza tym, w odniesieniu do miesięcznych danych MSDO wykorzystywano jeszcze rozkład gamma (14%), Weibulla (3%) i wykładniczy (0,3%). Punktowe serie czasowe maksymalnych sum dobowych opadów podlegają, według Pruchnickiego (1987), właśnie logarytmiczno-normalnym regułom rozkładu frekwencji. Dominacja tego porządku w ujęciu przestrzennym nie może zatem stanowić zaskoczenia.

## **6. Charakterystyka danych źródłowych za pomocą globalnych statystyk opisowych**

Do opisu zbioru danych MSDO wykorzystano zarówno statystyki globalne (globalne, nieprzestrzenne), jak i statystyki najbliższego sąsiada (lokalne, przestrzenne, podrozdz. V.7). Drugie z wymienionych dotyczą zbioru różnic między zmierzoną w danym punkcie pomiarowym wartością MSDO a MSDO zarejestrowaną w najbliższej leżącym innym stanowisku. Dają one zatem pogląd na temat zmienności lokalnej analizowanej cechy. Statystyki lokalne są istotne z wielu względów. Między innymi Stach i Tamulewicz (2005b, 2005c) stwierdzili, że to głównie one wpływają na jakość estymacji pola sum miesięcznych i rocznych opadów. Trzeba też wyraźnie podkreślić, że między wartościami statystyk należących do obu grup nie ma znaczących zależności. Oznacza to, że znajomość statystyk globalnych danego zbioru danych przestrzennych nie daje podstaw do wnioskowania o jego zmienności lokalnej i na odwrót.

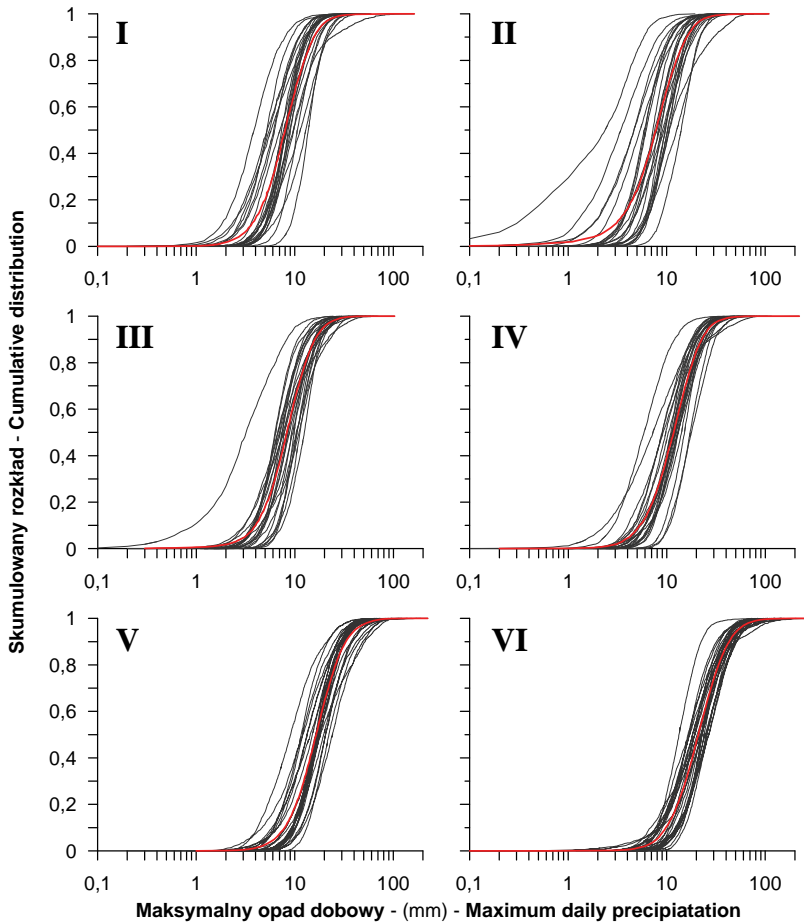
Pełne zestawienie statystyk opisowych analizowanych zbiorów miesięcznych i rocznych MSDO zamieszczono na dołączonym dysku DVD. Poniżej przedstawione zostaną – w formie tabelarycznej i graficznej – wybrane najważniejsze ich charakterystyki.

Wszystkie spośród 325 analizowanych zbiorów danych MSDO przedstawiono w postaci skumulowanych rozkładów na rycinach 35-38. Są tam również zamieszczone sumaryczne dystrybuanty dla poszczególnych miesięcy roku w całym wieloleciu i zbiorczy rozkład roczny. Ten sposób zestawienia umożliwi szerszą, niż za pomocą jedynie statystyk opisowych, charakterystykę danych.

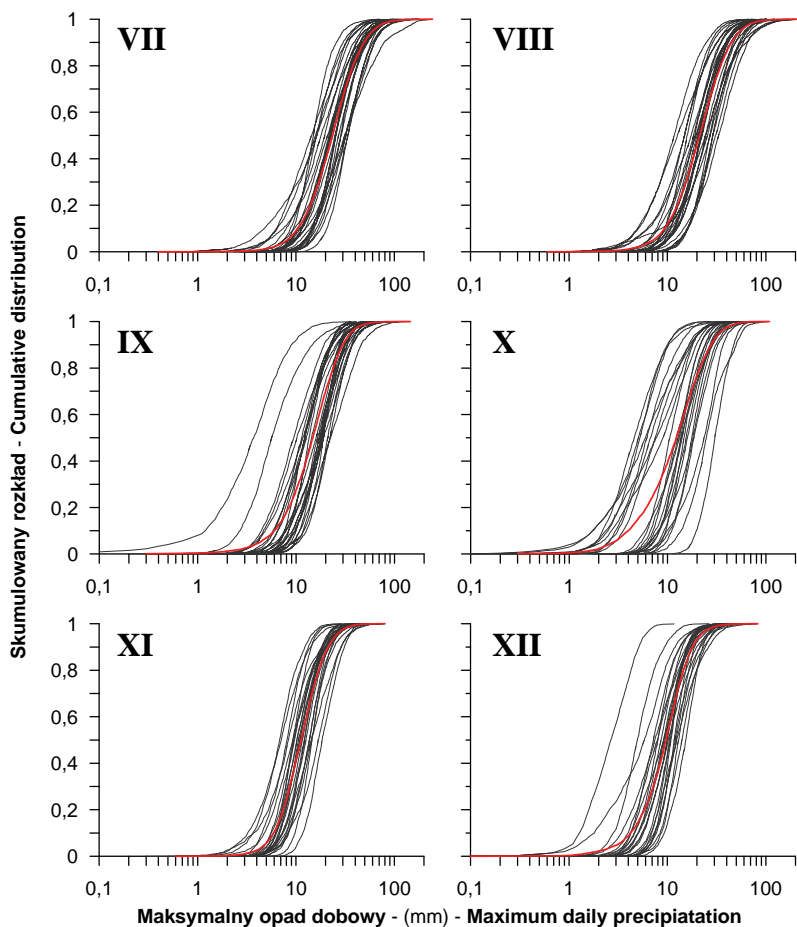
Porównanie wykresów krzywych skumulowanego rozkładu miesięcznych MSDO pozwala łatwo wskazać miesiące o stabilnym i zróżnicowanym „reżimie” wysokich opadów. Świadczy o tym zwartość wiązki krzywych. Do takich względnie jednorodnych w analizowanym wieloleciu należały

miesiące wiosenne od marca do czerwca, a wśród nich zwłaszcza czerwiec. Od głównej, bardzo wąskiej, wiązki krzywych odstają jedynie pojedyncze przypadki: marzec 1974, kwiecień 1974 i 1976, maj 1956 oraz czerwiec 1976. Są to wszystko miesiące posuszne, o bardzo niskich opadach. W marcu 1974 roku 95% zanotowanych maksymalnych opadów dobowych było na przykład mniejszych od 10 mm.

Drugą grupę stanowią miesiące, gdzie wiązki krzywych są szersze niż poprzednio, ale stosunkowo jednorodne. Należą do nich: lipiec, sierpień i listopad. Podobny charakter miała także zmienność MSDO we wrześniu,



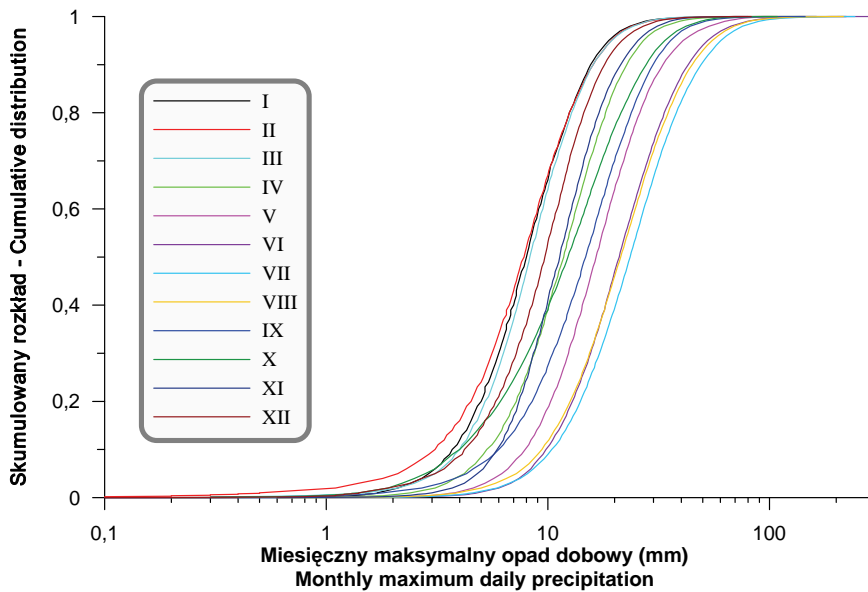
**Ryc. 35.** Dystrybuanty empiryczne MSDO dla poszczególnych miesięcy wielolecia 1956-1980: od stycznia do czerwca. Grubą czerwoną linią zaznaczono średnie rozkłady dla całego zbioru danych [4]



**Ryc. 36.** Dystrybuanty empiryczne MSDO dla poszczególnych miesięcy wielolecia 1956-1980: od lipca do grudnia. Grubą czerwoną linią zaznaczono średnie rozkłady dla całego zbioru danych [4]

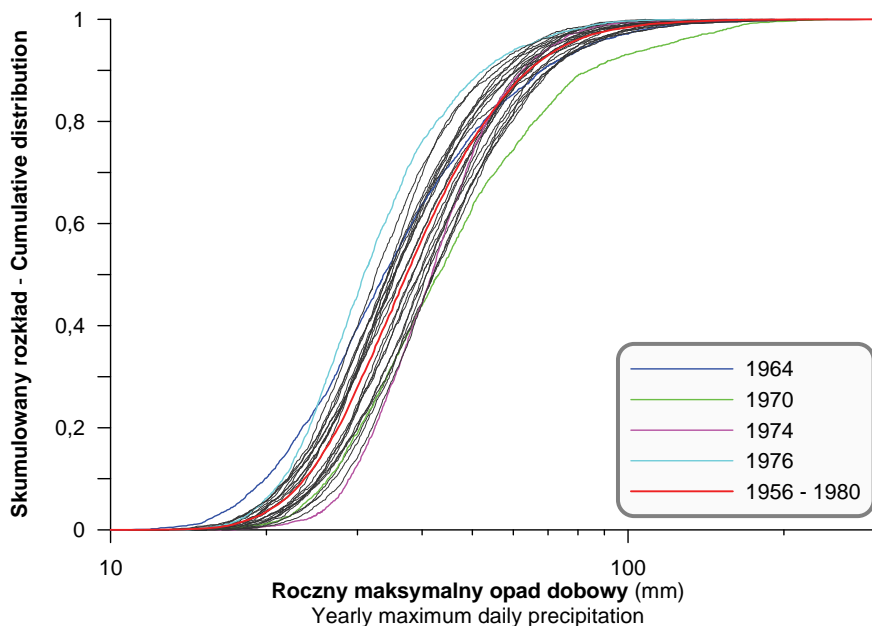
grudniu i styczniu, ale w wieloleciu notowano w grupie tych miesięcy po 2-3 anomalne, odstające znacznie od pozostałych, przypadki. Były to skrajnie suche miesiące: wrzesień 1959 i 1969, grudzień 1963, 1968 i 1972 oraz styczeń 1964.

Najbardziej zróżnicowany charakter mają rozkłady MSDO w lutym i październiku. W tym ostatnim miesiącu widoczna jest wręcz wyraźna dwudzielność wiązki, zwłaszcza dla dolnej połowy rozkładu. Grupa lat o niskich maksimach dobowych w październiku obejmuje: 1957, 1959, 1961, 1962, 1965, 1969, 1972, 1977, 1979. Jak widać z tego zestawienia, takie sytuacje w analizowanym 25-leciu powtarzały się dość systematycznie co 2-5 lat.



Ryc. 37. Skumulowane średnie rozkłady MSDO dla poszczególnych miesięcy w wieloleciu 1956-1980

Skumulowane krzywe rozkładów miesięcznych dla całego wielolecia pokazują, oczywiście, znaczącą zmienność sezonową (ryc. 37). Wyraźnie rysują się dwie grupy miesięcy o skrajnie odmiennym reżimie wysokich opadów i „strefa” przejściowa. Z jednej strony są to styczeń, luty i marzec, a z drugiej – czerwiec, lipiec i sierpień. Krzywe te nie są jednak idealnie równoległe względem siebie. Dla miesięcy zimowych charakterystyczna jest duża zbieżność częstości sum dobowych opadów wyższych od 10 mm, przy wyraźnie odmiennych dolnych fragmentach rozkładu. Najwyższą częstością bardzo niskich MSDO cechuje się luty:  $\frac{1}{4}$  zanotowanych wartości była mniejsza od 5 mm. W wypadku miesięcy o najwyższych MSDO sytuacja jest odmienna. To raczej częstość występowania niskich maksymalnych sum dobowych jest zbliżona. Dla opadów powyżej 8-9 mm na dobę zaznacza się wyraźna „przewaga” ich frekwencji w lipcu. Czerwiec i sierpień są pod tym względem bardzo podobne. Uwagę zwraca również charakterystyczne skrzyżowanie wieloletnich krzywych kumulacyjnych dla poziomu około 0,35 (35%) i sum opadów  $\leq 9$  mm dla kwietnia, października i listopada. Jest to odbicie wspomnianej wyżej odmienności rozkładu MSDO w październiku, z wysoką powtarzalnością występowania lat z bardzo niskimi opadami w tym okresie.

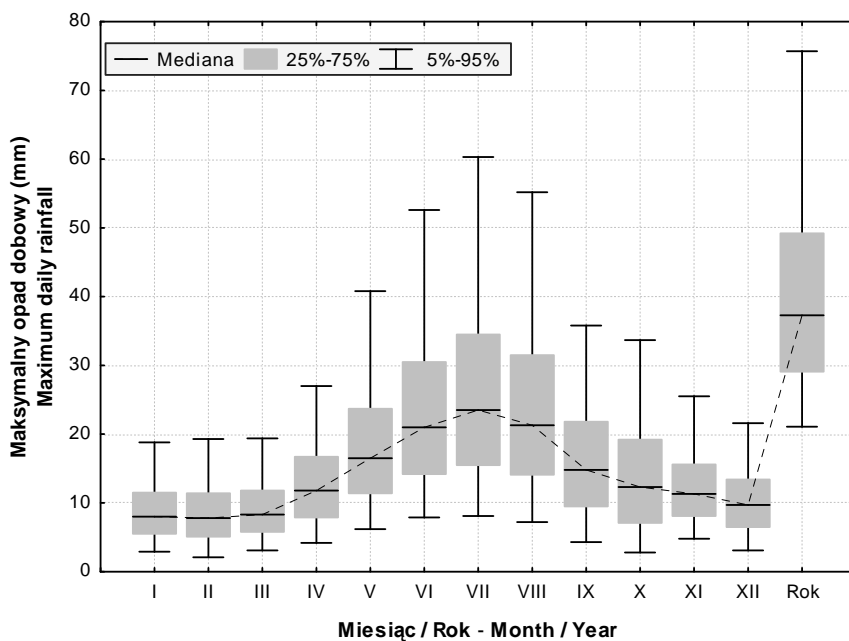


**Ryc. 38.** Skumulowane rozkłady rocznych MSDO dla poszczególnych lat w wieloleciu 1956-1980. Kolorem wyróżniono rozkład średni (dla całego zbioru danych) i odbiegające znacząco od przeciętnej lata ekstremalne

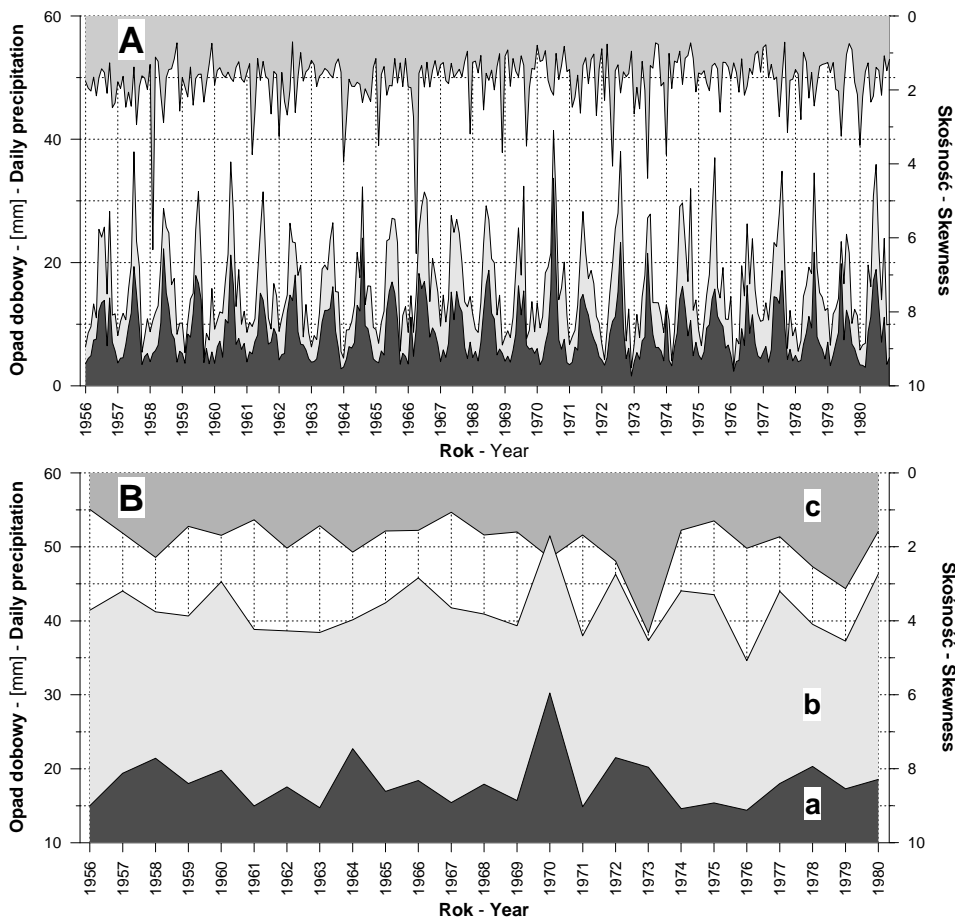
Na rycinie 38 zawierającej skumulowane roczne wartości MSDO dla poszczególnych lat analizowanego wielolecia najbardziej charakterystyczne jest wystąpienie zwartej wiązki krzywych, reprezentującej lata „przeciętne” i odstające pojedyncze przypadki „ekstremalne”. W odróżnieniu jednak od podobnych wykresów wykonanych dla poszczególnych miesięcy, roczne krzywe „ekstremalne” obejmują tylko część całego zakresu zmienności – odstają od głównej wiązki w pewnym odcinku. Najbardziej typowe przykłady to rok 1964 z anomalnie wysoką frekwencją MSDO mniejszych niż 25 mm, rok 1974, kiedy wyjątkowo rzadko notowano sumy dobowe opadów niższe od 35 mm, czy rok 1970 z ekstremalnie wysoką częstością sum przekraczających 40 mm na dobę (tab. 5). W pozostałym zakresie wysokości MSDO lata te mieszczą się w granicach „przeciętności”. Jedynym przypadkiem, który narusza tę regułę jest rok 1976, z konsekwentnie minimalnymi frekwencjami w każdej klasie wysokości MSDO.

Najniższe roczne MSDO w wieloleciu 1956-1980 zawierały się w przedziale od 10,0 (rok 1964) do 16,4 mm (rok 1980). Jest to zmienność stosunkowo niewielka wobec zróżnicowania najwyższych rocznych sum dobowych, które wynosiły od 108,1 mm w 1956 roku do 300,0 mm w roku 1973.

Miesięczne MSDO wykazują regularną i znaczącą cykliczność sezonową zarówno w ujęciu wieloletnim (ryc. 39, tab. 6), jak i przy porównywaniu kolejnych lat (ryc. 40). Dotyczy to głównie średniej (mediana, kwartyli) i odchylenia standardowego. Zmienność sezonową innych parametrów widać wyraźnie dopiero w ujęciu wieloletnim. Wartości miar centralnych i rozrzutu są minimalne w miesiącach zimowych. Od kwietnia następuje wyraźny wzrost, a maksimum jest osiągnięte w lipcu. W kolejnych miesiącach do grudnia następuje konsekwentny spadek. Przebieg ten nie ma charakteru symetrycznej sinusoidy. Gradient wzrostu wiosennego jest mniejszy niż spadku letnio-jesiennego. Przebieg krzywej absolutnych maksimumów w poszczególnych miesiącach jest nieco odmienny. Najniższe wartości zanotowano w listopadzie i grudniu (poniżej 100 mm na dobę). Zimą, a szczególnie w styczniu najwyższe MSDO przekraczały nieco 100 mm. W kolejnych miesiącach, od kwietnia do sierpnia, notowano opady przekraczające 200 mm, z maksimum w czerwcu wynoszącym 300 mm. Od września do listopada wieloletnie maksima miesięczne spadają od poziomu około 150 mm na dobę do około 80 mm.



Ryc. 39. Mediana, przedział międzykwartyłowy oraz przedział 90% miesięcznych i rocznych MSDO w wieloleciu 1956-1980



Ryc. 40. Wartości średnie (b), odchylenie standardowe (a) i skośność (c) MSDO w kolejnych miesiącach (A) i latach (B) wielolecia 1956-1980

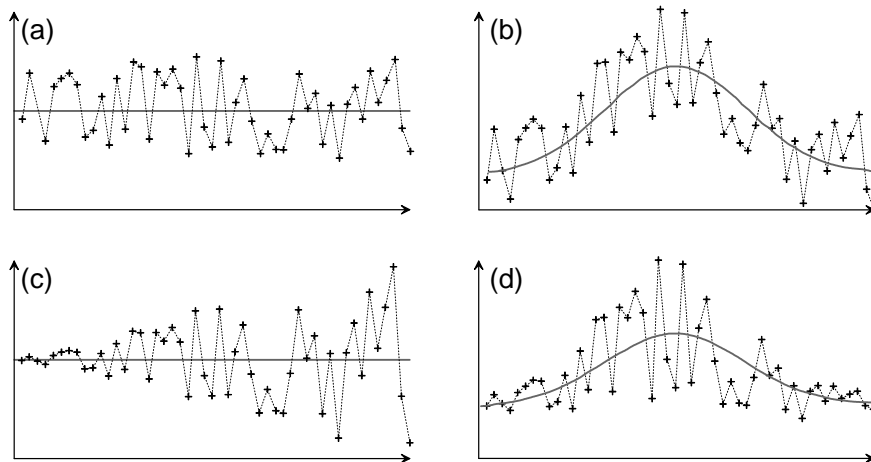
Charakterystyczną cechą zbiorów danych miesięcznych MSDO jest silna ( $R = 0,855$ ), wprost proporcjonalna relacja pomiędzy wartością średnią a odchyleniem standardowym (ryc. 42A). Jest to właściwość nazywana w statystyce heteroscedastycznością (gr. *hetero* – różny i *skedasis* – rozrzut, rozproszenie), kiedy sekwencja lub wektor zmiennych losowych charakteryzuje zróżnicowanie wariancji (Goovaerts 1997, Isaaks, Srivastava 1989). Dla zmiennych prawoskośnych zazwyczaj wariancja wzrasta wraz ze wzrostem średniej, dla zmiennych lewoskośnych – maleje (ryc. 41). Relacja między średnią a wariancją ma charakter funkcji kwadratowej, z czego wynika, że relacja pomiędzy średnią a odchyleniem kwadratowym powinna być liniowa. Przykładów takich zmiennych jest bardzo wiele zarówno w środowisku



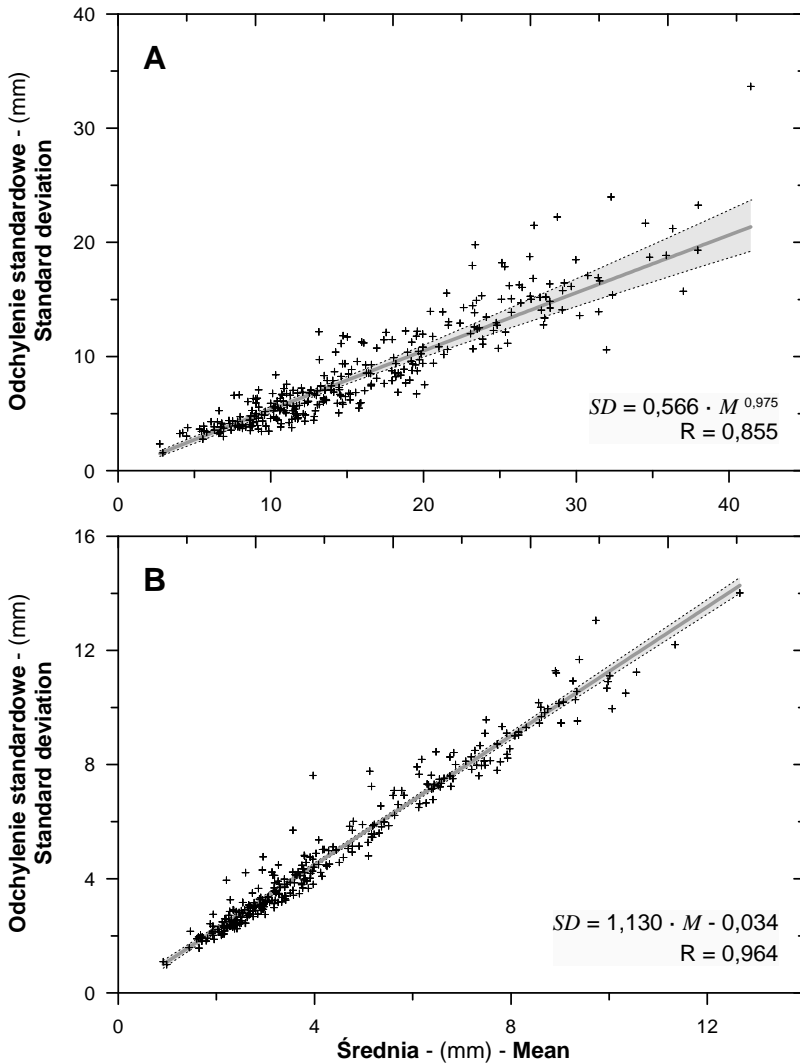
przyrodniczym, jak społeczno-gospodarczym. Kiedy porównujemy na przykład wysokości terenu z obszaru Wielkopolski (nizina) z notowanymi w Tatrach (góry), stwierdzamy, że rośnie nie tylko średnia wysokość ponad poziomem morza, ale także ich zróżnicowanie (przeciętne deniwelacje pomiędzy szczytami i obniżeniami). W praktyce statystycznej i ekonometrycznej (Hamilton 1994, Kennedy 1998, Salvatore, Reagle 2002) obecność heteroscedastyczności danych uważane jest za niekorzystne, ponieważ utrudnia identyfikację poprawnych modeli regresyjnych opartych na metodzie zwykłych najmniejszych kwadratów (ang. OLS – *Ordinary Least Squares*). W kontekście problemu niniejszej rozprawy jest to jednakże właściwość, która umożliwia ocenę relacji bezwzględnych wysokości MSDO na podstawie wiedzy o ich zmienności (por. rozdz. VI).

Omawiana zależność (ryc. 42A), bardzo zbliżona do liniowej, pokazuje, że odchylenie standardowe wynosi zazwyczaj około 56% wartości średniej. Od linii regresji najsilniej *in plus* odstaje przypadek z lipca 1970 roku, a *in minus* – z października roku 1974. Podobne, choć nieco słabsze, wprost proporcjonalne zależności występują między maksimum miesięcznych MSDO a ich odchyleniem standardowym ( $R = 77,8\%$ ), między skośnością a kurtozą ( $R = 77,2\%$ ) oraz między średnią a maksimum ( $R = 64,8\%$ ).

Do oceny możliwości występowania długookresowego trendu rocznych i miesięcznych statystyk MSDO z wielolecia 1956-1980 użyto nieparame-



**Ryc. 41.** Hipotetyczne profile danych w czasie bądź przestrzeni ilustrujące pospolite relacje pomiędzy lokalnymi średnimi i lokalną zmiennością (Isaaks, Srivastava 1989): (a) – średnia, reprezentowana przez poziomą linię oraz zmienność są stałe, (b) – trend średniej, zmienność pozostaje stała, (c) – średnia jest stała, a zmienność się zwiększa lub maleje, (d) – zarówno lokalne średnie, jak i lokalna zmienność rosną lub maleją



**Ryc. 42.** Relacja między średnią a odchyleniem standardowym MSDO w poszczególnych miesiącach wielolecia 1956-1980: A - statystyki globalne, B - różnice wartości opadów względem najbliższego sąsiada. Na wykresach zaznaczono krzywą regresji oraz jej 99% przedział ufności. Podano także wzór funkcji i współczynnik determinacji regresji

trycznego testu Manna-Kendalla, w jego modyfikacji uwzględniającej sezonowość (Hirsch i in. 1982). Jest on szeroko używany w naukach przyrodniczych i społeczno-gospodarczych ze względu na prostotę, stabilność oraz możliwość uwzględniania braków danych i wartości poniżej progu detekcji.

Tabela 5. Statystyki opisowe rocznych MSDO w latach 1956-1980

Rok	Minimum	Q1	Mediana	Q3	Maksimum	Średnia	SD	Skośność
Year	Minimum	First quartile	Median	Third quartile	Maximum	Mean	St. Dev.	Skewness
1956	14,6	30,1	38,1	50,1	108,1	41,4	15,0	0,988
1957	12,5	30,3	39,5	53,4	189,7	44,0	19,4	1,628
1958	12,6	28,0	35,2	46,6	181,9	41,2	21,4	2,287
1959	11,3	28,0	36,6	48,8	144,2	40,7	18,0	1,447
1960	12,5	31,7	40,7	54,3	181,8	45,3	19,8	1,689
1961	11,7	28,2	35,4	46,5	123,0	38,8	15,0	1,271
1962	13,1	26,8	34,4	45,4	177,7	38,6	17,6	2,038
1963	14,2	28,2	35,4	45,2	136,1	38,4	14,7	1,429
1964	10,0	25,4	33,7	48,0	191,8	40,1	22,7	2,143
1965	15,3	30,2	39,4	50,5	174,0	42,4	16,9	1,572
1966	10,8	32,2	42,0	55,0	221,0	45,8	18,4	1,555
1967	15,3	30,1	38,4	50,6	133,9	41,8	15,4	1,065
1968	14,3	28,2	36,6	49,3	190,6	40,9	17,9	1,678
1969	10,9	28,4	35,2	46,7	155,4	39,3	15,7	1,597
1970	15,4	32,3	43,0	60,5	242,9	51,5	30,2	2,298
1971	12,5	27,8	34,8	45,7	173,0	38,0	14,9	1,683
1972	14,6	31,9	41,9	54,8	223,5	46,3	21,5	2,389
1973	11,9	26,1	32,5	42,1	300,0	37,3	20,2	4,333
1974	15,4	33,8	41,7	50,4	163,0	44,1	14,6	1,550
1975	15,0	32,5	40,7	51,3	146,5	43,5	15,4	1,302
1976	13,0	25,1	30,8	39,4	166,1	34,6	14,4	2,040
1977	15,0	31,7	40,1	52,4	193,4	44,0	18,0	1,729
1978	12,8	27,1	34,2	45,0	210,8	39,5	20,3	2,538
1979	12,3	27,3	34,3	42,0	205,6	37,2	17,3	3,132
1980	16,4	33,4	42,0	54,2	168,3	46,4	18,6	1,580

W niniejszym opracowaniu posłużono się jego implementacją opracowaną przez Libiseller i Grimvall (2002, <http://www.mai.liu.se/~clib/welcome/PMKtest.html>). Roczne dane MSDO nie wykazują w analizowanym wieloleciu żadnych istotnych statystycznie tendencji. Nieco inaczej wygląda rezultat testu dla miesięcznych MSDO (tab. 7). Wskazuje on, że w latach 1956-1980 zmalały ich maksima, zróżnicowanie i skośność. Przegląd szczegółowych wyników dla poszczególnych miesięcy ujawnia, że zmiany dotyczą przede wszystkim miesięcy zimowych. Na przykład, maksima ( $r = -0,56$ ) i odchylenia standardowe ( $r = -0,55$ ) dla marcowych MSDO wykazują spadek na poziomie istotności  $p = 0,002$ .

**Tabela 6.** Statystyki opisowe MSDO dla poszczególnych miesięcy i całego roku w latach 1956-1980

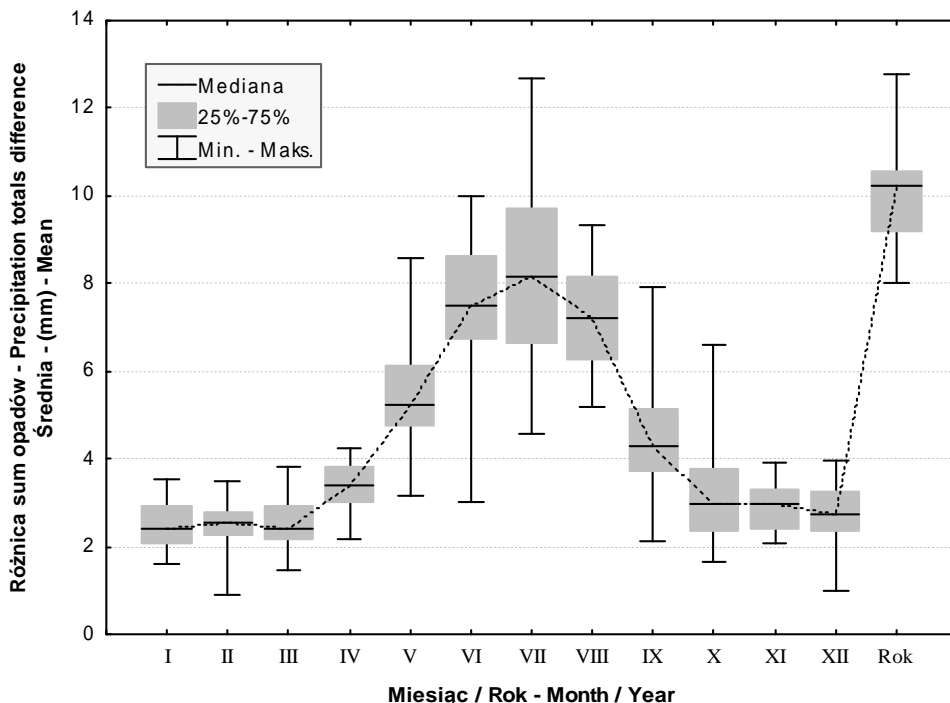
Miesiąc	Liczebność	Minimum	Maksimum	Średnia	Mediana	SD	Skośność	Kurtoza
Month	Count	Minimum	Maximum	Mean	Median	St. Dev.	Skewness	Kurtosis
<b>I</b>	61 512	0,1	163,0	9,1	8,0	5,6	3,050	32,410
<b>II</b>	61 615	0,0	109,0	8,9	7,8	5,8	2,130	14,080
<b>III</b>	61 954	0,0	103,0	9,4	8,3	5,4	1,850	10,720
<b>IV</b>	62 229	0,0	221,0	13,2	11,8	7,7	2,150	19,690
<b>V</b>	62 535	0,0	223,5	19,1	16,5	11,8	2,160	12,610
<b>VI</b>	62 580	0,1	300,0	24,5	21,0	15,4	2,360	15,400
<b>VII</b>	62 645	0,4	242,9	27,5	23,5	17,8	2,190	12,590
<b>VIII</b>	62 618	0,6	215,5	25,0	21,3	16,2	2,120	12,150
<b>IX</b>	62 679	0,0	145,0	16,8	14,8	10,4	1,680	9,170
<b>X</b>	62 580	0,0	108,1	14,5	12,3	10,1	1,530	6,940
<b>XI</b>	62 356	0,0	80,1	12,7	11,3	6,7	1,620	7,810
<b>XII</b>	62 182	0,0	82,7	10,7	9,7	6,1	1,730	10,020
<b>Rok</b>	61 940	10,0	300,0	41,6	37,3	18,8	2,213	9,715

**Tabela 7.** Ogólny wynik testu Manna-Kendalla na obecność liniowego trendu statystyk miesięcznych MSDO

Parametr	Parameter	n n	Stat. testu Test stat.	Odch. Std. Std. Dev.	Stat. MK MK stat.	Poziom <i>p</i> <i>p</i> -value
Minimum	Minimum	300	184	186,252	0,9879	0,3232
1Q (25%)	First Quartile	300	117	190,619	0,6138	0,5395
Mediana (50%)	Median	300	74	179,265	0,4128	0,6798
3Q (75%)	Third Quartile	300	48	168,408	0,2850	0,7756
Maksimum	Maximum	300	-333	139,026	-2,3952	<b>0,0166</b>
Zakres	Range	300	-350	135,056	-2,5915	<b>0,0096</b>
3Q - 1Q	Interquartile Range	300	6	148,400	0,0404	0,9677
Średnia	Mean	300	20	175,970	0,1137	0,9095
Odch. stand.	Stan. Deviation	300	-200	144,074	-1,3882	0,1651
Skośność	Skewness coef.	300	-374	171,422	-2,1818	<b>0,0291</b>

## 7. Charakterystyka danych źródłowych za pomocą lokalnych statystyk opisowych

Przedstawiona niżej analiza statystyk lokalnych ma istotne ograniczenia. O ile w każdym jej momencie dysponowano surowymi danymi pomiarowymi MSDO, to w drugim wymienionym zakresie używane oprogramowanie (Surfer wersja 8) takiej możliwości nie stwarzało. Oznacza to, że nie było bezpośredniego „dostępu” do zbioru różnic NN; uzyskano jedy-



Ryc. 43. Mediana, przedział międzykwartyłowy oraz zakres miesięcznych i rocznych średnich różnic MSDO w najbliższych stanowiskach w wieloleciu 1956-1980

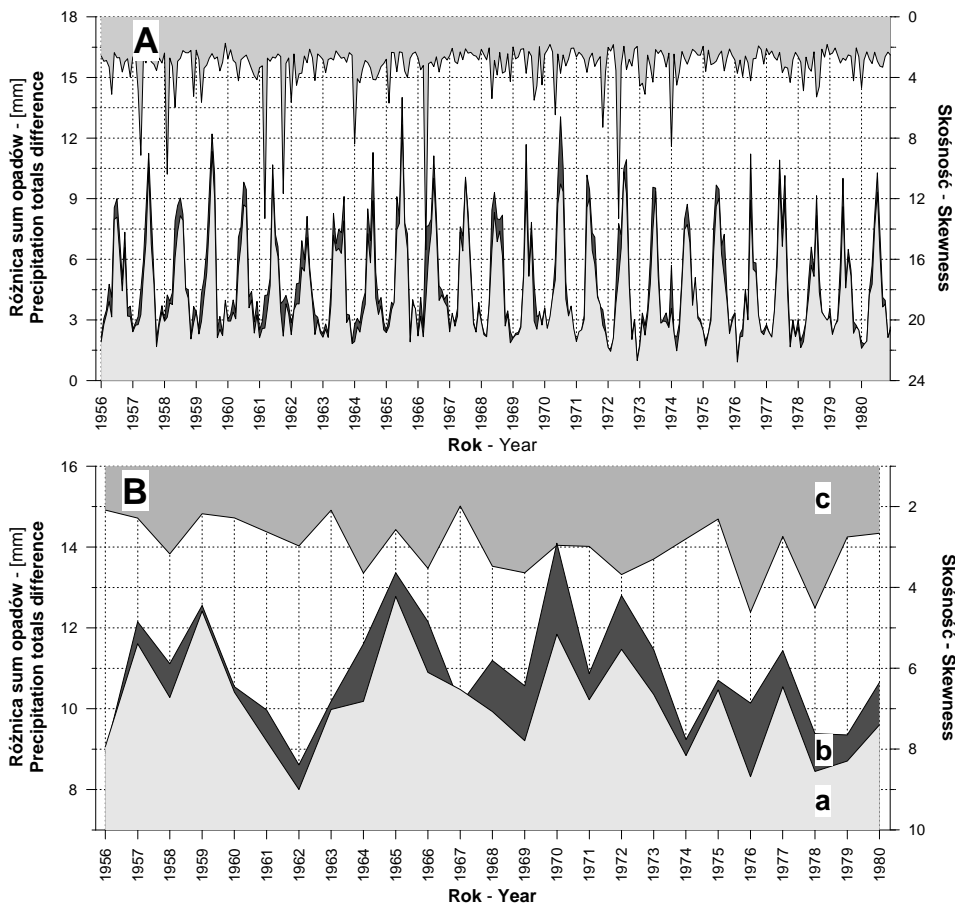
nie pełen zestaw ich statystyk opisowych dla każdego z 325 analizowanych zbiorów MSDO. Tak więc w niektórych z zamieszczonych poniżej rycin przedstawione są „statystyki ze statystyk”.

Zmienność sezonowa statystyk różnic MSDO w najbliższych stanowiskach jest jeszcze silniej zarysowana i dotyczy wszystkich, oprócz minimum (które w każdym miesiącu wynosiło oczywiście 0 mm) i skośności, parametrów (ryc. 43 i 44). Średnia różnica, dla całości danych miesięcznych wynosząca 4,4 mm, waha się od 2,5-2,6 mm w styczniu, lutym i marcu do 8,2 mm w lipcu. Rozkład wartości jest jednak zazwyczaj silnie prawoskośny (średnia skośność to 3,221, maksymalnie nawet 15,37 w kwietniu roku 1966), co powoduje spore rozbieżności między średnią a medianą. Przeciętna mediana dla miesięcy wynosi zatem jedynie 2,9 mm i zmienia się od 1,7 mm w styczniu oraz lutym do 5,4 mm w lipcu. Przebieg sezonowy jest bardziej symetryczny, niż miało to miejsce przy statystykach globalnych. Wyraźnie odróżniają się od siebie dwa półrocza: od października do marca i między kwietniem a wrześniem. W pierwszym z nich różnice NN MSDO

są niskie i mało zróżnicowane, w drugim zmieniają się znacznie z miesiąca na miesiąc. Średnie miesięczne odchylenie standardowe różnic dla wszystkich 300 przypadków wynosi 4,9 mm i waha się od 1,0 do 14,0 mm. Jego sezonowa zmienność mieści się między 2,7 mm w styczniu i lutym a 9,1 mm w lipcu. Bardzo charakterystyczny jest rozkład maksimum analizowanej cechy zbioru danych miesięcznych MSDO. Średnia miesięczna maksymalna różnica dla 25-lecia 1956-1980 wynosiła 53,0 mm i wahała się w przedziale od 7,6 mm w grudniu 1972 roku do 199,8 mm w maju tego samego roku. Również ta cecha ma wyraźny rozkład sezonowy, jednak w pewnych szczegółach jest on odmienny od opisywanych poprzednio. Podobnie jak poprzednio, najmniejsze maksymalne różnice notowane są w styczniu (średnio 30,3 mm), a największe w lipcu (średnio 86,9 mm), ale inne są charakterystyki jej zmienności. Miesiącem wyraźnie dominującym w tym aspekcie jest maj i nie wynika to jedynie z powodu wystąpienia jednej ekstremalnej wartości we wspomnianym powyżej maju roku 1972. Zmienność maksimum różnic również w czerwcu i sierpniu była większa niż w lipcu, a dwa pozostałe przypadki, kiedy cecha ta osiągnęła więcej niż 195 mm wystąpiły w kwietniu i sierpniu. Między majem a sierpniem zdarzają się w Polsce średnio więcej niż raz w roku sytuacje, kiedy na dystansie kilku/kilkunastu kilometrów różnice maksymalnych sum dobowych przekraczają 100 mm. W pozostałych miesiącach są to przypadki sporadyczne.

Zależność między średnią a odchyleniem standardowym, jeszcze silniejsza niż opisana powyżej, istnieje również w statystykach najbliższego sąsiada (ryc. 42B). Współczynnik kierunkowy regresji jest prawie dwukrotnie większy niż poprzednio, czyli odchylenie standardowe jest przeciętnie o 10% większe od średniej. Przypadki odstające są mniej widoczne niż poprzednio. Wyraźnie zaznaczają się jedynie dwa usytuowane powyżej krzywej regresji. Są to, ponownie, lipiec roku 1970 i czerwiec roku 1966.

Różnice NN rocznych MSDO są wyraźnie wyższe niż w każdym z miesięcy (ryc. 43 i 44). Średnie wahały się od 8,0 mm w roku 1962 do 12,8 mm w roku 1965 wynosząc przeciętnie 10,1 mm. Rozkład danych rocznych był nieco mniej skośny niż miesięcznych (2,953) i znacznie bardziej stabilny pod tym względem (1,985-4,611). Mediana jest zatem również wyraźnie niższa od średniej i wynosi 6,9 mm (5,2-9,0 mm). Odchylenie standardowe wynosiło przeciętnie 10,9 mm (8,6-14,1 mm) i tylko w latach 1956 oraz 1967 było niższe od średniej. Maksyma różnic NN rocznych MSDO zmieniały się od 58,3 mm w 1956 roku do 181,5 mm w roku 1978 i wynosiły przeciętnie 112,8 mm. Zastosowanie testu Manna-Kendalla nie ujawniło obecności żadnego istotnego liniowego trendu rocznych statystyk NN MSDO.



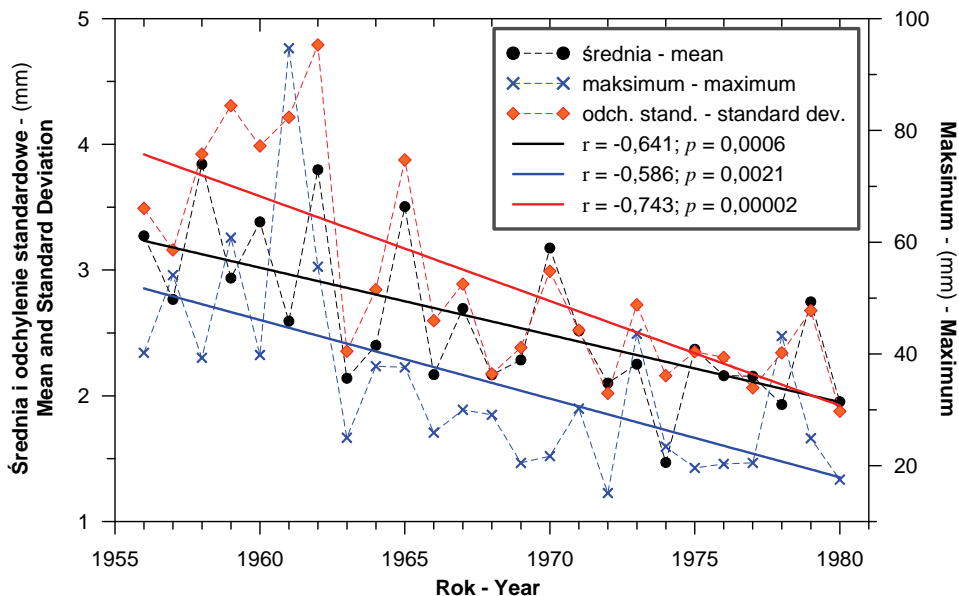
Ryc. 44. Wartości średnie (a), odchylenie standardowe (b) i skośność (c) różnic MSDO w najbliższych stanowiskach w kolejnych miesiącach (A) i latach (B) wielolecia 1956-1980

Ogólny wynik testu Manna-Kendalla (tab. 8) wskazuje na obecność w analizowanym zbiorze statystyk NN MSDO szeregu istotnych tendencji. Są to wszystkie trendy ujemne i dotyczą wartości trzeciego kwartyła, maksimum, zakresu, zakresu międzykwartyłowego, średniej, odchylenia standardowego i skośności. Przegląd statystyk M-K dla poszczególnych miesięcy wykazał, że rezultaty te są w głównej mierze uwarunkowane zmianami w owych charakterystykach w miesiącach zimowych i wiosennych, a zwłaszcza w marcu i lutym. Koresponduje to nieco z przedstawionymi poprzednio wynikami tego samego testu statystyk globalnych. W kilku przypadkach (marzec: maksimum, zakres, średnia, odchylenie standardowe; luty: odchylenie standardowe) trendy były istotne na poziomie  $p < 0,001\%$ . Jako przy-

kład umieszczono wykres pokazujący tendencję średnich i maksymalnych różnic NN miesięcznych MSDO oraz ich odchyłeń standardowych dla marca (ryc. 45). Proces zmniejszania się, w ujęciu wieloletnim, lokalnej zmienności ekstremalnych opadów to jeden z ciekawszych wyników uzyskanych w tej części niniejszej rozprawy. Ten aspekt wymagałby osobnej i bardziej wnikliwej analizy dokonanej na danych z dłuższego wielolecia.

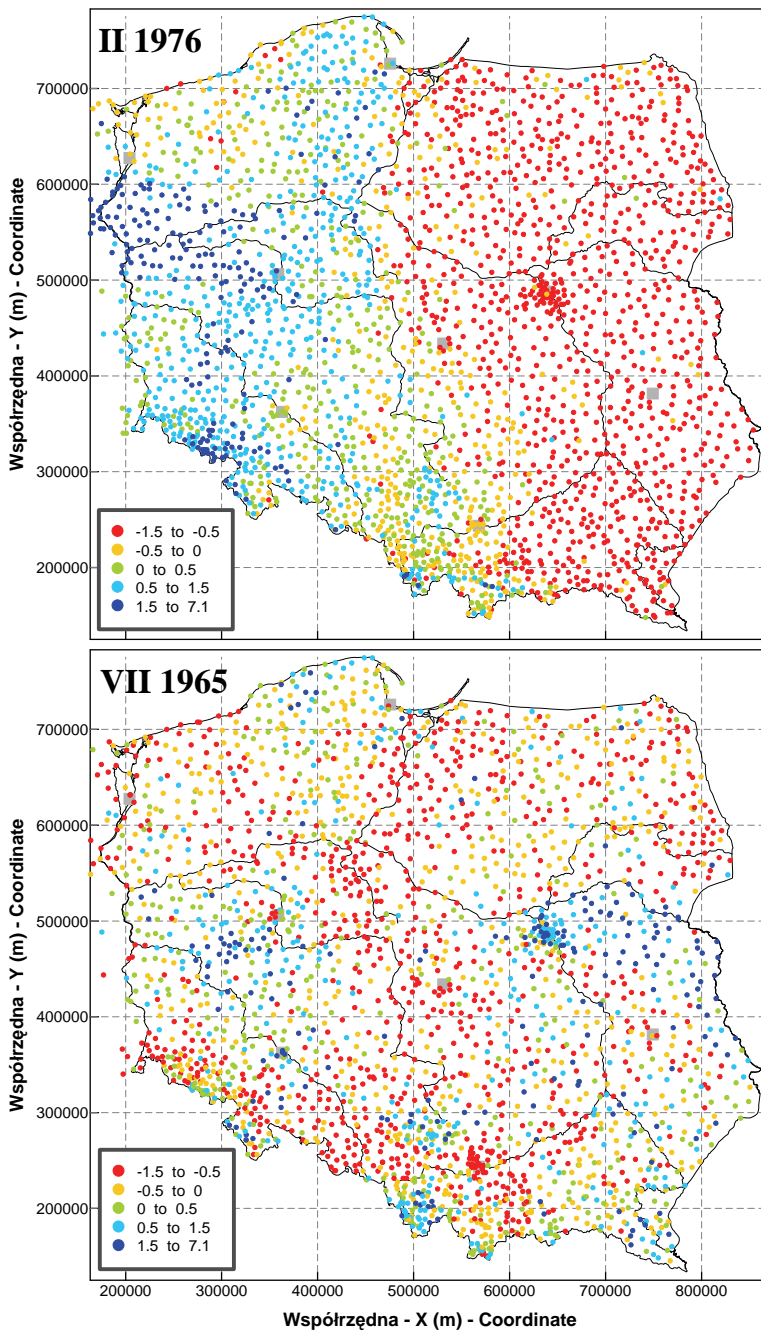
**Tabela 8.** Ogólny wynik testu Manna-Kendalla na obecność liniowego trendu statystyk różnic wysokości opadów w najbliższych stanowiskach miesięcznych MSDO

Parametr	Parameter	n	Stat. testu Test stat.	Odch. Std. Std. Dev.	Stat. MK MK stat.	Poziom <i>p</i> <i>p</i> -value
1Q (25%)	First Quartile	300	-251	187,271	-1,3403	0,1801
Mediana (50%)	Median	300	-323	195,647	-1,6509	0,0988
3Q (75%)	Third Quartile	300	-477	194,512	-2,4523	<b>0,0142</b>
Maksimum	Maximum	300	-515	161,531	-3,1882	<b>0,0014</b>
Zakres	Range	300	-515	161,531	-3,1882	<b>0,0014</b>
3Q - 1Q	Interquartile Range	300	-522	195,349	-2,6721	<b>0,0075</b>
Średnia	Mean	300	-540	196,109	-2,7536	<b>0,0059</b>
Odch. stand.	Stan. Deviation	300	-720	186,269	-3,8654	<b>0,0001</b>
Skóśność	Skewness coef.	300	-356	176,560	-2,0163	<b>0,0438</b>

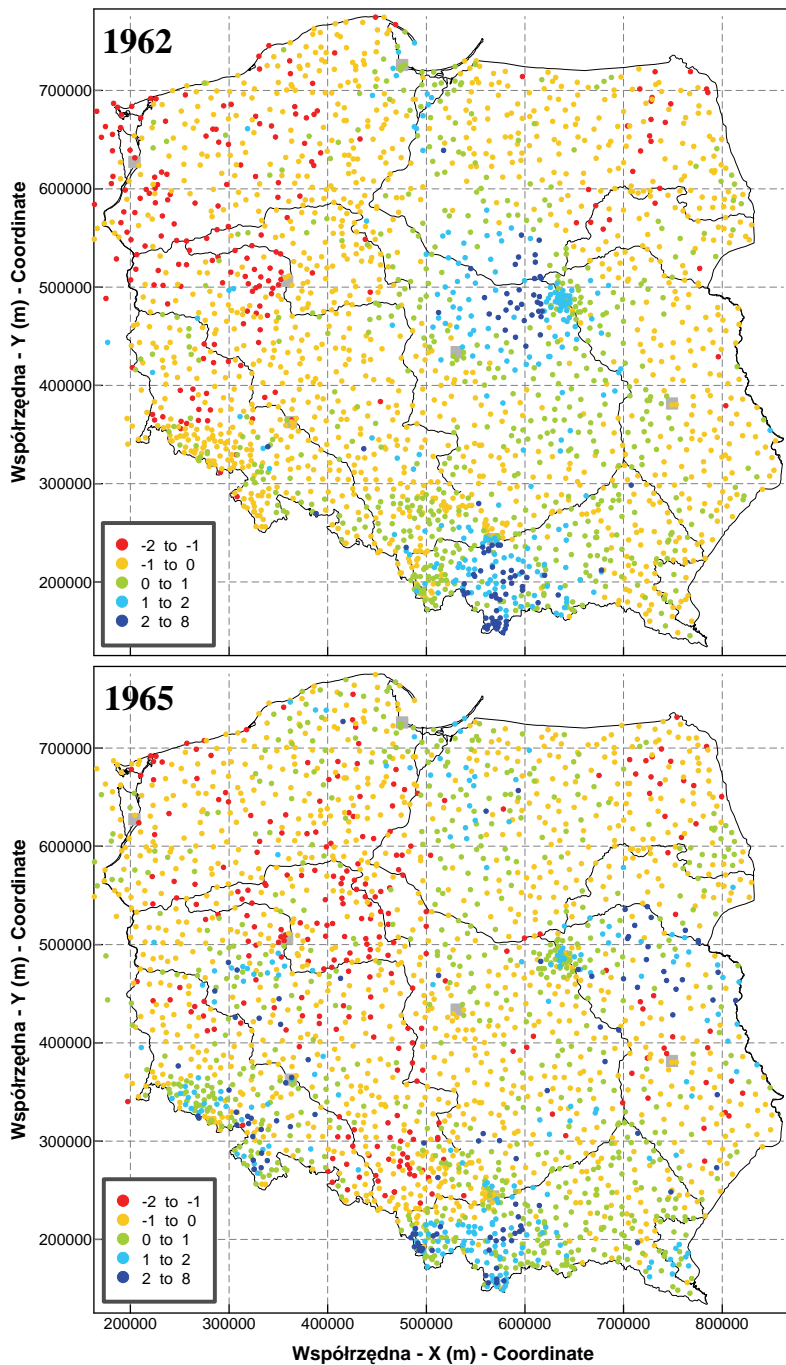


**Ryc. 45.** Trendy statystyk różnic względem najbliższego sąsiada miesięcznych MSDO dla marca w wieloletniu 1956-1980





Ryc. 46. Mapy punktowe standaryzowanych wysokości opadu dla miesięcy charakteryzujących się najmniejszą (II-1976) i największą (VII-1965) średnią zmiennością lokalną MSDO



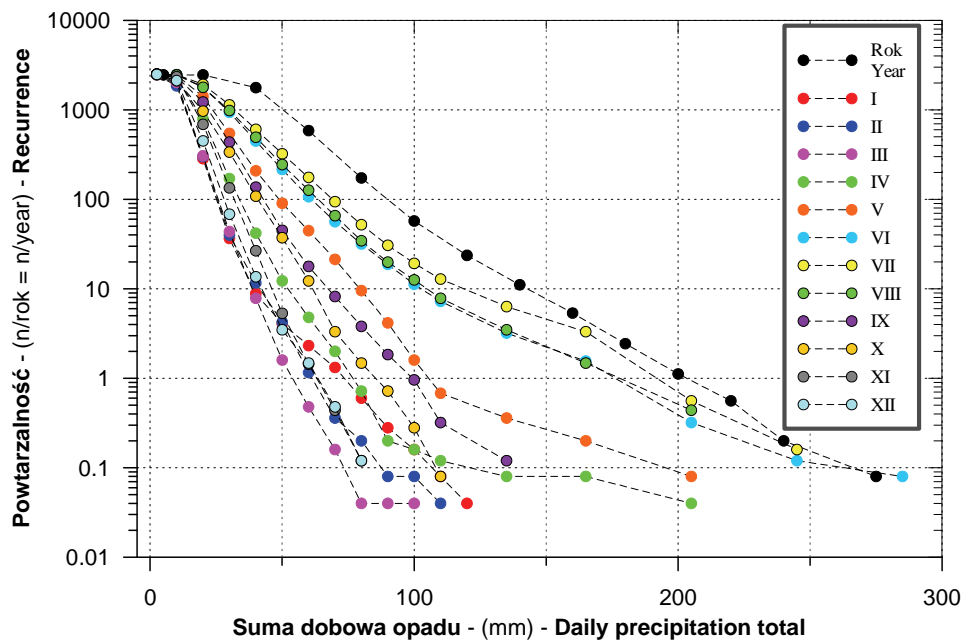
Ryc. 47. Mapy punktowe standaryzowanych wysokości opadu dla lat charakteryzujących się najmniejszą (1962) i największą (1965) średnią zmiennością lokalną MSDO

Na rycinach 46 i 47 znajdują się mapy punktowe pokazujące rozkład MSDO w miesiącach i latach o najniższych i najwyższych wartościach średnich różnic NN. Aby z obrazu wyeliminować wpływ bezwzględnej wysokości opadów, ich wartości zostały zestandaryzowane. Istota odmienności jest szczególnie wyraźnie widoczna na mapach miesięcznych MSDO. W lutym 1976 cały obszar na wschód od górnej Warty, Noteci i dolnej Wisły cechował się mało zróżnicowanymi skrajnie niskimi lub wręcz zerowymi opadami. Również na zachód od tej linii, mimo wyraźnie większych opadów, ich zmienność przestrzenna jest niewielka. W zasadzie, na całym terytorium kraju nie zaobserwowano sytuacji, aby punkty należące do skrajnych klas (czerwonej i ciemnoniebieskiej) ze sobą bezpośrednio sąsiadowały. Obraz z lipca 1965 jest wyraźnie odmienny. Dominuje mozaika niewielkich, nieregularnych powierzchni z opadami w zbliżonych klasach wysokości, a w wielu miejscach punkty ze skrajnie niskimi i wysokimi MSDO są położone bardzo blisko siebie. Tak silnych kontrastów nie widać na mapach rocznych MSDO z 1962 i 1965 roku. Mimo to różnice są, przy bardziej wnikliwym oglądzie, wyraźne, szczególnie w postaci liczby pojedynczych punktów w skrajnie wysokich i niskich klasach opadów występujących pośród stanowisk, gdzie zanotowano przeciętne MSDO (w zakresie  $\pm 1$  SD).

## 8. Relacje czasowe

Analizowany zbiór danych rocznych i miesięcznych MSDO przedstawiono także poprzez określenie powtarzalności sum opadów w określonych przedziałach ich wysokości (ryc. 48). Nie można jednakże zestawionych w taki sposób wartości w żaden sposób utożsamiać z rzeczywistą częstością występowania punktowych sum dobowych opadów o różnej wysokości. Jakkolwiek sieć pluwiometryczna na terenie naszego kraju była w wieloleciu 1956-1980 bardzo gęsta, to jednak z pewnością nie każdy opad został w niej zarejestrowany, a w wielu przypadkach lokalizacje maksymalnych wydajności stwierdzonych opadów znajdowały się gdzieś pomiędzy posterunkami pomiarowymi. Poza tym, specyfiką zbioru MSDO jest rejestrowanie dla poszczególnych stanowisk tylko jednej, najwyższej sumy dobowej w miesiącu lub roku. Z pewnością zdarzały się wielokrotnie sytuacje, że w danej lokalizacji równie wysokie opady mogły występować częściej niż jeden raz. Nie powinny jednak być to różnice diametralne. Wskazuje na to regularny raczej układ krzywych na rycinie 48. Wynika z nich, że średnio każdego roku w analizowanym wieloleciu notowano pięćdziesiąt kilka przypadków sum dobowych przekraczających 100 mm, 7 do 8 wystąpień opadów powyżej 150 mm i jednokrotnie sumy wyższe od 200 mm. Największy odsetek takich zdarzeń miał miejsce w lipcu – odpowiednie wartości częstości były na-

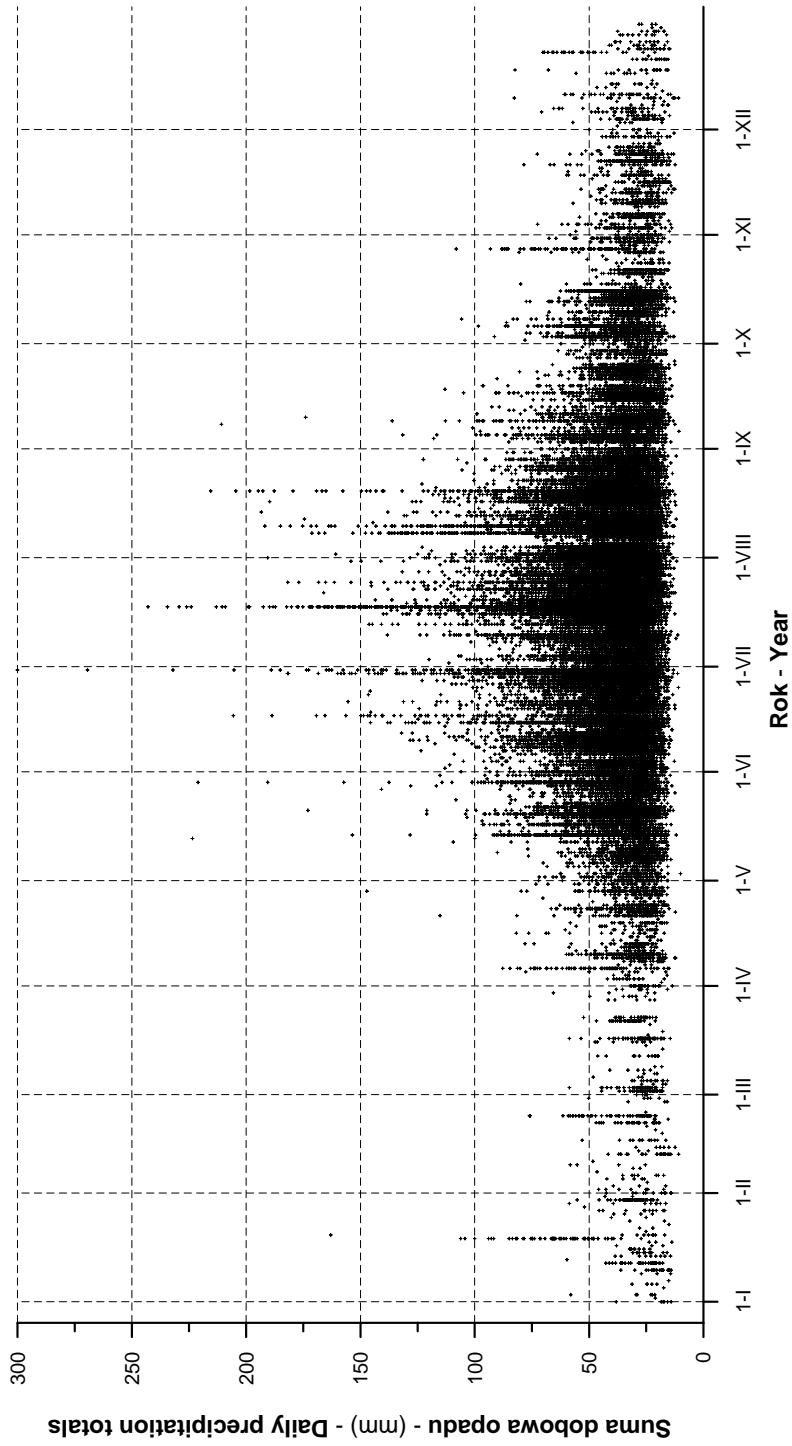
stępujące: 20, 4 do 5 i 0,7 do 0,8 (co 1,25 do 1,43 roku). Zbliżony do lipcowego rozkład częstości wysokich sum dobowych opadów mają także czerwiec i sierpień; pozostałe miesiące różnią się już znacznie pod tym względem. Bardzo charakterystyczne i zbliżone do siebie kształtem są krzywe powtarzalności dla kwietnia i maja. Wykazują one wyraźne załamanie i dla MSDO powyżej 110 mm (w maju) lub 90 mm (w kwietniu) przebieg tych krzywych zbliżony jest raczej do typowego dla miesięcy letnich. Poniżej tych wartości układają się one konsekwentnie w obrębie grupy miesięcy jesienno-zimowo-wiosennych. Najniższe częstości MSDO, szczególnie dla wartości powyżej 40 mm, występują w marcu. Ze względu na wspomniane wyżej ograniczenia posiadanego zbioru danych do oceny częstości punktowych MSDO na terenie naszego kraju, do obliczonych empirycznych wartości powtarzalności nie dopasowywano żadnych funkcji umożliwiających ich ekstrapolację. Z wykresów (ryc. 48) odczytano jedynie przybliżone wartości miesięcznych i rocznych MSDO o powtarzalności 0,1 (raz na 10 lat). Wynosiły one dla kolejnych miesięcy roku kalendarzowego: 106, 87, 73, 121, 195, 263, 262, 250, 139, 108, 81 i 81 mm. Roczna MSDO o takiej powtarzalności wynosi około 267 mm.



Ryc. 48. Powtarzalność (liczba przypadków rocznie) na terytorium Polski miesięcznych i rocznych MSDO o sprecyzowanej wysokości określona na podstawie danych z wielolecia 1956-1980

Zestawienie dat i wysokości rocznych MSDO (ryc. 49) pokazuje, że w Polsce takie zdarzenie może być zarejestrowane każdego dnia w roku. Prawdopodobieństwo jego wystąpienia do pierwszej dekady kwietnia i po połowie października jest jednak bardzo niskie. Wysokie sumy dobowe opadów częściej występują jesienią, niż zimą i na początku wiosny. Również wysokości rocznych MSDO są silnie zróżnicowane sezonowo. Opady o sumach przekraczających 100 mm na dobę przed połową maja i po połowie sierpnia występują bardzo rzadko. Przez 25 lat w okresie obejmującym blisko 75% roku zanotowano tylko 89 takich maksimów rocznych, o łącznej liczbie 892. Losowy charakter wysokich sum opadów dobowych widoczny jest nawet w 25-letniej serii obserwacyjnej dla przekraczającej 2400 stanowisk sieci obserwacyjnej. Na rycinie 49 wyraźnie zaznaczają się zarówno pojedyncze opady rejestrowane w setkach stanowisk, jak i kilkudniowe okresy, w których takie zdarzenie przez 25 lat ani razu nie wystąpiło. Do tej pierwszej grupy należą między innymi opady z 18 lipca 1970 roku, kiedy maksimum roczne zostało zanotowane w 742 stanowiskach, 8 sierpnia 1978 roku (558 stanowisk), 10 sierpnia 1964 (468 stanowisk), 14 maja 1962 roku (573 stanowisk) i kilka innych. „Dziury” w wykresie w okresach wysokiej częstotliwości występowania rocznych dobowych maksimów opadowych są widoczne na przykład dla przedziału 5-7 sierpnia (tylko 623 przypadki), czy 1-2 lipca (368 przypadków). Na uwagę zasługuje również najdłuższy, bo 4-dniowy okres pomiędzy 24 a 27 marca, w trakcie którego przez 25 lat w żadnym punkcie pomiarowym nie zanotowano sumy dobowej opadu, która byłaby zaklasyfikowana jako najwyższa w roku.

W tabeli 9 zestawiono 18 przypadków sum dobowych opadów przekraczających 200 mm, a zanotowanych w wieloleciu 1956-1980. Aż 12 wystąpiło w trakcie jedynie dwóch epizodów, które miały miejsce w zachodnich Karpatach: 18 lipca 1970 i 30 czerwca 1973 roku. Najwyższy opad dobowy, wynoszący 300 mm, zarejestrowano w tym drugim terminie na Hali Gąsienicowej w Tatrach. Jak wynika z zestawienia dokonanego przez Prokopa (2007), był to prawdopodobnie najwyższy opad dobowy na terenie naszego kraju w całym okresie obserwacji instrumentalnych. Jednakże, nie jest to, jak można by sądzić, jakiś poziom maksymalny, wynikający z położenia geograficznego Polski. Blisko bowiem granic naszego kraju zanotowano sumy dobowe wyraźnie wyższe (Prokop 2007). Było to bowiem 353 mm w Zinnwald-Georgenfeld (Saksonia, Niemcy, 12/13.08.2002 r.), oraz 345,1 mm w Nova Louka (Góry Izerskie, Czechy, 30.07.1897 r.). Z szacunków maksymalnych prawdopodobnych opadów (ang. *probable maximum precipitation*) na terenie Czech o różnym czasie trwania, wykonanych przez Rezacovą i in. (2005), wynika, że w kilku przygranicznych górskich i pogórskich obszarach punk-



**Ryc. 49.** Daty i wysokości rocznych MSDO zarejestrowanych na obszarze Polski i w strefie przygranicznej Niemiec w wieloletiu 1956-1980 (61 940 przypadków)

**Tabela 9.** Zestawienie rocznych MSDO przekraczających 200 mm, które zarejestrowano na terytorium Polski w latach 1956-1980

Stanowisko	Data	Współrzędne			Opad (mm/doba)
		X <sup>35</sup> (m)	Y (m)	Hs (m n.p.m.)	
Station name	Date	Easting (m)	Northing (m)	Elevation (m a.s.l.)	Precipitation (mm/day)
Hala Gąsienicowa	73-06-30	572 600	154 100	1 520	300,0
Hala Kondratowa	73-06-30	570 200	154 000	1 333	269,4
Leskowiec	70-07-18	532 200	214 800	870	242,9
Stańcowa	70-07-18	538 400	187 100	850	234,4
Kasprowy Wierch	73-06-30	571 400	152 200	1 991	232,0
Magurka	70-07-18	509 500	212 900	900	229,3
Zubrzyca Górna	70-07-18	545 700	189 000	700	226,0
Szczyrk	70-07-18	502 300	205 400	520	224,0
Wądołki Borowe	72-05-14	715 900	569 800	130	223,5
Niepołomice	66-05-29	587 000	243 200	194	221,0
Wapienica, zaporą	72-08-21	497 500	212 900	480	215,5
Lipowa	70-07-18	505 900	199 900	532	213,0
Zawodzie	78-09-08	659 400	615 800	120	210,8
Istebna-Młoda Góra	70-07-18	490 200	190 600	840	210,5
Bieńkówka	70-07-18	555 100	211 300	475	209,1
Wałbrzych	79-06-17	310 700	325 600	490	205,6
Hala Ornak	73-06-30	563 000	152 100	1 109	205,3
Ustroń-Równica, wieś	72-08-21	489 100	205 400	640	204,5

towo sumy dobowe mogą przekroczyć nawet 400 mm<sup>36</sup>. Jakkolwiek wraz ze wzrostem powierzchni średnie obszarowe sumy opadu wyraźnie maleją, to z cytowanego wyżej opracowania wynika, że nawet dla zlewni większych niż 1000 km<sup>2</sup> średni opad dobowy może przekraczać 250 mm.

## 9. Podsumowanie rozdziału

- W niniejszej rozprawie analizowano pomiary maksymalnych miesięcznych sum dobowych opadów (MSDO) oraz daty wstąpienia rocznych MSDO z wielolecia 1956-1980 na obszarze obejmującym terytorium Polski i strefę przygraniczną na terenie Niemiec szerokości około 30 km. Dysponowano bazą danych, na którą składało się ponad 809 000 indywidualnych wartości. Dane pochodzące z Niemiec stanowiły 1-2% całości.

<sup>35</sup> Współrzędne prostokątne w układzie 1992/19 zaokrąglone do pełnych setek metrów.

<sup>36</sup> Maksymalny opad dobowy zarejestrowany na Ziemi wystąpił z 7 na 8 stycznia 1966 roku w Foc Foc na wyspie Reunion i wynosił 1825 mm (WMO 1994 za Hense, Friederichs 2006).

- W Rocznikach Opadowych PIHM/IMGW, które stanowiły główne źródło danych, stwierdzono występowanie licznych błędów. Nie wszystkie podczas archiwizacji i analizy danych udało się wyeliminować.

- Istotnym źródłem niepewności uzyskanych wyników analizy struktury przestrzennej MSDO była niska dokładność lokalizacji stanowisk pomiarowych, wynosząca 1 minutę długości i szerokości geograficznej. Związany z tym błąd mógł stanowić nawet jedną czwartą średniej odległości między najbliższymi stanowiskami.

- W analizowanym zbiorze danych występowały też przypadki stanowisk o takich samych współrzędnych lokalizacji. Stanowiły one od około 0,1 do ponad 1 procenta całości uwzględnianych punktów. Ze względu na ich duże znaczenie dla oceny zmienności MSDO na krótkich dystansach zastosowano specjalną procedurę losowego przypisywania im różnego położenia w obrębie niepewności współrzędnych.

- Na podstawie zmian liczebności sieci pomiarów pluwiometrycznych można w analizowanym wieloleciu wyróżnić cztery główne okresy: wzrostu (1956-1963), stabilizacji (1964-1972), ponownego wzrostu (1973-1975) i stopniowego spadku (1976-1980). Średnio w każdym roku wielolecia dysponowano danymi z około 2485 punktów.

- Oprócz zmian liczebności w latach 1956-1980 dokonywała się zmiana organizacji sieci. Przez cały ten czas trwała generalna tendencja do bardziej równomiernego rozmieszczenia punktów pomiarowych. Jednocześnie stopniowo likwidowano lokalne skupienia (głównie specjalne sieci miejskie) i lokowano dodatkowe punkty pomiarowe w miejscach, gdzie ich zagęszczenie było najmniejsze.

- Średnia odległość do najbliższego stanowiska wynosiła w analizowanym wieloleciu około 7,3 km, a jej odchylenie standardowe - 3,3 km. Rozkład przestrzenny stanowisk pluwiometrycznych przez cały czas miał charakter zbliżony do losowego.

- Zagęszczenie sieci pomiarowej opadów atmosferycznych było najwyższe na obszarach wyżynnych i górskich południowej Polski, szczególnie w Sudetach, na Górnym Śląsku i w zachodnich Karpatach.

- Biorąc pod uwagę częstość dni z opadem, miesięczne MSDO stanowią około 5-8% całego zbioru sum dobowych, zaś roczne MSDO - jedynie około 0,5% zmierzonych opadów dobowych. Selekcja danych miesięcznych eliminuje około 20-50% najniższych sum opadów, natomiast rocznych - ponad 95%.

- Zbiory miesięcznych i rocznych MSDO w dwóch trzecich przypadków mają dystrybuantę empiryczną zbliżoną do logarytmiczno-normalnej. Poza tym, najczęściej dopasowywano jeszcze rozkłady wartości ekstremalnych (Gumbela I) i gamma.



- Biorąc pod uwagę zmienność rozkładu MSDO w poszczególnych miesiącach wielolecia, wyróżniono trzy grupy miesięcy: o stabilnym (III-VI), pośrednim (VII-IX, XI-I) i bardzo zmiennym reżimie powtarzalności wysokich opadów (II i X). Anomalie rozkładu MSDO są związane z suszami atmosferycznymi. Najwyższą częstością bardzo niskich MSDO charakteryzuje się luty: dwie trzecie zanotowanych wartości było mniejsze od 10 mm. W lipcu takich opadów było tylko około 9%.

- Średnie miesięczne MSDO wahają się od 8,9 mm w lutym do 27,5 mm w lipcu.

- Najniższe roczne MSDO w wieloleciu 1956-1980 zawierały się w przedziale od 10,0 mm (rok 1964) do 16,4 mm (rok 1980). Jest to zmienność stosunkowo niewielka wobec zróżnicowania najwyższych rocznych sum dobowych, które wynosiły od 108,1 mm w 1956 roku do 300,0 mm w roku 1973. Średnie rocznych MSDO wahały się od 34,6 mm w roku 1976 do 51,5 mm w roku 1970.

- Dużą uwagę poświęcono analizie zmienności lokalnej MSDO, wyrażonej poprzez różnice wartości zmierzonych w sąsiadujących stanowiskach. W ujęciu sezonowym zmienia się ona od średnio od 2,5 mm w styczniu do 8,2 mm w lipcu. Najwyższe i jednocześnie najbardziej zróżnicowane maksima zmienności lokalnej występują w maju. Między majem a sierpniem zdarzają się w Polsce, średnio więcej niż w raz w roku, sytuacje, kiedy na dystansie kilku/kilkunastu kilometrów różnice maksymalnych sum dobowych przekraczają 100 mm. W pozostałych miesiącach są to przypadki sporadyczne. Średnie różnice rocznych MSDO wahały się od 8,0 mm w roku 1962 do 12,8 mm w roku 1965 i wynosiły przeciętnie 10,1 mm.

- Zarówno w globalnych, jak i lokalnych statystykach miesięcznych MSDO stwierdzono silną wprost proporcjonalną zależność między średnią a odchyleniem standardowym. W pierwszym przypadku odchylenie wynosi średnio 57, a w drugim 113% średniej. W rocznych zbiorach MSDO te relacje również istnieją, ale są zdecydowanie słabsze.

- W analizowanym wieloleciu stwierdzono wysoce istotne tendencje malejące niektórych statystyk globalnych i lokalnych MSDO (odchylenia standardowego, maksimumów i in.). Zmiany dotyczą tylko miesięcy zimowych, a wśród nich zwłaszcza marca.

- Maksymalne roczne sumy dobowe opadów mogą być w Polsce zarejestrowane każdego dnia w roku. Prawdopodobieństwo ich wystąpienia przed drugą dekadą kwietnia i po połowie października jest jednak bardzo niskie. W 90% przypadków sumy dobowe przekraczających 100 mm notowane były od połowy maja do połowy sierpnia. Losowy charakter wysokich sum opadów dobowych widoczny jest wyraźnie nawet w 25-letniej serii obserwacyjnej dla przekraczającej 2400 stanowisk sieci obserwacyjnej.

- Średnio każdego roku w analizowanym wieloleciu notowano pięćdziesiąt kilka przypadków sum dobowych przekraczających 100 mm, 7-8 wystąpień opadów powyżej 150 mm i jednokrotnie sumy wyższe od 200 mm. Największy odsetek takich zdarzeń miał miejsce w lipcu. Wysokości miesięcznych MSDO o powtarzalności 0,1 (raz na 10 lat) szacować można dla kolejnych miesięcy roku kalendarzowego na 106, 87, 73, 121, 195, 263, 262, 250, 139, 108, 81 i 81 mm. Roczna MSDO o takiej powtarzalności wynosi około 267 mm. Podane wartości trzeba, ze względu na specyfikę analizowanej bazy danych, traktować jako minimalne.

- Większość z najwyższych, przekraczających 200 mm na dobę, opadów zanotowano w trakcie dwóch epizodów w Karpatach. Maksymalny, w okresie pomiarów instrumentalnych, opad dobowy wynoszący 300 mm zarejestrowano 30 czerwca 1973 roku na Hali Gąsienicowej, w Tatrach. Przypadki ekstremalnych opadów stwierdzonych w bezpośrednim sąsiedztwie granic Polski, a także oceny teoretyczne sugerują, że maksymalne ze względów geograficzno-klimatycznych sumy dobowe mogą sięgać w górach do 400 mm.

## **Analiza struktury przestrzennej MSDO opartej na semiwariogramach danych znormalizowanych**

### **1. Wprowadzenie**

Wstępnym etapem symulacji *p-field* jest wygenerowanie inicjalnego, bezwarunkowego, pola prawdopodobieństwa analizowanej cechy zachowującego najważniejsze właściwości jej struktury przestrzennej (patrz dodatek XI.2). W przyjętej w niniejszej pracy metodyce (ryc. 129) etap ten był wykonywany dla każdego z 325 analizowanych zbiorów danych poprzez (1) normalizację danych, (2) obliczanie empirycznego izotropowego semiwariogramu danych o zasięgu 212,5 km (85 odstępów po 2,5 km), (3) modelowanie matematyczne uzyskanego semiwariogramu, (4) generowanie na podstawie uzyskanego modelu metodą symulacji spektralnej inicjalnego pola prawdopodobieństwa. Z przedstawionego w skrócie opisu wynika, że jednym z efektów owej procedury było 325 modeli struktury przestrzennej, po jednym dla każdego analizowanego zbioru danych miesięcznych i rocznych. Ich zaletą jest duży, ponadregionalny zasięg, - w skali setek kilometrów. Pominięcie w analizie potencjalnej anizotropii miało dwa aspekty: praktyczny (czasochłonność) i merytoryczny (częściowe uniknięcie maskowania heterogeniczności pola MSDO na tak dużym obszarze, jaki zajmuje Polska). Zagadnienie to zostało omówione szczegółowo w dodatku X.3. Normalizacja danych ma przy analizie struktury przestrzennej wielką zaletę, umożliwia bowiem bezpośrednio porównywanie wartości semiwariancji (patrz dodatek X.4). Nie są one uzależnione już od zmienności bezwzględnych wartości analizowanych zbiorów danych.

## 2. Częstość występowania i charakterystyka elementarnych modeli

Wszystkie 325 modeli danych znormalizowanych miało charakter złożony, czy jak to się określa w literaturze geostatystycznej – zagnieżdżony (ang. *nested*). Oznacza to, że składały się z co najmniej dwóch modeli elementarnych. Zagadnienie to zostało wyjaśnione w podrozdziale III.2.5.2 (s. 43), łącznie z przedstawieniem podstawowych właściwości używanych w niniejszym opracowaniu typów struktur elementarnych.

We wszystkich modelach znormalizowanych danych MSDO, zarówno miesięcznych jak i rocznych, konieczne było użycie modelu nuggetowego. Fakt ten jest oczywisty, biorąc pod uwagę nieciągłość opadów dobowych oraz asynchroniczność danych MSDO. Dodatkowym składnikiem tej części zmienności wyników pomiarów, która nie wykazuje żadnej struktury przestrzennej, są systematyczne i losowe błędy określenia objętości opadu, oraz ich różnicowanie w małej skali (na dystansie mniejszym niż przeciętny odstęp pomiędzy stanowiskami).

Drugą, najczęściej stosowaną w modelowaniu strukturą elementarną, zarówno przy miesięcznych, jak i rocznych zbiorach danych MSDO (tab. 10), jest struktura sferyczna. Oznacza to liniowy spadek podobieństwa wraz z odległością i zmienność przestrzenną o postaci przeplatających się płatów wysokich i niskich wartości opadów. Rozmiar płatów (ich średnica) jest zbliżony i można go ocenić na podstawie zasięgu struktury (parametr  $a$  modelu sferycznego, patrz s. 41-42). Proporcje przypadków użycia funkcji sferycznej są podobne dla poszczególnych składowych przy miesięcznych i rocznych zbiorach danych MSDO. Stanowi on bowiem 95-100% modeli zastosowanych dla składowej 2 i 3, 70-82% dla składowej 4 (trendu) i prawie idealnie tak samo, bo około 64% dla składowej 1 (tab. 10). Wykorzystanie funkcji sferycznej do modelowania składowej 4, długodystansowej, określanej dalej jako trend, oznacza jedynie, że spadek podobieństwa wraz z odległością miał w jej obrębie charakter liniowy. Aby uzyskać dobre dopasowanie manipulowano arbitralnie przyjętą wartością zasięgu w granicach 265-3800 km. W 84% przypadków były to jednakże odległości 350 i 380 km. Nie mają te wartości jednak realnego znaczenia w odniesieniu do zasięgu funkcjonowania procesów – umożliwiały jedynie regulację nachylenia krzywej. Trzeba również pamiętać, że wśród możliwych do wykorzystania w niniejszej pracy elementarnych funkcji, także wykładnicza ma w początkowym odcinku charakter liniowy. Mogłaby zatem być zamiennie stosowana do modelowania trendu. To, że używano modelu sferycznego, wynikało jedynie z przyjętej *a priori* konwencji.

Trzecią, biorąc pod uwagę częstość używania do modelowania semiwariogramów danych znormalizowanych, elementarną funkcją była wykładnicza (tab. 10). Wykorzystywano ją przede wszystkim przy składowej pierwszej,

**Tabela 10.** Typy funkcji (modeli) użytych do dopasowania poszczególnych składowych złożonych modeli struktury przestrzennej miesięcznych znormalizowanych danych MSDO

Typ modelu Model type	Składowa 1 Component 1		Składowa 2 Component 2		Składowa 3 Component 3		Trend	
	Liczebność Number	(%)	Liczebność Number	(%)	Liczebność Number	(%)	Liczebność Number	(%)
<b>Modele miesięczne - Monthly models</b>								
Sferyczny Spherical	181	63,5	172	94,5	146	94,8	134	70,5
Wykładniczy Exponential	104	36,5	10	5,5	0	0,0	1	0,5
Gaussowski Gaussian	0	0,0	0	0,0	8	5,2	55	28,9
<b>Suma - Total</b>	<b>285</b>		<b>182</b>		<b>154</b>		<b>190</b>	
<b>Modele roczne - Yearly models</b>								
Sferyczny Spherical	16	64,0	25	100,0	9	100,0	9	82,0
Wykładniczy Exponential	9	36,0	0	0	0	0,0	0	0,0
Gaussowski Gaussian	0	0,0	0	0,0	0	0,0	2	18,0
<b>Suma - Total</b>	<b>25</b>		<b>25</b>		<b>9</b>		<b>11</b>	

gdzie występuje zarówno w modelach struktury przestrzennej miesięcznych, jak i rocznych zbiorów danych MSDO, w około 36% przypadków. Oznacza to, że dominowała w tych okresach na krótkich dystansach mozaika nieregularnych powierzchni wysokich i niskich opadów, a zasięg (średnica) owych płatów zmieniał się w szerokim zakresie w sposób losowy. W modelach struktury przestrzennej rocznych MSDO funkcji wykładniczej, poza pierwszą składową, nie stosowano do żadnej innej; w miesięcznych jej rola była marginesowa: 10 przypadków (5,5%) modeli składowej 2 i 1 przypadek (0,5%) trendu (składowa 4). Oznaczałoby to, że w takich skalach przestrzennych i czasowych MSDO mają raczej zasięgi stałe (powtarzalne), a nie losowe. Wniosek ten należy jednak traktować bardziej jako hipotezę niż stwierdzenie faktu.

Najrzadziej przy modelowaniu struktury przestrzennej znormalizowanych danych MSDO używano czwartej dopuszczalnej w programie IKSIM (Ying 2000) funkcji elementarnej – gaussowskiej (tab. 10). Przy miesięcznych zbiorach danych stanowiła ona komponent w około 60 modelach złożonych. Osem razy wykorzystano funkcję gaussowską do dopasowania składowej trzeciej (5,2%), a pięćdziesiąt pięć – składowej czwartej, czyli trendu (28,9%). W dwudziestu pięciu modelach rocznych MSDO zastosowano ją jedynie

dwukrotnie i to tylko do dopasowania trendu. Jej użycie ma bardziej charakter „techniczny” niż związany ściśle ze specyfiką procesu przestrzennego, jaki reprezentuje (patrz podrozdz. III.2.5.2, s. 42 i 43). Był to bowiem jedyny z elementarnych modeli, jakimi można było operować, mający częściowo charakter paraboliczny – „wklęsły”. Zazwyczaj w takiej sytuacji wykorzystuje się funkcję potęgową o postaci:

$$g(h) = h^\omega \quad \text{gdzie } 0 < \omega < 2 \quad [25]$$

Przy  $\omega > 1$  ma on kształt wklęsły – paraboliczny. Funkcja potęgowa różnie w nieskończoność, nie ma zatem zarówno zasięgu, ani progów. Tego typu modele nazywa się w geostatystyce nieograniczonymi (ang. *unbounded*). Interpretuje się jako reprezentację procesu losowego mającego charakter ruchów Browna (Webster, Oliver 2001). Przyjmując dla uproszczenia ruch cząsteczki w jednym wymiarze, jej prędkość lub moment pędu w położeniu  $\mathbf{x} + \mathbf{h}$  zależy od jej prędkości lub momentu pędu w bezpośrednio poprzednim bliskim położeniu  $\mathbf{x}$ . Relację tę można zapisać równaniem:

$$Z(\mathbf{x} + \mathbf{h}) = \beta Z(\mathbf{x}) + \varepsilon \quad [26]$$

gdzie  $\varepsilon$  to niezależne, losowe odchylenie gaussowskie, a  $\beta$  to parametr. W najprostszym przypadku  $\beta = 1$ , a wariogram procesu wygląda wówczas następująco:

$$2\gamma(\mathbf{h}) = E\left[\{Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})\}^2\right] = |\mathbf{h}|^k \quad [27]$$

Jeśli wykładnik  $k$  w równaniu [27] wynosi 1, to wówczas mamy do czynienia z modelem liniowym, gdzie  $\gamma(|\mathbf{h}|) \rightarrow \infty$  gdy  $|\mathbf{h}| \rightarrow \infty$ . Nazywany jest on często modelem ścieżki losowej (ang. *random walk model*).

Przy typowych ruchach Browna składnik  $\varepsilon$  jest dla każdego kroku niezależny. Jeśli jednak  $\varepsilon$  z równania [26] są przestrzennie skorelowane, to ślad ruchu cząsteczki jest bardziej „gładki” niż w klasycznym ruchu Browna. Wówczas to wykładnik  $k$  jest większy od 1, a krzywa wklęsła. Z drugiej strony, kiedy odchylenia  $\varepsilon$  są odwrotnie (ujemnie) skorelowane, wówczas ścieżka ruchu jest bardziej chaotyczna. Wykładnik  $k$  z równania [27] jest wtedy mniejszy od 1, a krzywa wariogramu wypukła. Kiedy  $\varepsilon$  są idealnie skorelowane, wówczas  $k = 2$ , a ślad ruchu cząsteczki jest płynny (gładki), co oznacza, że nie ma już charakteru losowego. W sytuacji gdy  $k \rightarrow 0$  chaotyczność ruchu rośnie aż do poziomu białego szumu, którego obrazem jest opisany w podrozdziale III.2.5.2 model nuggetowy.

Do sprawdzenia, czy częstość użycia poszczególnych typów modeli elementarnych nie wykazuje zmienności sezonowej wykorzystano test Chi-kwadrat. Istotne statystycznie zróżnicowanie stwierdzono jedynie w odniesieniu do składowej drugiej ( $p = 0,00084$ ). Decydujące znaczenie ma w tym

wypadku wysoka reprezentacja modelu wykładniczego w lipcu i sierpniu. W pozostałych miesiącach częstości obserwowane i oczekiwane poszczególnych modeli elementarnych nie różnią się istotnie. Podobny układ, obejmujący jednak okres od maja do sierpnia, występował przy składowej pierwszej. Odchylenia nie były jednak na tyle duże, żeby wynik testu potwierdził ich istotność ( $p = 0,589$ ). Biorąc pod uwagę wcześniejsze omówienia specyfiki procesów losowych reprezentowanych przez poszczególne modele, można przypuszczać, że latem częściej pojawiają się rozkłady opadów o szerszym, bardziej zróżnicowanym, spektrum zasięgów.

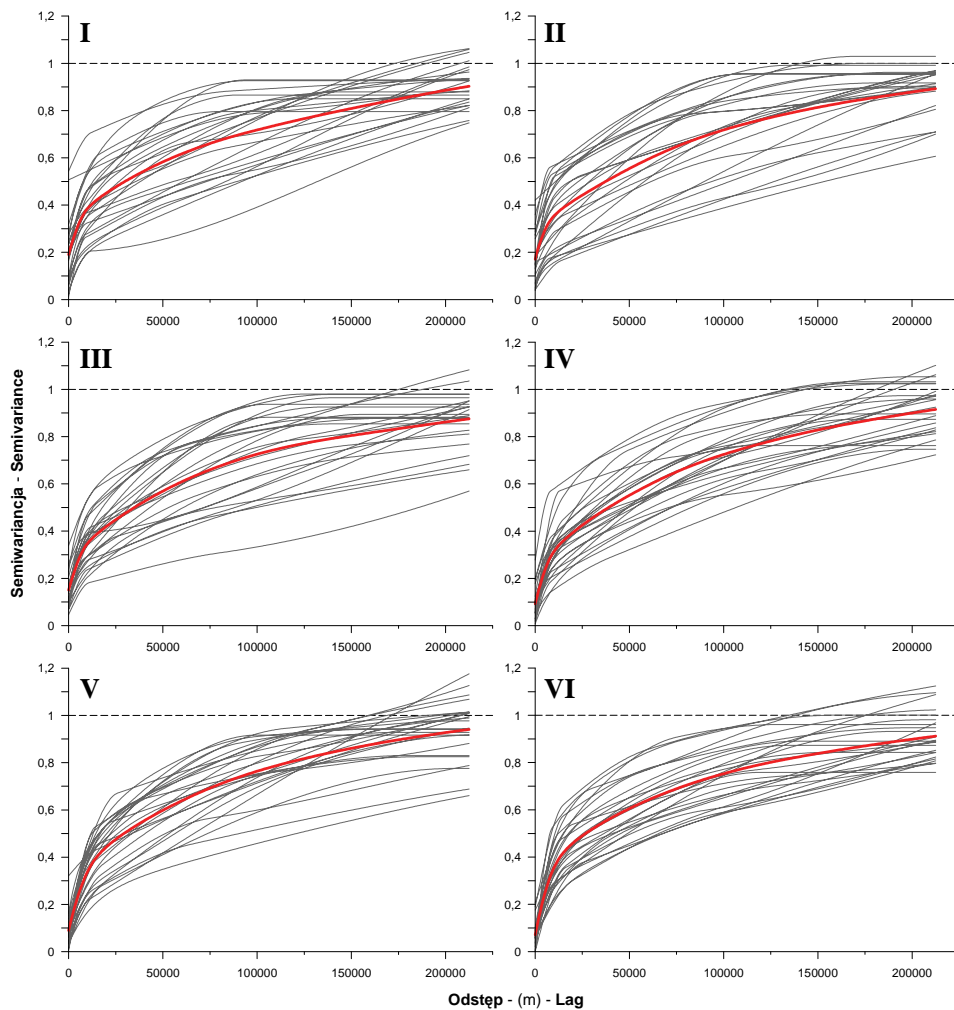
### **3. Sezonowa i wieloletnia zmienność struktury przestrzennej MSDO**

Cały zbiór analizowanych 325 modeli semiwariogramów danych znormalizowanych przedstawiono na rycinach 50-52. Zostały one pogrupowane w kolejnych miesiącach roku (i latach) w celu wychwycenia ewentualnej zmienności sezonowej. Z rycin poniższych wynika jednak, że w zasadzie w każdym miesiącu może się pojawić dowolny z zarejestrowanych przebiegów autokorelacji pola MSDO. Poszczególne miesiące różnią się jednak nieco rozrzutem (szerokością wiązki) krzywych, czyli zmiennością obserwowanych układów struktur, a także występowaniem pojedynczych, anomalnych układów. I tak, kwiecień, czerwiec, październik i listopad charakteryzują się większą zawartością wiązki krzywych, podczas gdy styczeń, luty, lipiec i grudzień – jej większym rozproszeniem. Nie są to jednak różnice duże. Bardziej zwraca uwagę występowanie (lub brak) pojedynczych, odstających przypadków, na przykład: styczeń 1979, styczeń 1960, marzec 1964, lipiec 1965, lipiec 1967, lipiec 1968, lipiec 1970, wrzesień 1956, listopad 1963, listopad 1964, listopad 1965, czy grudzień 1967. Semiwariogram to miara niepodobieństwa, zatem krzywe układające się poniżej głównej wiązki świadczą o występowaniu mało zmiennych opadów na dużym obszarze. I odwrotnie, w miesiącach, dla których modele semiwariogramów ulokowane są powyżej głównej wiązki opady były bardzo zmienne – obejmowały niewielkie powierzchnie. Pierwsze z wymienionych powyżej charakteryzują się zazwyczaj skrajnie niskimi wartościami wariancji nuggetowej (zmienności losowej), a drugie – wysokimi. Taka prawidłowość jest jak najbardziej oczekiwana.

Najistotniejsza jednak zmienność struktur przestrzennych MSDO w ujęciu sezonowym dotyczy wariancji nuggetowej. Z jednej strony mamy bowiem miesiące zimowe (zwłaszcza styczeń i luty) o wysokiej i bardzo zmiennej wartości tej cechy, a z drugiej miesiące wiosenno-letnio-jesienne o niskim i mało zmiennym nuggetcie (zwłaszcza czerwiec, sierpień i wrzesień).

Jest to prawdopodobnie z jednej strony odbicie większych błędów pomiarów opadów śniegu zimą, a także zróżnicowania sezonowego prędkości wiatrów (błędy związane z turbulencją). Problem ten będzie jeszcze omawiany dalej.

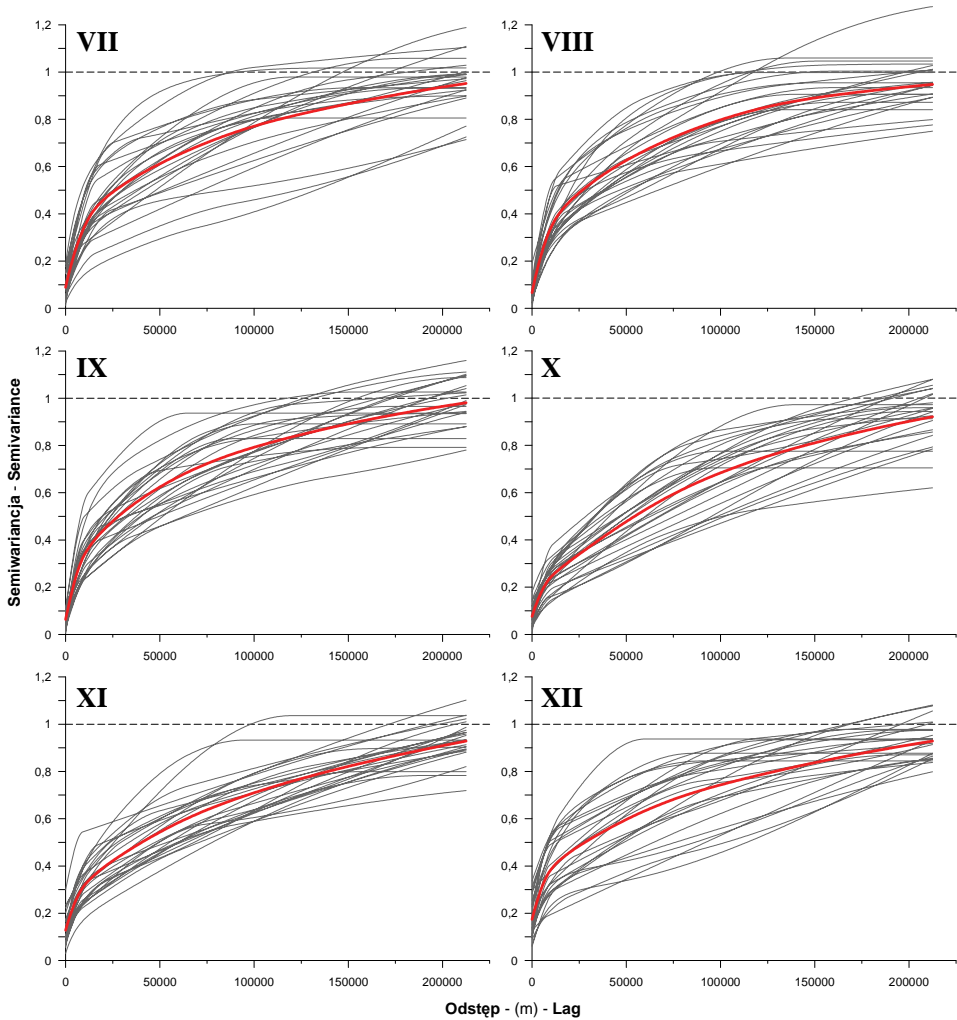
Rycina 52, pokazująca przebieg średnich semiwariogramów dla poszczególnych miesięcy, na pozór wskazuje na bardzo niewielkie zróżnicowanie. Jedynie krzywa dla października wyraźnie odstaje od reszty ukazując częstsze występowanie „dużych” deszczy o małej zmienności przestrzennej



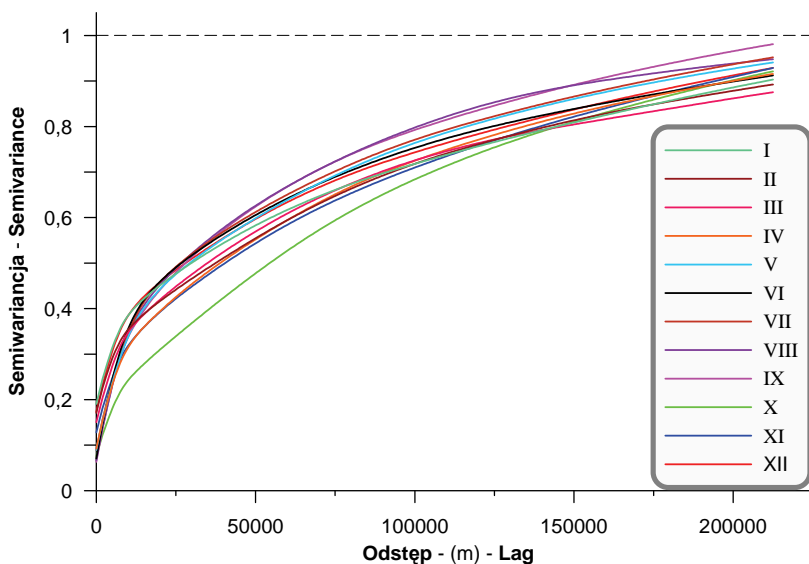
Ryc. 50. Bez kierunkowe (izotropowe) modele semiwariogramów danych znormalizowanych dla miesięcznych zbiorów MSDO od stycznia do czerwca. Grubą czerwoną linią zaznaczono średni semiwariogram dla 25-lecia 1956-1980




i sporym zasięgu. Jednakże, bardziej drobiazgowo analiza pokazuje, że krzywe w obrębie wiązki układają się w uporządkowany, konsekwentny sposób. W dolnej części zgrupowane są semiwariogramy miesięcy jesienno-zimowych i wczesnowiosennych, w górnej – późnowiosennych i letnich. Szczególnie silne jest zróżnicowanie pomiędzy wrześniem, leżącym w skrajnie górnej części wiązki krzywych (obok lipca i sierpnia), a październikiem – usytuowanym na samym dole. Charakterystyczne jest też krzyżowanie się



**Ryc. 51.** Bezkierunkowe (izotropowe) modele semiwariogramów danych znormalizowanych dla miesięcznych zbiorów MSDO od lipca do grudnia. Grubą czerwoną linią zaznaczono średni semiwariogram dla 25-lecia 1956-1980



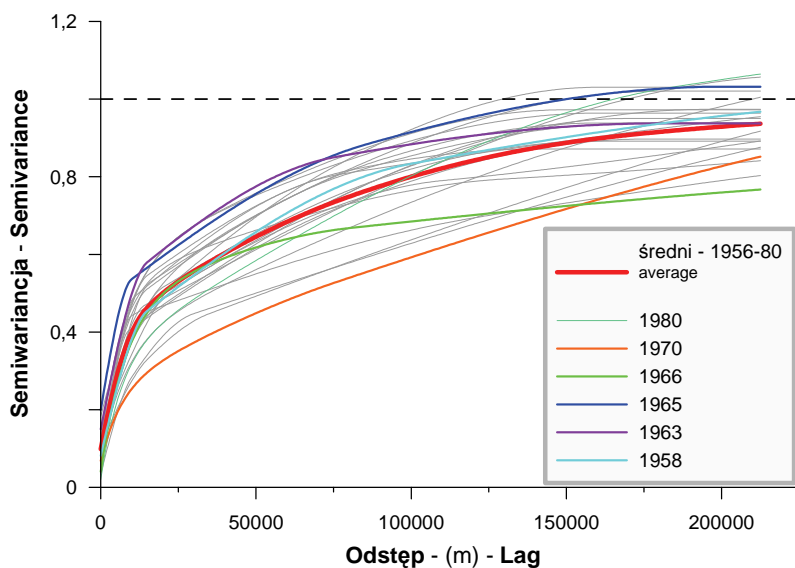
Ryc. 52. Średnie 25-letnie z izotropowych modeli semiwariogramów danych znormalizowanych dla poszczególnych zbiorów miesięcznych MSDO (patrz ryc. 50 i 51) 

przebiegów linii autokorelacji obu wymienionych wyżej grup przy odległościach rzędu kilku do kilkunastu kilometrów. Ten układ związany jest z, wspomnianą już wcześniej, większą względną zmiennością losową MSDO (nuggetem) w miesiącach jesienno-zimowych i wczesnowiosennych.

Skala rozrzutu i kształty semiwariogramów rocznych MSDO (ryc. 53) są bardzo podobne do omówionych wyżej struktur przestrzennych miesięcznych MSDO. Charakterystyczną różnicę stanowi jedynie wyraźnie wyższy udział w całości zmienności przestrzennej wysokich opadów struktury pierwszej, o zasięgu kilkunastu kilometrów. Wskazywałoby to na większą, w ujęciu rocznym, rolę opadów konwekcyjnych, o niewielkim zasięgu przestrzennym. Skrajnymi charakterystykami struktury przestrzennej MSDO odznaczały się rok 1970 i lata 1963 oraz 1965. Rok 1970 to duże obszary objęte opadami o względnie zbliżonej wydajności; pozostałe dwa lata to „mozaika” małych powierzchni o silnie zróżnicowanych wartościach MSDO. Specyficzny był także rok 1966, kiedy na rozległą, mało zróżnicowaną „powierzchnię” opadową „nałożone” zostały „izolowane wzniesienia”.

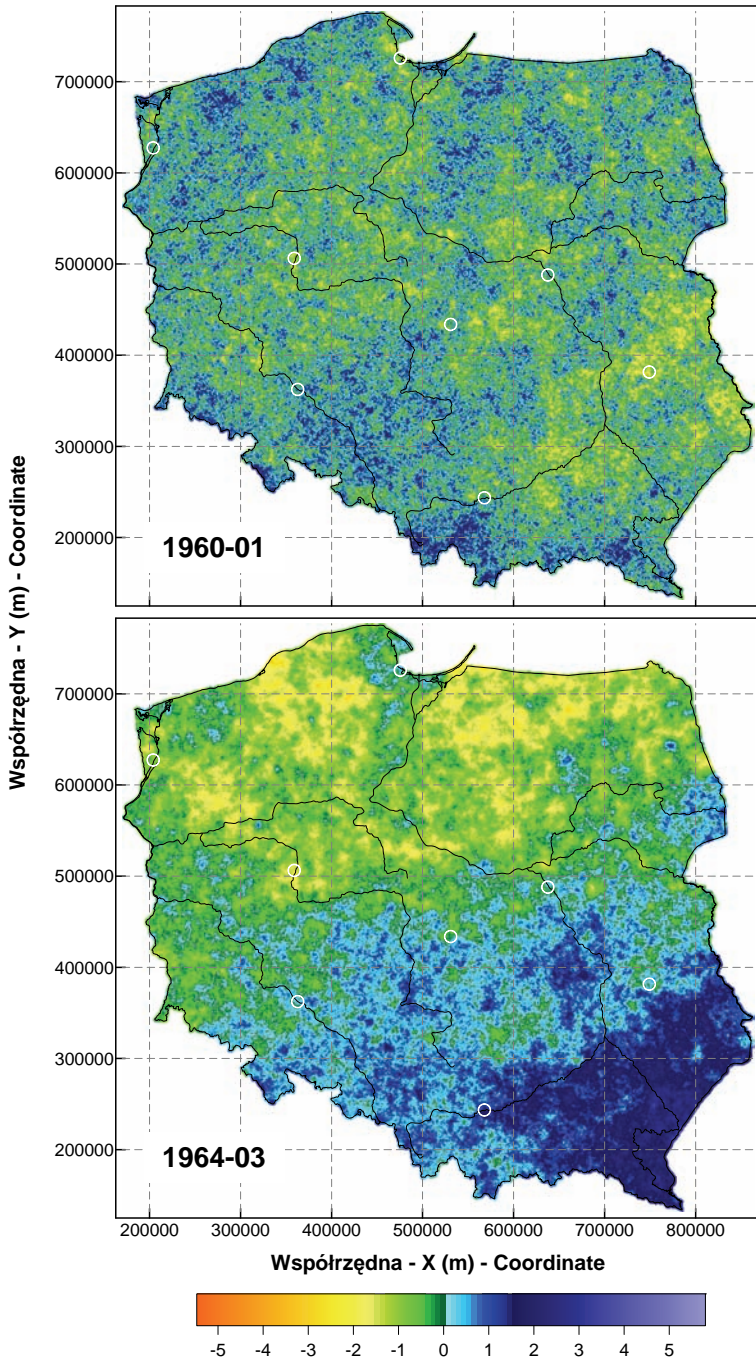
Czytelnikowi, dla którego zastosowana w niniejszej rozprawie metodyka stanowi *novum*, wyobrażenie sobie zmienności przestrzennej reprezentowanej przez poszczególne modele będzie z pewnością sprawiało dużą trudność. Dlatego, charakterystyczne (przeciętne i ekstremalne), omawiane w dalszej części niniejszego rozdziału, typy struktury przestrzennej MSDO

przedstawiono również w postaci map wysokości opadów<sup>37</sup>. Zdecydowano, że będą to obrazy pojedynczych symulacji warunkowych wykonywanych w przestrzeni danych znormalizowanych (patrz dodatek X.2 i X.4). W dodatku X scharakteryzowano szeroko wady i zalety map uzyskanych za pomocą geostatystycznych estymacji i symulacji. Największym plusem tych drugich, w kontekście tematu rozprawy, jest dobre odwzorowanie modelu struktury przestrzennej oraz honorowanie danych pomiarowych w ich lokalizacjach. MSDO w poszczególnych miesiącach i latach analizowanego wielolecia różniły się bardzo swą bezwzględną wielkością. Aby to zróżnicowanie nie maskowało różnic i podobieństw rozkładów przestrzennych, symulacje wykonywano w przestrzeni danych znormalizowanych. Dzięki takiemu zabiegowi możliwe jest bezpośrednie porównywanie map, a także wykonywanie na nich operacji algebraicznych. W trakcie porównywania map należy zwrócić uwagę na wielkość i kształt płatów wysokich i niskich wartości, a także na ich amplitudy wyrażone kolorami depresji i elewacji symulowanego pola MSDO. Pierwszy przykład takich map stanowi rycina 54, na której umieszczono symulowane obrazy zmienności przestrzennej miesięcznych MSDO ze stycznia roku 1960 i marca roku 1964. Są to ekstremalne przypadki wyraźnie odróżniające się na rycinie 50.

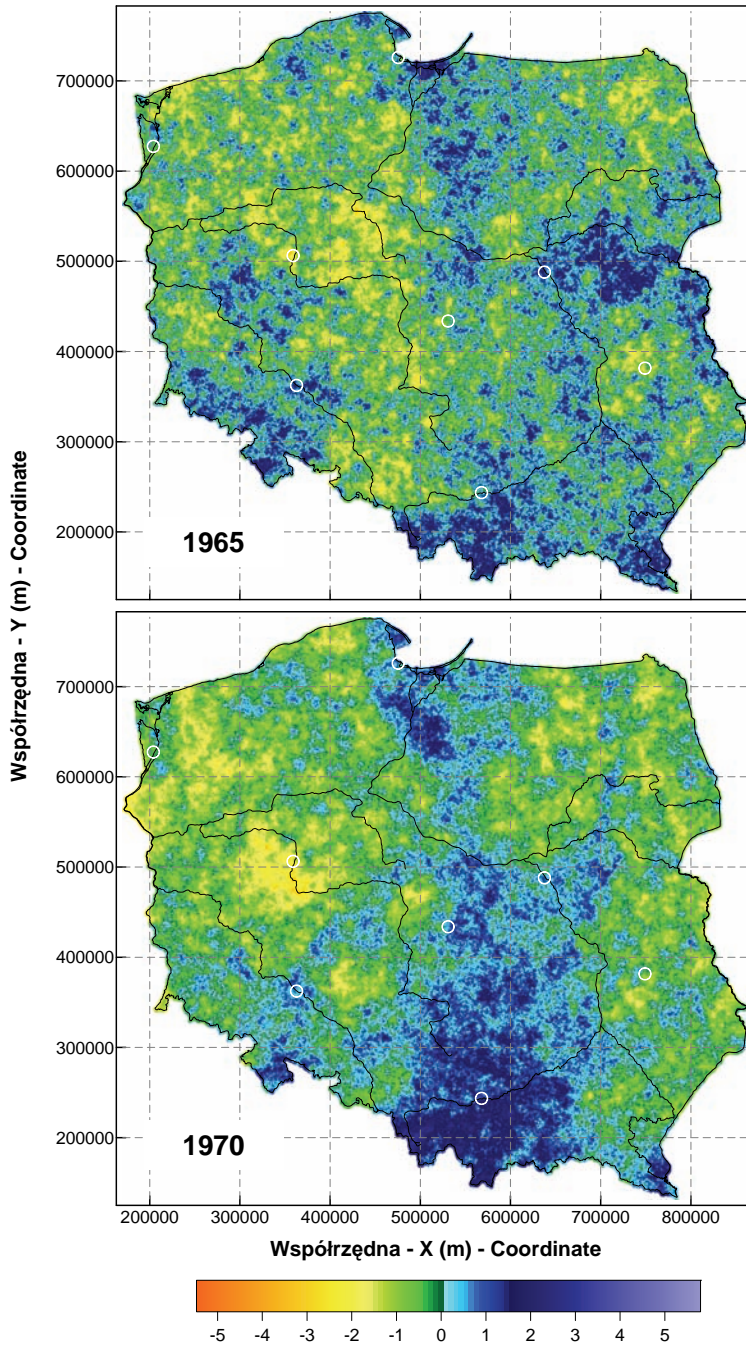


**Ryc. 53.**Bez kierunkowe (izotropowe) modele semiwariogramów danych znormalizowanych dla rocznych zbiorów MSDO. Grubą czerwoną linią zaznaczono średni semiwariogram dla 25-lecia 1956-1980, a kolorem – kilka przypadków ekstremalnych

<sup>37</sup> Pojedyncze siatki  $1 \times 1$  km wartości symulowanych dla wszystkich 325 zbiorów analizowanych MSDO zamieszczono na dołączonym dysku DVD (patrz załącznik XII.2).



Ryc. 54. Pojedyncze obrazy symulacji warunkowych pola MSDO wykonanych dla danych znormalizowanych ze stycznia roku 1960 i marca 1964



Ryc. 55. Pojedyncze obrazy symulacji warunkowych pola rocznych MSDO wykonanych dla danych znormalizowanych z roku 1965 i 1970



Wyróżniającą cechą struktury MSDO ze stycznia 1960 był najwyższy w całym analizowanym wieloleciu udział semiwariancji nuggetowej wynoszący 0,54. Maksymalna była zatem składowa losowa związana z błędami pomiarowymi i/lub zmiennością krótkodystansową opadów. Obraz takiego pola miesięcznych MSDO (ryc. 54) charakteryzuje się silnie zarysowaną „ziarnistością” związaną z bliskim sąsiedztwem skrajnie niskich i wysokich opadów. Nie ma wyraźnych, zwartych obszarów zbliżonych ich sum. Model dla marca roku 1964 reprezentuje z kolei krzywa o najmniejszym nachyleniu i bardzo łagodnym przebiegu. Świadczy to o dominacji procesów o małej zmienności i dużym zasięgu przestrzennym przekraczającym zakres analizy, to jest ponad 200 km. Symulowane w siatce 1×1 km MSDO dla tego miesiąca pokazują rozległy obszar niskich opadów obejmujący północ kraju i sięgający po Warszawę. W obrębie niego występują niewielkie „wyspy” wyższych sum dobowych – na przykład na Pojezierzu Kaszubskim. Południowo-wschodnia część Polski, poniżej linii łączącej Kraków i Lublin, to zwarty obszar względnie jednolitych wysokich opadów. Istotny dla zestawionych na rycinie 54 przypadków jest również fakt, iż średnie MSDO w obu miesiącach były bardzo zbliżone i wynosiły około 9,1 mm. Podobny był również zakres zmienności.

Należy jednak wyraźnie zaznaczyć, że tak istotne zróżnicowanie rozkładu przestrzennego, widoczne doskonale gołym okiem, jest wyjątkiem, a nie regułą. Gdyby było inaczej, nie byłoby potrzeby stosowania wyrafinowanych metod numerycznych. Generalnie jednak, duża liczba danych pomiarowych i ich skomplikowane relacje czasoprzestrzenne powodują, że wyróżnienie układów typowych i anomalnych jest, na podstawie wizualnej analizy map, praktycznie niemożliwe. Ilustracją tego faktu może być rycina 55, na której zamieszczono symulowane obrazy rozkładu przestrzennego rocznych MSDO dla dwóch lat o ekstremalnych charakterystykach struktury – roku 1965 i 1970 (por. ryc. 53). Można było z góry przypuszczać, że dłuższy okres agregacji danych powoduje mniejsze zróżnicowanie obrazów rocznych MSDO w stosunku do miesięcznych, lecz ekstremalne przypadki w okresie 25-letnim powinny różnić się znacząco. Na rycinie 55 wyraźnej odmienności nie widać. Bardziej w oczy rzucają się podobieństwa lokalizacji obszarów o wysokich i niskich maksymalnych sumach dobowych opadów. Dopiero szczegółowy ogląd ujawnia istotne różnice. W roku 1965 powierzchnie o podobnych wysokościach MSDO były mniejsze i bardziej nieregularne, wręcz „porozrywane”.

#### **4. Klasyfikacja struktury przestrzennej MSDO - liczba i charakter składowych**

Klasyfikacja struktury przestrzennej powinna uwzględniać dwa elementy: podobieństwo kształtu krzywej (niezależnie od aktualnej wartości semiwariancji) i podobieństwo wartości semiwariancji.

Podstawą pierwszej klasyfikacji była liczba i charakter elementarnych składowych, z których zbudowany był model. Na podstawie tego kryterium wyróżniono 6 klas, przy czym do jednej z nich należał tylko 1 z 325 modeli (ryc. 56 i 57):

1. nugget + 1 składowa + trend,
2. nugget + 2 składowe + trend,
3. nugget + 3 składowe + trend,
4. nugget + 2 składowe,
5. nugget + 3 składowe
6. nugget + 1 składowa

Anomalnym przypadkiem był luty 1957 roku. Strukturę przestrzenną MSDO w tym miesiącu najlepiej opisywał model składający się z nuggetu i jednej struktury sferycznej o zasięgu 110 km (typ 6 na ryc. 57).

Roczne MSDO wykazują struktury przestrzenne jedynie typu drugiego, czwartego lub piątego (ryc. 57B). Najczęściej występuje typ drugi: w 45% MSDO miesięcznych i 44% rocznych. Udział struktur typu 1, 4 i 5 jest dla zbioru danych miesięcznych podobny – od około 15 do 18%. Struktura MSDO rocznych typu 4 występuje w 20% przypadków, typu 5 – w 36%. Występowanie pięciu podstawowych typów struktury przestrzennej dla danych miesięcznych nie wykazuje istotnej zmienności sezonowej (ryc. 58). Typy 1, 2, 4 i 5 notowano we wszystkich miesiącach roku, a różnice liczebności były niewielkie i nie wykazywały żadnej regularności. Występowania struktur typu 3 nie zarejestrowano wprawdzie w miesiącach zimowych (XII, I i II), ale liczebność tej grupy jest tak niska (w pozostałych miesiącach było 1-2 przypadków), że nie można tego faktu traktować jako prawidłowość. Hipotezę o braku zmienności sezonowej występowania poszczególnych typów struktury przestrzennej MSDO potwierdzono za pomocą jednoczynnikowej analizy wariancji.

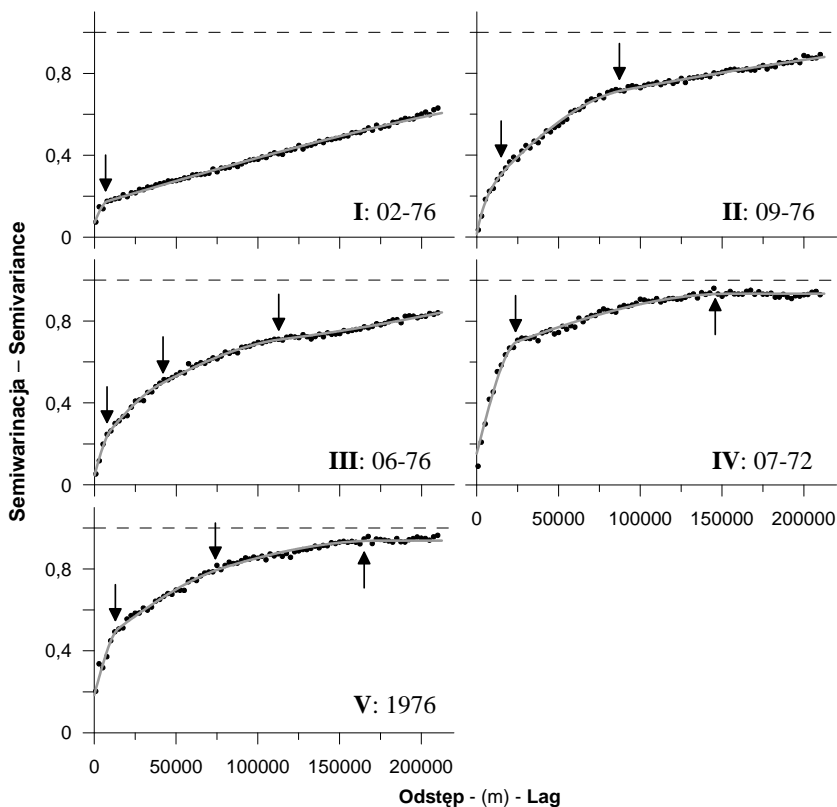
## **5. Charakterystyka składowych złożonych modeli struktury przestrzennej MSDO**

Każdy złożony model struktury przestrzennej można scharakteryzować poprzez podanie zasięgów autokorelacji poszczególnych jego składowych oraz odpowiadających im wariancji progowych (oraz oczywiście nuggetu – składowej o zasięgu zerowym). Zbiór 300 modeli miesięcznych MSDO został w ten sposób podsumowany za pomocą histogramów na rycinach 59 i 60.

Zmienność losowa (nugget) wynikająca z nieciągłości MSDO i ich asynchroniczności, błędów pomiarowych oraz niedokładności określenia lokalizacji stanowisk pomiarowych, a także zmienności krótkodystansowej opadów (na odległościach mniejszych niż 2,5 km) wynosi dla danych miesięcznych 11,3% (SD = 8,4%) całej wariancji danych. Dla rocznych MSDO średnia wa-

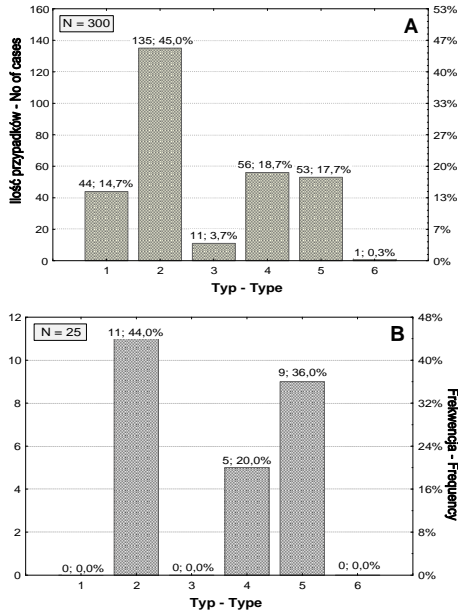
riancja nuggetowa jest nieco niższa (9,9%), ale przede wszystkim znacznie mniej zróżnicowana (SD = 4,8%). Podobną wielkość semiwariancji nuggetowej (ok. 10%) ekstremalnych opadów dobowych z sierpnia 2002 stwierdził Haberlandt (2007). Rozrzut wielkości nuggetu jest w poszczególnych miesiącach bardzo duży, bo zamyka się w przedziale 0,1-54,3% i silnie skośny – ponad 55% mieści się w klasie 0-10%.

Najbardziej powszechnie występująca (285 przypadków) jest autokorelacja MSDO opadów o zasięgu od poniżej 10 do 40 km (średnio ok. 15,5 km, SD = 6,4 km, ryc. 60 i 61). Istnienie takiej struktury przestrzennej jednego z dwóch wyróżnionych typów silnych opadów w okolicach Oslo podawał Skaugen (1997). Jej średni udział w całkowitej wariancji danych wynosi dla danych miesięcznych ponad 20% (ryc. 59), wahając się w szerokich granicach od 3,5 do 55,9%. Rozkład tych danych jest również asymetryczny, ale znacznie słabiej niż w przypadku nuggetu.

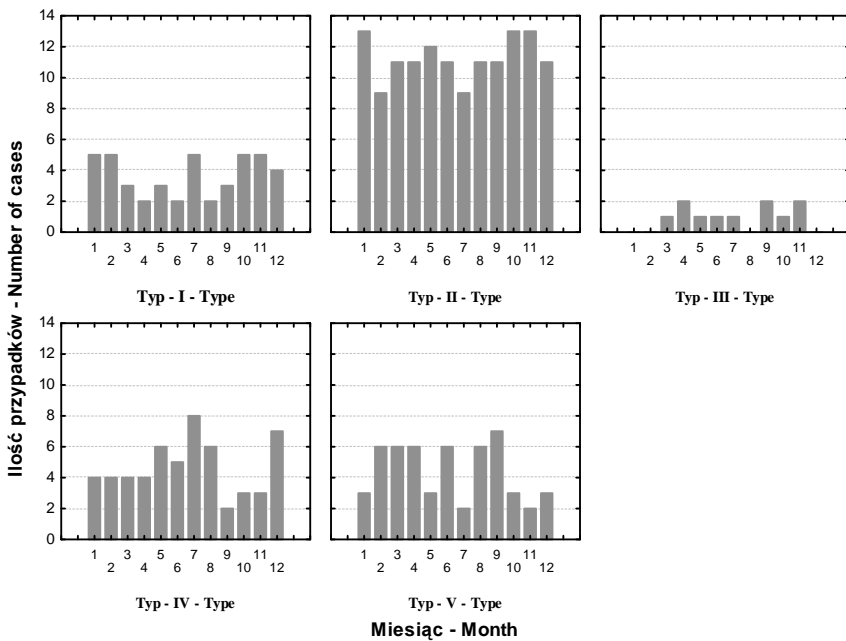


Ryc. 56. Przykładowe semiwariogramy empiryczne danych znormalizowanych oraz ich modele dla pięciu wyróżnionych typów struktury przestrzennej MSDO. Strzałkami zaznaczono orientacyjne zasięgi poszczególnych struktur

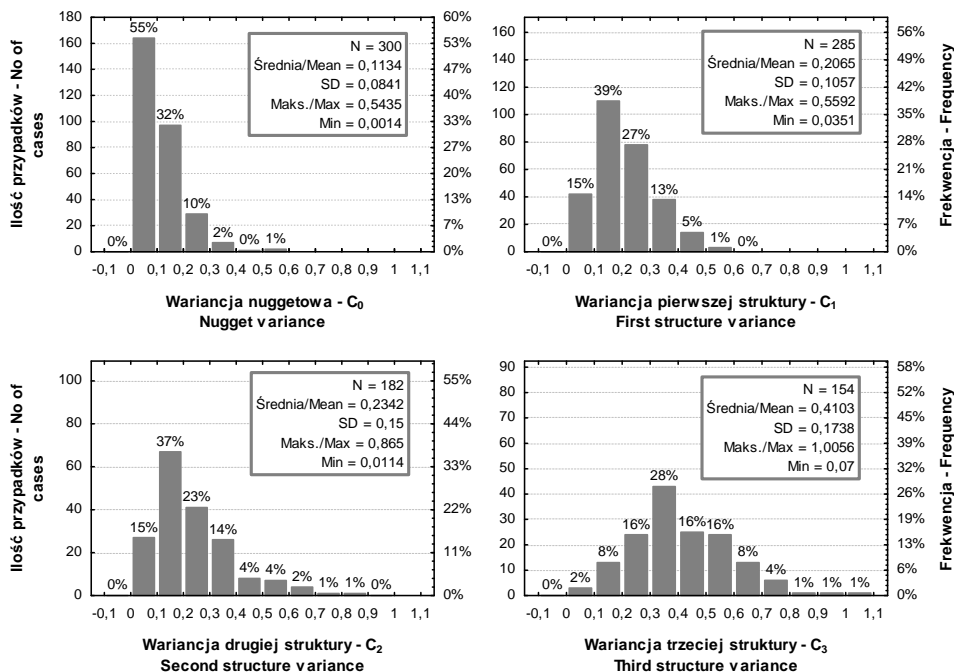




Ryc. 57. Frekwencja wyróżnionych typów struktury przestrzennej MSDO dla miesięcznych (A) i rocznych (B) zbiorów danych



Ryc. 58. Sezonowa zmienność pięciu podstawowych typów struktury przestrzennej MSDO w 25-leciu 1956-1980

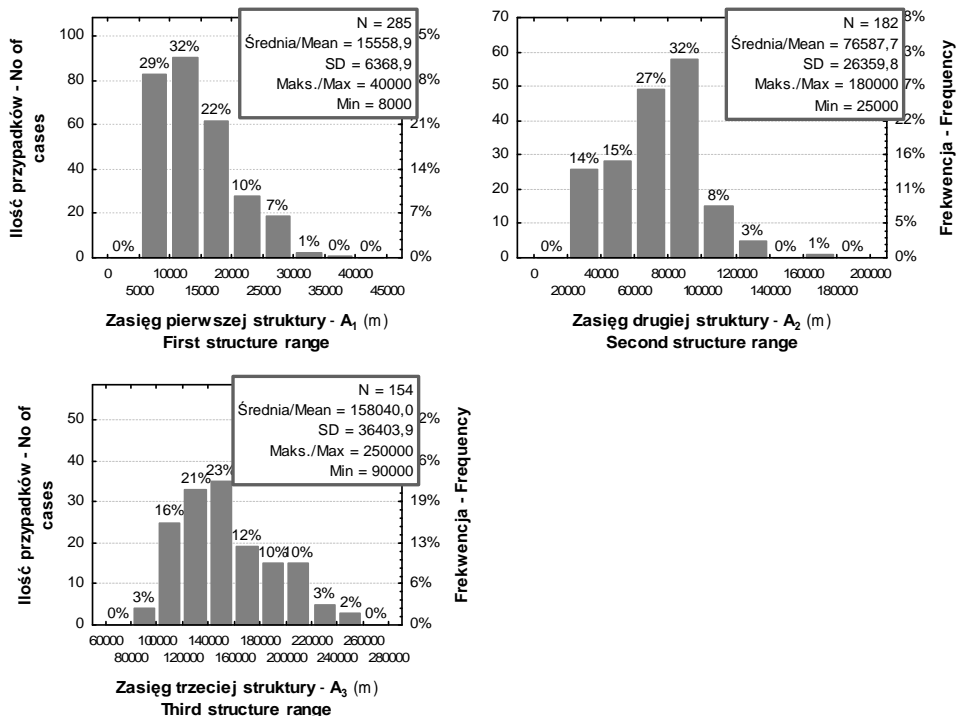


**Ryc. 59.** Frekwencje wielkości semiwariancji nuggetowej (C<sub>0</sub>) oraz wielkości wariancji progowych kolejnych składowych (C<sub>1</sub>, C<sub>2</sub> i C<sub>3</sub>) modeli struktury przestrzennej znormalizowanych miesięcznych danych MSDO z wielolecia 1956-1980. Obydwie osie wyskalowano w sposób umożliwiający bezpośrednie porównanie wykresów

Drugie załamanie krzywej autokorelacji występuje przy odległościach wahających się od 25 do 180 km (średnio ok. 76,6 km, SD = 26,4 km, ryc. 60). Udział tej struktury waha się w przedziale od 1 do ponad 86%, wynosząc przeciętnie 23,4%. Również i tutaj *gros* przypadków lokuje się w lewej części histogramu (ryc. 59) – frekwencja wartości bardzo wysokich (> 40%) jest niewielka.

Zasięg trzeciej struktury podobieństwa MSDO wynosi przeciętnie około 158 km (SD = 36,4 km) i waha się w przedziale od 90 do 250 km. Udział jej w całkowitej wariancji danych wynosi średnio 41%, a jego zmienność wykazuje symetryczny rozkład od minimum 7,0 do maksimum 100,6%.

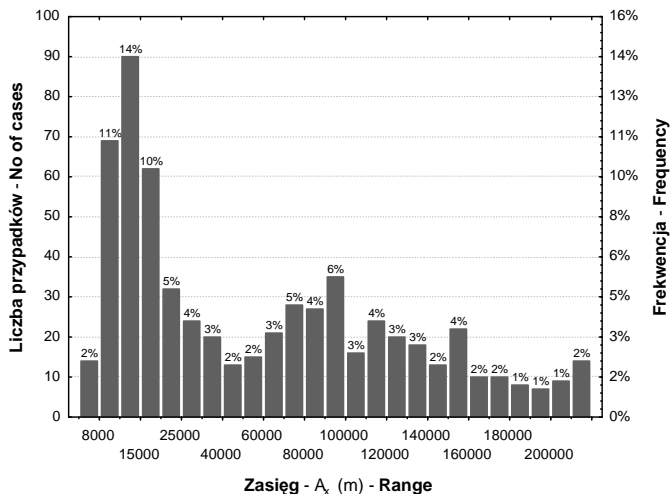
Haberlandt (2007) wykonując analizę struktury przestrzennej ekstremalnych opadów z sierpnia 2002 roku, które wywołały katastrofalną powódź w dorzeczu Łaby, zidentyfikował dla dobowych sum opadów autokorelację o zasięgu około 80/90 km. Uwzględnienie w analizie strukturalnej anizotropii wykazało, że wzdłuż głównej jej osi opady dobowe wykazywały podobieństwo na dystansie 140 km, w kierunku prostopadłym do niej – 80 km.



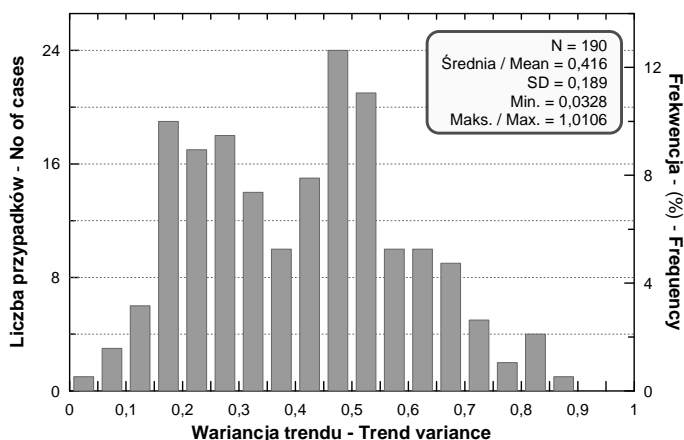
Ryc. 60. Frekwencje zasięgu kolejnych składowych ( $A_1$ ,  $A_2$  i  $A_3$ ) modeli struktury przestrzennej znormalizowanych miesięcznych danych MSDO z wielolecia 1956-1980. Jedynie oś częstości została wyskalowana w sposób umożliwiający bezpośrednie porównanie wykresów

Wykresy, które zaprezentował cytowany autor (Haberlandt 2007) wskazywałyby na istnienie również struktury o zasięgu mniejszym niż 20 km. Przyjęta wielkość jednostkowego odstepu (ang. *lag*) przy analizie danych z deszczomierzy była jednak zbyt duża, aby ten fakt można było jednoznacznie stwierdzić. Jakkolwiek wyżej cytowana publikacja dotyczyła jednostkowego przypadku, zgodność średnich zasięgów drugiej i trzeciej struktury miesięcznych MSDO, z podawanymi przez Haberlandta (2007) jest zastanawiająca. Nie można bowiem w świetle tych danych wykluczyć hipotezy, że mamy do czynienia ze statystycznym obrazem działania tego samego anizotropowego zjawiska - intensywnych opadów frontalnych. Zatem, w uzyskanych modelach izotropowych raz pojawia się wyraźnie autokorelacja o zasięgu zgodnym z głównym kierunkiem anizotropii, raz prostopadła do niej; w innym zaś przypadku - obie wartości. Efekt zależy od liczby i konfi-

guracji stanowisk objętych opadem (czy strefa opadowa w całości, czy w części znajdowała się na terytorium naszego kraju). Należy również zauważyć, że między parametrami składowej drugiej i trzeciej (wariancjami cząstkowymi  $C_2$  i  $C_3$  i zasięgami cząstkowymi  $A_2$  i  $A_3$ ) istnieją słabe, ale statystycznie istotne (odpowiednio  $p = 0,0028$  i  $p = 0,0039$ ) relacje. Wzrost  $C_2$  pociąga za sobą spadek  $C_3$ . Zależność  $A_2$  i  $A_3$  jest wprost proporcjonalna.



Ryc. 61. Częstości zasięgów wszystkich składowych ( $A_1$ ,  $A_2$  i  $A_3$  łącznie) modeli struktury przestrzennej znormalizowanych miesięcznych danych MSDO z wielolecia 1956-1980

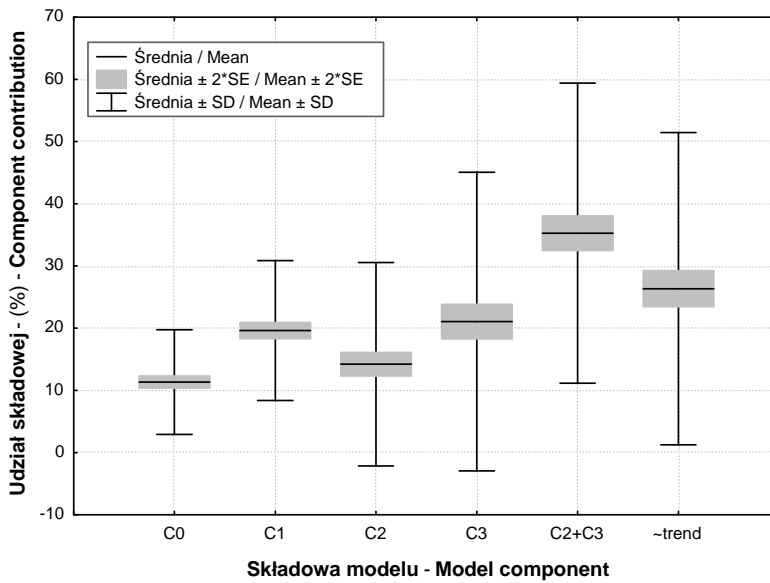


Ryc. 62. Częstość wartości wariancji składowej trendu oceniona dla odstępów 212,5 km

Ocena zasięgu trendu występującego w wyróżnionych 1., 2. i 3. typie struktury przestrzennej MSDO może być jedynie szacunkowa. Prawdopodobnie są to odległości rzędu 250-350 km. To, jaka część zmienności maksymalnych sum dobowych opadów jest skutkiem istnienia zjawisk o takim zasięgu można również tylko szacować. Jako podstawę do takiej oceny przyjęto wartość semiwariancji składowej trendu, wyliczonej z modelu dla arbitralnie przyjętej odległości 212,5 km (ryc. 62). Średnia udziału składowej trendu w 190 przypadkach miesięcznych struktur MSDO, w których wystąpiła, wynosił przy powyższym założeniu 41,6% (SD = 18,9%) i wahał się 3-101%. Histogram tych wartości (ryc. 62) ma wyraźnie charakter bimodalny z jednym maksimum wynoszącym około 20, a drugim – 50%. Cecha ta wykazuje pewien związek z numerem „porządkowym” struktury zawierającej składową trendu. Średni udział tej składowej w strukturze pierwszej wynosi bowiem 56,7%, a w trzeciej – 23,4%. Zakresy międzykwartylowe tych wartości kompletnie się nie pokrywają. Jednakże, najbardziej liczba struktura druga sama wykazuje również bimodalność rozkładu udziału składowej trendu o bardzo podobnym do całości kształcie. Być może ma to związek z jakimś zróżnicowaniem genetycznym wysokich opadów o dużej rozciągłości przestrzennej.

Składowe modeli rocznych MSDO różnią się wyraźnie swoimi przeciętnymi zasięgami od tych, które stwierdzono dla zbiorów miesięcznych. Są one regularnie nieco dłuższe.  $A_1$  wynosi bowiem średnio 17,5 km (9,0-30,0),  $A_2$  – 84,6 km (30,0-130,0) a  $A_3$  – 173,8 km (140,0-230,0).

Ponieważ wszystkie składowe oprócz nuggetowej występują jedynie w części analizowanych zbiorów modeli, ich rzeczywiste znaczenie w okresie wieloletnim można ocenić poprzez takie zestawienie, w którym brakującemu w danym przypadku elementowi przypisywana jest wartość zero. Pierwszym etapem takiej analizy jest porównanie na jednym wykresie frekwencji wszystkich wyróżnionych zasięgów autokorelacji miesięcznych MSDO (ryc. 61). Wyraźnie widoczne są na nim dwie, wcześniej już sygnalizowane prawidłowości. Po pierwsze, zdecydowanie ilościowo dominują MSDO o krótkim zasięgu (średnio ok. 15 km), a po drugie rozkład frekwencji opadów obejmujących obszary o średnicy 50-150 km wskazuje na ich identyczną genezę (konsekwentne zmiany frekwencji w poszczególnych klasach ze śladem bimodalności). Potwierdzałoby to sugerowany powyżej związek tych struktur z anizotropowymi opadami frontalnymi. Po depresji frekwencji zasięgów w przedziale 180-200 km zaznacza się na wykresie (ryc. 61) ich lekki wzrost. Jakkolwiek może to być efekt przypadkowy, to jednak nie powinno się wykluczać możliwości, że sygnalizowane jest w ten sposób częstsze występowanie opadów o tak długim zasięgu autokorelacji (~ 250-300 km).



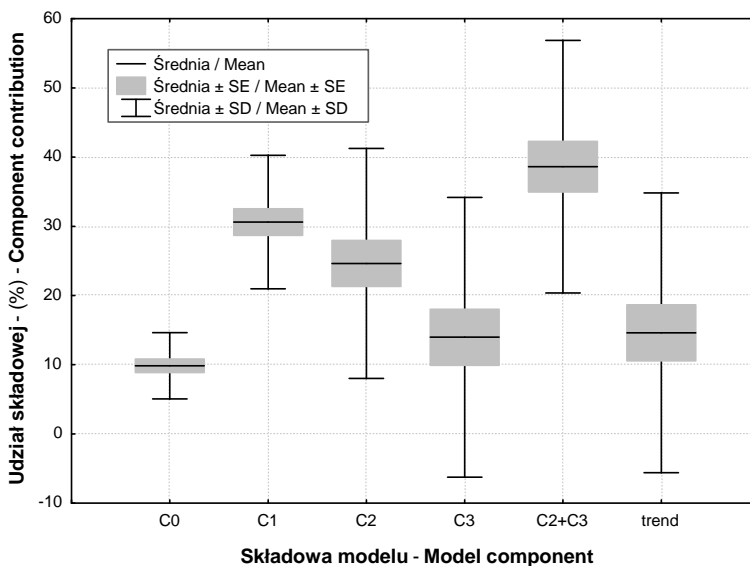
Ryc. 63. Udział poszczególnych składowych w całkowitej zmienności przestrzennej miesięcznych MSDO w wieloletniu 1956-1980

Porównanie pierwiastków kwadratowych z wariancji poszczególnych składowych przedstawiono na rycinie 63. Wartości te można w tym momencie interpretować jako udział każdej z nich w całkowitej zmienności analizowanej cechy. Ponieważ składowe są efektem działania różnych procesów, prawdopodobnie o zróżnicowanej genezie, odzwierciedlają one ich udział w całkowitej zmienności MSDO. W rozdziale V wykazano istnienie bardzo silnej, wprost proporcjonalnej zależności pomiędzy średnimi i odchyleniami standardowymi zarówno statystyk globalnych, jak i lokalnych MSDO. Można zatem z dużym prawdopodobieństwem przypuszczać, że wysoka zmienność oznacza także wysokie bezwzględne sumy opadu.

Ponieważ, jak wspomniano powyżej, składowa nuggetowa występowała w każdym z 300 analizowanych modeli, jej udział nie uległ zmianie. Zmalała, ale bardzo nieznacznie (285 przypadków na 300), do poziomu 19,6% (SD = 11,2%), rola pierwszej składowej. Znacznie większa redukcja znaczenia dotyczy składowej drugiej i trzeciej, które notowano odpowiednio w 182 i 154 miesiącach. Ich udział spadł do 14,2 (SD = 16,4%) i 21,1% (SD = 24,0%). Jeśli jednak poprawna jest hipoteza zakładająca, że stanowią one w rzeczywistości rezultat działania tego samego anizotropowego procesu, to w sumie ich znaczenie jest największe: 35,3% (SD = 24,1%). Ich sumaryczna frekwencja, czyli liczba miesięcy, w których notowano albo jedną, albo drugą, albo

obie jednocześnie jest również bardzo wysoka, bo wynosi 87,7% (263 miesiące). Udział trendu, oceniony według omówionej wyżej procedury i po uwzględnieniu przypadków, w których nie był notowany, wynosi średnio dla analizowanego wielolecia 26,3% (SD = 25,1%). Suma wariancji wszystkich składowych dla całego zbioru 300 miesięcy wynosi średnio 92,6% całej wariancji próby. Brakujące (do 100%) 7,4% (SD = 10,9%) to zmienność pola MSDO o zasięgu przekraczającym 212,5 km, która powinna zostać dodana do wariancji trendu. Tak więc, wieloletnie proporcje udziału zmienności poszczególnych składowych ( $C_0$ ,  $C_1$ ,  $C_2 + C_3$ , trend) tworzących pole miesięcznych MSDO na obszarze Polski wynoszą jak 1 : 1,73 : 3,11 : 2,98.

Podobne zestawienie dokonane dla rocznych MSDO (ryc. 64) daje obraz różniący się w kilku istotnych szczegółach. Przede wszystkim, zdecydowanie wyższy jest udział składowej pierwszej (średnia - 30,6%, SD = 9,6%), a mniejszy trendu (średnia - 14,6%, SD = 20,2%). Odwrotna jest również relacja między składową drugą a trzecią - ta pierwsza jest tym razem większa. Wynika to z faktu, że w modelach struktury przestrzennej rocznych MSDO składowa pierwsza i druga występowały w każdym roku z analizowanego wielolecia, podczas gdy strukturę trzecią i trend notowano tylko odpowiednio w 9 i 11 przypadkach. Suma elementu drugiego i trzeciego jest zbliżona do tej, jaką stwierdzono dla zbioru danych miesięcznych - wynosi bowiem



Ryc. 64. Udział poszczególnych składowych w całkowitej zmienności przestrzennej rocznych MSDO w wieloleciu 1956-1980

średnio 38,6% (SD = 18,2%). Z analizowanego zbioru 25 modeli struktury przestrzennej rocznych MSDO wynika, że w zakresie do 212,5 km było uchwycone 93,7% całej ich zmienności. Brakujące 6,3% powinno zostać, jak to zrobiono poprzednio, dodane do wariancji trendu. Podobnie jak w wypadku danych miesięcznych, zestawiono proporcje udziału zmienności poszczególnych składowych ( $C_0$ ,  $C_1$ ,  $C_2 + C_3$ , trend) tworzących pole rocznych MSDO na obszarze Polski i uzyskano szereg 1 : 3,11 : 3,92 : 2,12.

Pole MSDO na obszarze Polski jest zatem najczęściej efektem sumowania się skutków działania trzech (a może czterech) typów procesów operujących w różnych skalach przestrzennych: lokalnej (< 10-20 km), regionalnej (50-150 km) i ponadregionalnej (> 200 km). Wstępnie można postawić hipotezę o związku wyróżnionych skal przestrzennych z konwekcyjną/orograficzną, synoptyczną (frontalną) i „klimatologiczną” genezą wysokich opadów. Ich udział jest bardzo zmienny i zróżnicowany, w zależności od okresu agregacji - miesięcznego czy rocznego. Generalnie dominują jednak wysokie sumy dobowe opadów o rozciągłości przestrzennej od 50 do 150 km, związanych z migracją frontów atmosferycznych (35-38%). Przy miesięcznym kroku analizy MSDO, na drugim miejscu, z przeciętnym udziałem około 32%, sytuują się opady o zasięgu ponadregionalnym, a trzecim - konwekcyjno/orograficzne (ok. 20%). W ujęciu rocznym kolejność i rola tych dwóch składowych jest odwrotna. Należy również zwrócić uwagę, że jakkolwiek prawdopodobnie największy udział w składowej nuggetowej mają błędy pomiarowe i nieciągłość/asynchroniczność MSDO, to jednak pewna, nieznaną jej część związana jest z rzeczywistą zmiennością opadów na dystansie krótszym niż minimalny odstęp stanowisk i powinna być „doliczona” do składowej pierwszej.

Ważny punkt odniesienia do uzyskanych w niniejszym opracowaniu charakterystyk struktury przestrzennej sum dobowych opadów na obszarze Polski stanowi cytowana już na początku rozdziału praca Moszkowicza (2000). Przy porównaniach pamiętać jednak należy o jej, wspomianej uprzednio, mocno ograniczonej, czasowo i przestrzennie, reprezentatywności, jak też innym znaczeniu, jakie ów autor nadaje niektórym terminom. Nie można na przykład „promień korelacji” (ang. *radius of correlation*) Moszkowicza (2000) utożsamiać z zasięgiem autokorelacji używanym w niniejszej pracy. U wspomnianego bowiem autora oznacza to odległość, przy której funkcja autokorelacji spada poniżej 0,5.

Z izotropowej analizy pomiarów z sieci deszczomierzy cyfrowych wynika (fig. 1 w: Moszkowicz 2000), że zasięg autokorelacji, niezależnie od przyjętego kroku agregacji opadu (10, 30 minut, 1, 3, 6, 12 i 24 godziny), jest z pewnością większy niż 5 km. Na tej odległości współczynnik korelacji wynosi jeszcze od około 0,4 dla opadów 10-minutowych, do około 0,85 - dla dobowych. Spadek korelacji wraz z odległością jest zatem tym większy, im



czas kumulacji opadu jest krótszy. Ten sam czynnik wpływa również na zmniejszanie się błędu losowego (określanego w niniejszej pracy jako wariancja nuggetowa), który przy 24-godzinnym kroku agregacji wynosi praktycznie 0. Wartość ta nie jest zaskoczeniem, biorąc pod uwagę fakt, że analizowano pojedyncze, ciągle epizody opadów, a w deszczomierzach rejestrujących praktycznie nie ma strat związanych z parowaniem i zwilżaniem.

Dane z radaru meteorologicznego w Legionowie analizował Moszkowicz (2000) oddzielnie dla opadów konwekcyjnych i frontalnych (ang. *stratiform*) oraz stosując zarówno izotropowe, jak i anizotropowe funkcje autokorelacji. Jest to, jak wspomniano uprzednio, opracowanie uwzględniające jedynie pojedyncze, wyselekcjonowane przypadki. Bardzo istotny jest także fakt, że kierunek anizotropii nie był określany empirycznie z danych, ale przyjmowany *a priori* na podstawie znanego kierunku adwekcji pola opadów<sup>38</sup>. Nie zawsze jednak może to być ściśle. Z zamieszczonych w omawianym artykule (Moszkowicz 2000) wykresów i zestawień tabelarycznych wynika znacząca różnica struktury przestrzennej opadów konwekcyjnych i frontalnych. Dotyczy to zarówno ich zasięgu, jak i anizotropii. Ta druga cecha w wypadku opadów konwekcyjnych wydaje się być stabilna niezależnie od kroku czasowego agregacji opadów. Stosunek anizotropii wynosił bowiem od około 0,25 do 0,3 ( $Y/X$ , gdzie  $Y$  to oś prostopadła do kierunku adwekcji, a  $X$  – równoległa). Przy opadach pochodzących z chmur warstwowych Moszkowicz (2000) stwierdził konsekwentne zmniejszanie się anizotropii wraz ze zwiększaniem czasu sumowania opadów, poczynając od 0,438 przy 20 minutach, do 1,000 przy 12 godzinach. Po połączeniu obu zbiorów danych (opady konwekcyjne i frontalne) współczynnik anizotropii oscylował około wartości 0,3 dla czasów poniżej 2 godzin, a następnie wzrastał, osiągając 0,77 dla opadów 24-godzinnych. Opierając się zatem na tych danych i pamiętając ciągle o niewielkiej próbie, jaką dysponował Moszkowicz (2000), można próbować, na podstawie stosunku anizotropii analizowanego pola opadów, szacować udział składowej konwekcyjnej i frontalnej.

Analiza izotropowa była przeprowadzona w promieniu 70 km, a anizotropowa – 50 km. W zasadzie, tylko w wypadku opadów konwekcyjnych można było na takim dystansie stwierdzić, jaki był ich zasięg autokorelacji. Niestety, autor omawianego opracowania (Moszkowicz 2000) zamieścił wykresy tylko dla wybranych przedziałów czasu agregacji opadów. Dla najkrótszego analizowanego izotropowo kroku czasowego (10 minut) charakterystyczny jest bardzo szybki spadek autokorelacji opadów konwekcyjnych wraz z odległością, o kształcie funkcji wykładniczej osiągającej praktyczny zasięg około 15 km. Również krzywa dla opadów frontalnych przy tej odle-

---

<sup>38</sup> Główne osie anizotropii równoległe i prostopadłe do kierunku adwekcji.

głości wykazuje wyraźne załamanie – zmniejszenie tempa spadku. Ten sam próg odległości widoczny jest także na wykresie autokorelacji opadów konwekcyjnych z godzinnego okresu sumowania. Tym razem jest to jednak tylko silne załamanie krzywej przy autokorelacji wynoszącej około 0,3, natomiast zasięg osiągany – przy odległości przekraczającej 50 km. Opady z chmur warstwowych w tym samym przedziale czasu wykazują liniowy spadek podobieństwa z odległością; przy odległościach rzędu 70 km współczynnik korelacji wynosi jeszcze 0,2. Prawie identycznie układa się krzywa autokorelacji tego typu opadów także przy 24-godzinnym okresie ich sumowania. Znaczące są natomiast, przy tym samym skoku interwału czasu, zmiany struktury opadów konwekcyjnych. Spada bowiem rola struktury krótkodystansowej, a także zmniejsza się jej zasięg, do około 6-7 km. Pojawia się za to załamanie krzywej autokorelacji, znaczące prawdopodobnie praktyczny zasięg opadów konwekcyjnych, przy odległości rzędu 55 km.

Interpretacja wyników uzyskanych przez Moszkowicza (2000) w trakcie analizy anizotropowej jest niestety utrudniona, ze względu na niską jakość wykresów w dostępnej elektronicznej kopii owej publikacji<sup>39</sup>. Można z nich jednak wyczytać, że przy krótkich czasach agregacji (do 1 godziny), długość krótszej osi anizotropii opadów konwekcyjnych wynosi około 10 km<sup>40</sup>. W trakcie opadów trwających 24 godziny struktura przestrzenna jest bardziej złożona. Po pierwszym załamaniu na dystansie około 15 km, krzywa autokorelacji dla krótszej osi w dalszym ciągu maleje, osiągając zasięg przy odległościach rzędu 30-35 km. Opady z chmur warstwowych wykazywały, jak wspomniano już poprzednio, anizotropię przy czasach poniżej 12 godzin. Jej układ jest bardzo charakterystyczny. Spadek autokorelacji wzdłuż krótszej osi ma bowiem w początkowym odcinku charakter liniowy, a wzdłuż dłuższej – paraboliczny. Wskazywałoby to, że na dystansie 5-7 km na kierunku adwekcji zmienność przestrzenna opadów z chmur warstwowych jest znikomo mała i praktycznie niemierzalna. Zasięg autokorelacji z danych przedstawionych przez Moszkowicza (2000) można określić jedynie wzdłuż krótszej osi anizotropii, która wynosi, jak się wydaje, około 45-50 km.

Przedstawione powyżej, obszerne omówienie analizy struktury przestrzennej opadów, zawartej w pracy Moszkowicza (2000), a opartej na ciągłej rejestracji naziemnej i pomiarach radarowych, potwierdza w wielu punktach wnioski uzyskane na podstawie danych MSDO. Szczególnie powtarzająca się w kilku kontekstach odległość 15 km dla zasięgu cząstkowego opadów konwekcyjnych świadczy o dużym znaczeniu owej struktury, powiązanej z pewnością z mechanizmem tworzenia się, rozwoju i zaniku komórek cyrkulacji konwekcyjnej w atmosferze. Charakteryzując znaczenie poszczegól-

---

<sup>39</sup> Pozyskanej z bazy danych ScienceDirect.

<sup>40</sup> Zasięgu dłuższej osi anizotropii z zamieszczonych wykresów odczytać nie sposób.

nych składowych struktury przestrzennej MSDO wspomniano, że pewna, nieznaną, część zmienności losowej (wariancji nuggetowej) jest związana z krótkodystansowym zróżnicowaniem opadów na odległościach mniejszych niż minimalny odstęp stanowisk. W całkowitym bilansie powinna ona być doliczona do pierwszej składowej. Dane Moszkowicza (2000) wskazują, że zmienność krótkodystansowa może stanowić nawet 50% wariancji nuggetowej<sup>41</sup>!

## **6. Klasyfikacja struktury przestrzennej MSDO - wartość semiwariancji**

Przeprowadzono ją następująco: najpierw, korzystając z programu Vmodel należącego do biblioteki GSLIB (Deutsch, Journal 1998), obliczono wartości modeli wszystkich miesięcznych i rocznych zbiorów danych co 0,5 km do odległości 212,5 km (425 liczb)<sup>42</sup>. Następnie do oceny ich podobieństwa zastosowano algorytm klasyfikacji hierarchicznej Warda (ryc. 65-67). Ponieważ modele były wyliczone z danych znormalizowanych, jako miarę podobieństwa można było wykorzystać zwykłą odległość euklidesową. Aby ocenić optymalną liczbę grup, wykorzystano krzywą aglomeracji pokazującą, jak w miarę łączenia poszczególnych obiektów i ich grup rośnie odległość wiązania (spada podobieństwo). Jako kryterium podziału przyjęto odległość wiązania, przy której następuje „przebiegnięcie” krzywej aglomeracji (ryc. 68). Oznacza to, że łączone są już wówczas obiekty o bardzo małym stopniu podobieństwa. Kryterium to pozwoliło na jednoznaczne wyróżnienie trzech grup w wypadku modeli rocznych (ryc. 67 i 68). Tak klarownego podziału nie można było dokonać w odniesieniu do modeli struktury przestrzennej miesięcznych MSDO. Gwałtowny spadek podobieństwa nastąpił bowiem już przy relatywnie małej odległości wiązania obiektów. Wydaje się jednak, że podział na 7 grup jest bliski optymalnemu. Dla porównania wykonano również wykresy i zestawienia tabelaryczne dla podziału na 4 grupy, który na dendrycie (ryc. 65) jest bardzo wyraźny. Grupa 2 w obu podziałach jest identyczna, a grupa 4 z czterech odpowiada 5 z siedmiu. Należy jednakże przy interpretacji wyników tych klasyfikacji pamiętać, że struktura przestrzenna MSDO w analizowanym wieloleciu, opisywana za pomocą modeli semiwariogramów, nie wykazuje jakiegos drastycznego zróżnicowania (por.

---

<sup>41</sup> Cytowany autor (Moszkowicz 2000) w tekście swojego artykułu pisze, że istnienia efektu nuggetowego nie stwierdził, podczas gdy z zamieszczonych wykresów wynika coś przeciwnego.

<sup>42</sup> Obliczone semiwariancje modelowe oraz wyniki klasyfikacji zamieszczono na dołączonym dysku DVD.

ryc. 50, 51 i 53). To raczej *continuum* stopniowych zmian od jednego ekstremum do drugiego. Wyróżnianie ostrych granic jest w tym przypadku zawsze dość arbitralne. Z pewnością bardziej odpowiednio byłoby w takim wypadku zastosowanie klasyfikacji rozmytych, gdzie przynależność do konkretnej klasy jest wyrażana w kategoriach prawdopodobieństwa.

Wyróżnione klasy podobieństwa wartości autokorelacji przedstawiono na rycinach 69, 70 i 76. Obrazują one średnią wartość semiwariancji modeli należących do danej grupy i zakres jednego odchylenia standardowego. Wykresy te jednoznacznie wskazują, że uzyskane grupy różnią się nie tylko bezwzględną wartością semiwariancji, ale także kształtem ich przebiegu. Wskazywać to może, że różnice te mogą mieć także charakter jakościowy – genetyczny. Z drugiej strony, zakresy zmienności grup w dużym stopniu się pokrywają, świadcząc, że mamy do czynienia raczej z „mieszaniną”, w której proporcje poszczególnych składowych ulegają ciągłym zmianom.

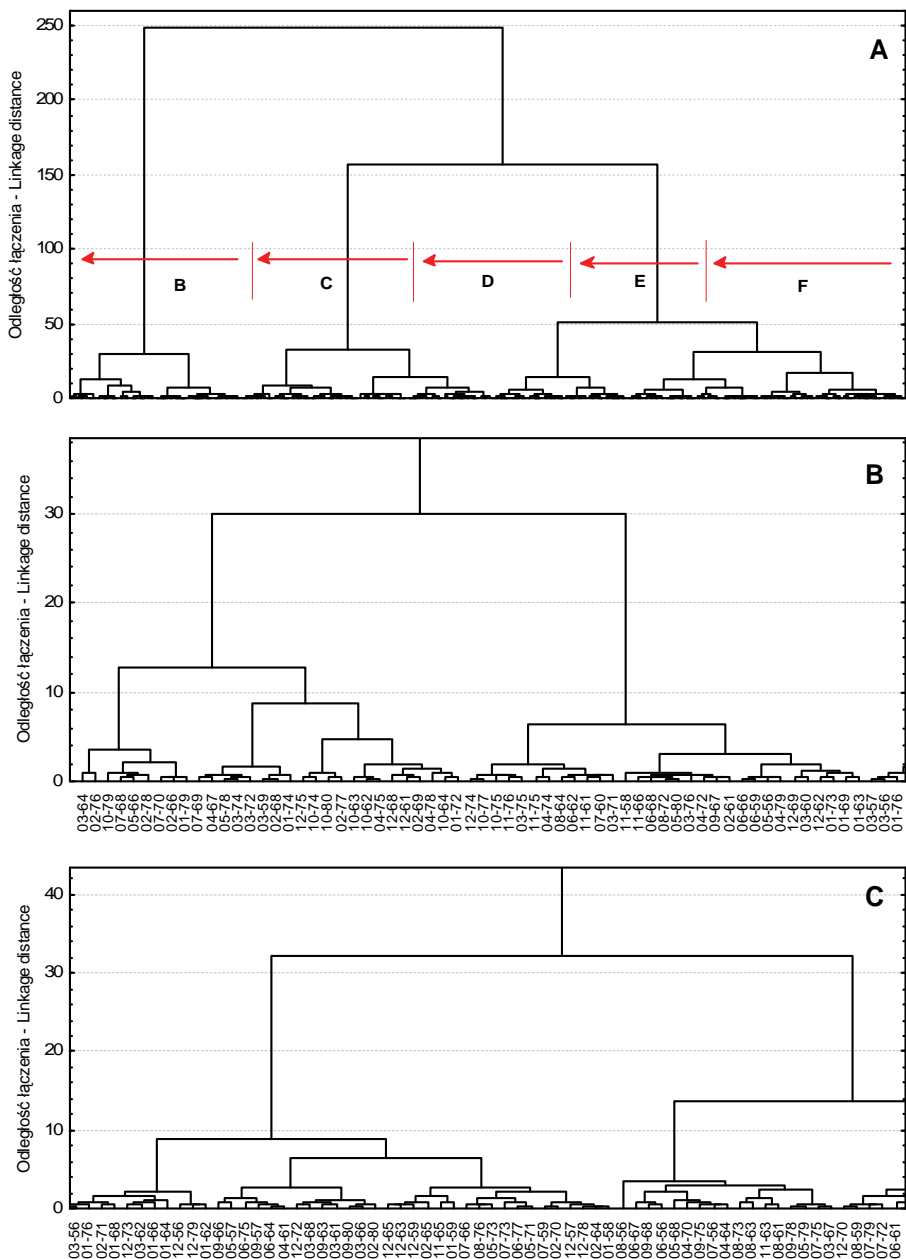
Z modeli należących do każdej grupy wybrano jeden – najbardziej zbliżony do średniej. Jako kryterium użyto minimum sumy kwadratów odchyleń między poszczególnymi, dla odstępów o szerokości 0,5 km, wartościami kolejnych modeli a średnią grupową. Każdy z tak wytypowanych modeli struktury przestrzennej miesięcznych lub rocznych MSDO został zaznaczony na wyżej wymienionych wykresach (ryc. 69, 70 i 76). Przede wszystkim jednak przypadki te przedstawiono w postaci map symulowanych, w przestrzeni danych znormalizowanych, pół opadów (ryc. 71-75 oraz 77 i 78).

Pojawianie się wyróżnionych typów modeli miesięcznych MSDO przy podziale na 4 grupy nie wykazuje istotnej statystycznie zmienności sezonowej ( $p$  dla klasycznego testu  $\chi^2 = 0,190$ ; przy  $\chi^2$  obliczonym zgodnie z teorią maksymalną wiarygodności  $p = 0,088$ ). Częstość występowania poszczególnych przypadków przy podziale na 7 grup wydaje się taką zmienność wykazywać ( $p = 0,026$  i  $p = 0,007$ ). Większość wyróżnionych typów występowała jednak w każdym miesiącu, a różnice częstości były przy próbie 25-letniej dość niewielkie. Względnie wyraźnie cykl sezonowy zaznacza się przy grupie drugiej i trzeciej. W tej pierwszej maksimum widoczne jest w lipcu i sierpniu, a minimum – jesienią (październik-grudzień). W grupie trzeciej stabilne maksimum częstości występuje między grudniem a marcem; w pozostałych miesiącach liczba przypadków była generalnie niska i zmienna. Pewne oznaki zmienności sezonowej widać również w odniesieniu do grup 4, 5 i 6. W czwartej stabilne maksimum zaznacza się między lipcem a wrześniem, w piątej – w porach przejściowych (kwiecień i październik-listopad), w szóstej – wyższe częstości występowały od maja do października z kulminacją w dwóch ostatnich miesiącach.

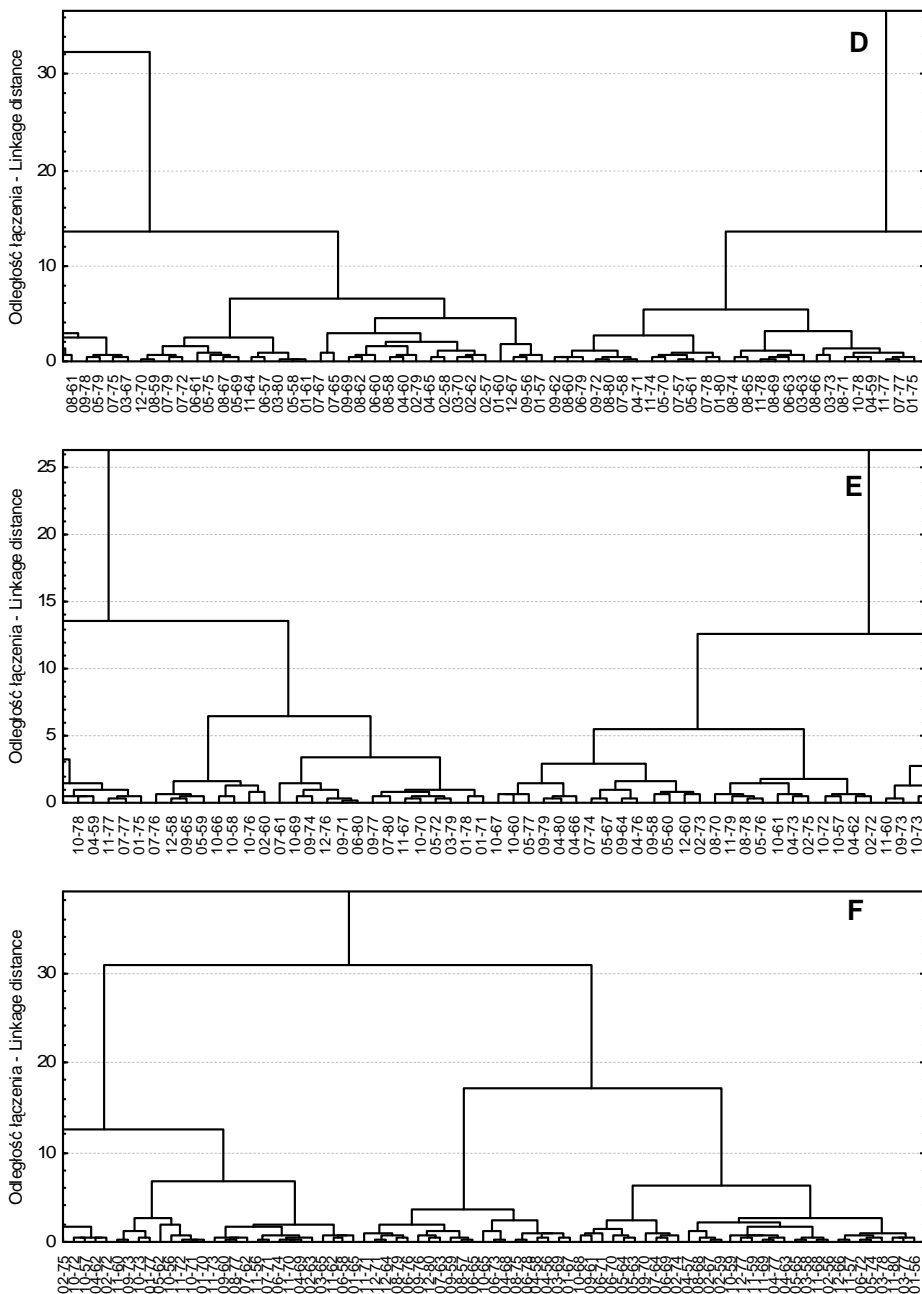
Przy podziale miesięcznych modeli struktury MSDO na 4 grupy najbardziej wyróżnia się grupa trzecia (ryc. 70), zawierająca 49 przypadków. Charakteryzuje się ona najniższą wariancją nuggetową (śr. 7,0%), strukturą


pierwszą o kształcie wykładniczym i średnim udziale 12,4% oraz liniowym wzrostem niepodobieństwa po pierwszym załamaniu krzywej. Ten kształt wskazuje na zdecydowaną dominację opadów o dużym zasięgu przestrzennym i regularnej zmienności. Składowa trendu występuje w 47 na 49 przypadków, a jej średni udział przekracza 42,3%. Ponieważ jednak suma wariancji wszystkich składowych jest w tej grupie najniższa i osiąga dla maksymalnego analizowanego zasięgu 212,5 km zaledwie 0,781, oznacza to, że rzeczywisty udział opadów wielkoobszarowych jest znacznie większy – ponad 64%. Składowa druga i trzecia w 14 przypadkach w ogóle nie występuje, a ich średni sumaryczny udział wynosi zaledwie 16,6%. Stosunki poszczególnych składowych bardzo odbiegają od średniej wieloletniej:  $C_0 : C_1 : C_2 + C_3 : \text{trend} = 1,00 : 1,77 : 2,38 : 9,13$ . Rozrzut wartości modeli semiwariancji zaklasyfikowanych do grupy 3 jest największy ze wszystkich 4 grup – od odstepu większego niż 50 km wykazuje jednak względną stałość (ryc. 70). Dla grupy 3 charakterystyczny jest także najkrótszy zasięg składowej 1 – średnio 14,2 km i najwyższy stosunek zasięgów składowej 3 do 2 wynoszący 2,17. Obraz rozkładu przestrzennego MSDO w styczniu 1974 – miesiącu o modelu semiwariancji najbardziej zbliżonym do średniej grupowej – pokazuje wyraźny podział Polski na trzy strefy (ryc. 72): Karpaty z najwyższymi sumami, obszar na wschód od Wisły z niskimi opadami i zachodnia Polska z mocno zróżnicowanymi, ale generalnie wysokimi opadami. Poza Karpatami pole MSDO tworzy generalnie mozaikę niewielkich, nieregularnych plam wyższych i niższych sum opadów.

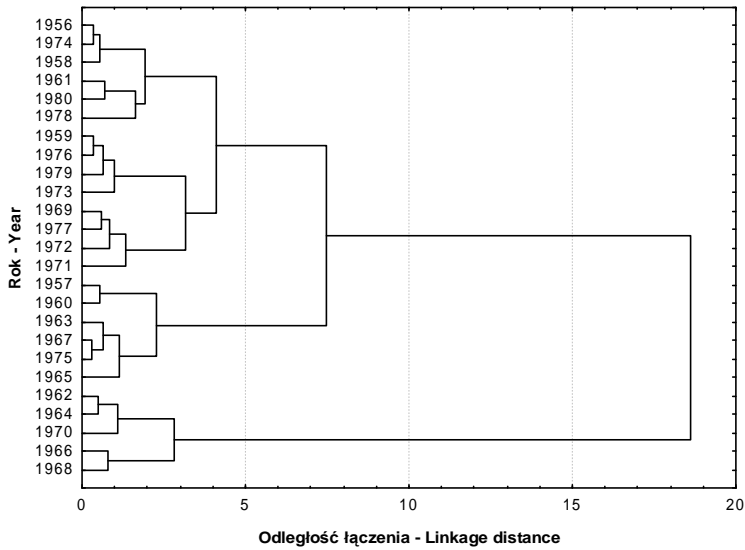
Kontrastowo odmienny charakter ma grupa 2 (ryc. 70), składająca się z 47 miesięcy. Udział semiwariancji nuggetowej jest tu najwyższy (14,1%), a składowej trendu, nawet z uwzględnieniem zmienności przekraczającej zasięg analizy – najniższy (3,9%). W tej grupie, jako jedynej, wykres średnich wartości semiwariancji wykazuje osiągnięcie, przy odległościach przekraczających 200 km, maksimum (próg). Pierwsza składowa ma kształt sferyczny, dalszy wzrost niepodobieństwa ma charakter wykładniczy. Jej udział był ponad dwukrotnie wyższy niż w grupie 3 (27,5%). Największe znaczenie mają tutaj jednak składowe druga i trzecia – w sumie 54,5%. Względne stosunki składowych wynoszą w tej grupie  $C_0 : C_1 : C_2 + C_3 : \text{trend} = 1,00 : 1,95 : 3,86 : 0,27$ . Wydaje się również, że znacząco różni obie grupy zasięg pierwszej struktury. W grupie 2 wynosi on bowiem średnio 16,7 km i jest o ponad 2 km dłuższy niż przeciętna dla grupy 3. Modele należące do tej grupy są też stosunkowo do siebie podobne (niskie odchylenie standardowe), zwłaszcza w przedziałach do 120 km. W sierpniu 1958 roku MSDO układały się w pasy wysokich i niskich wartości (ryc. 71). Były one mocno „porozrywane” i składały się z szeregu płatów o nieregularnym kształcie. „Elementarne” płaty były wyraźnie większe niż w przypadku grupy 3.



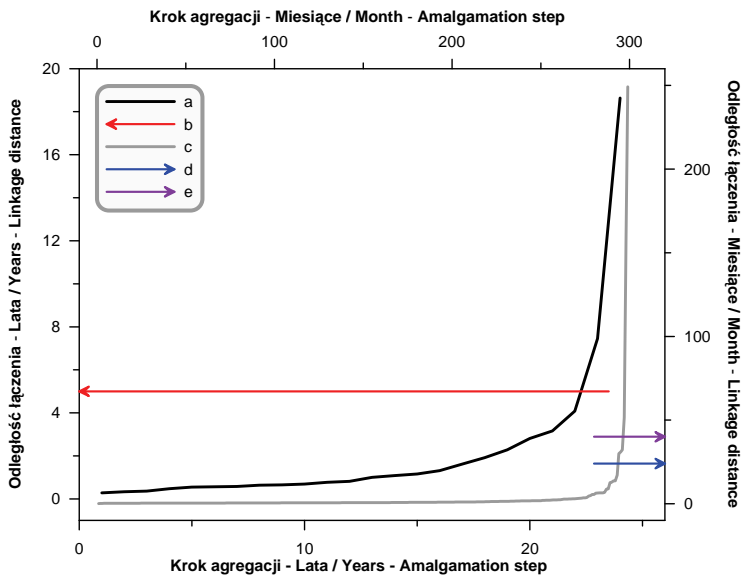
Ryc. 65. Dendryt podobieństwa modeli struktury przestrzennej znormalizowanych danych miesięcznych MSDO uzyskany metodą aglomeracji hierarchicznej Warda. A: cały dendryt; B - C - jego fragmenty powiększone tak, aby widoczne były oznaczenia poszczególnych przypadków. Przybliżony zasięg fragmentów zaznaczono czerwonymi strzałkami na rycinie A. Części D - F przedstawiono na rycinie 66



Ryc. 66. Powiększone fragmenty dendrytu podobieństwa modeli struktury przestrzennej znormalizowanych danych miesięcznych MSDO uzyskanego metodą aglomeracji hierarchicznej Warda. Cały dendryt i przybliżone zasięgi fragmentów przedstawiono na rycinie 65A. Części B i C dendrytu znajdują się również na rycinie 65 

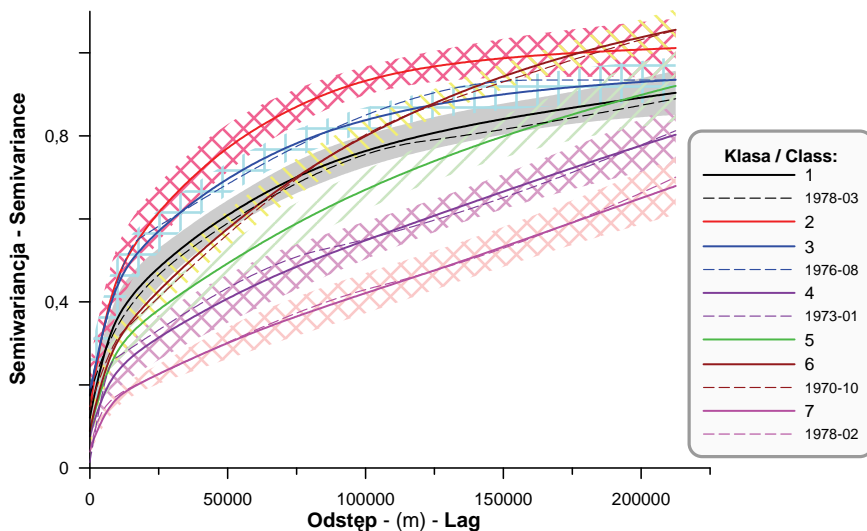


Ryc. 67. Dendryt podobieństwa modeli struktury przestrzennej znormalizowanych danych rocznych MSDO uzyskany metodą aglomeracji hierarchicznej Warda

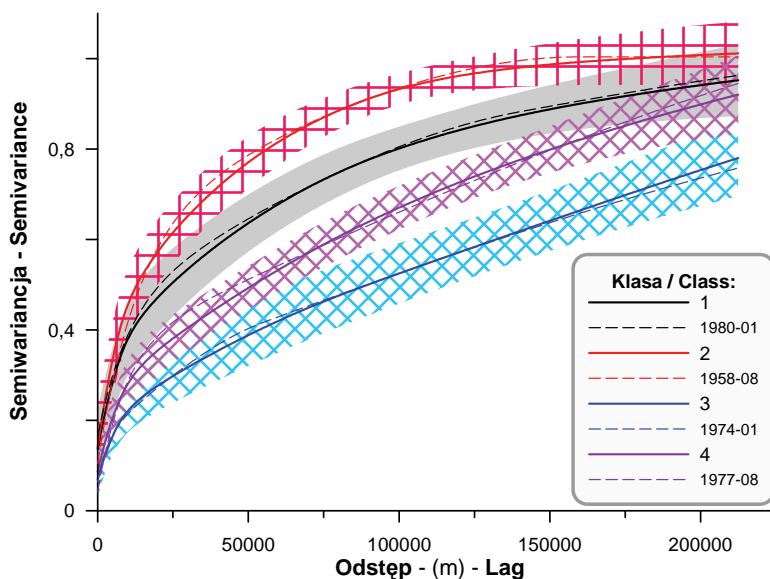


Ryc. 68. Kryterium podziału dendrytów podobieństwa modeli struktury przestrzennej znormalizowanych danych MSDO (ryc. 65A i 76): a – krzywa aglomeracji danych rocznych; b – poziom podziału dendrytu danych rocznych na 3 klasy przy odległości łączenia równej 5; c – krzywa aglomeracji danych miesięcznych; d – poziom podziału dendrytu danych miesięcznych na 7 klas przy odległości łączenia równej 24; e – poziom podziału dendrytu danych miesięcznych na 4 klasy przy odległości łączenia równej 40

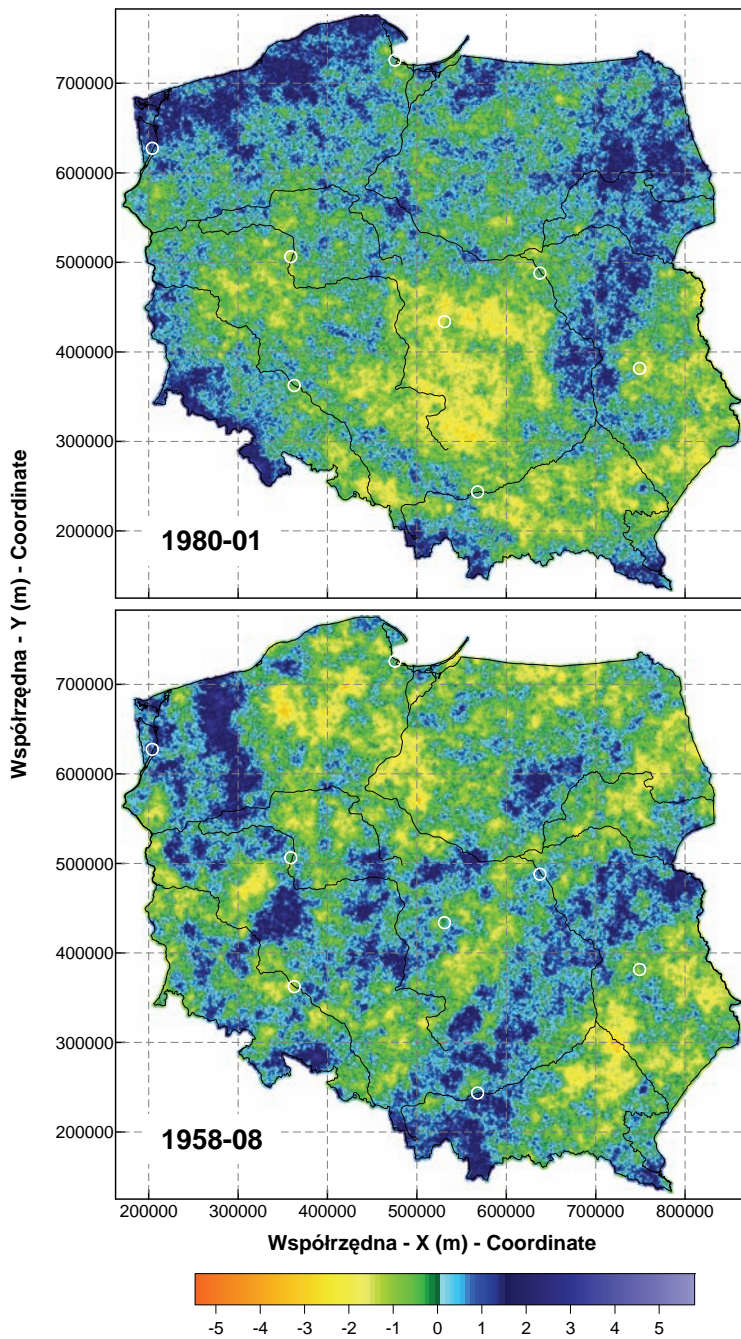





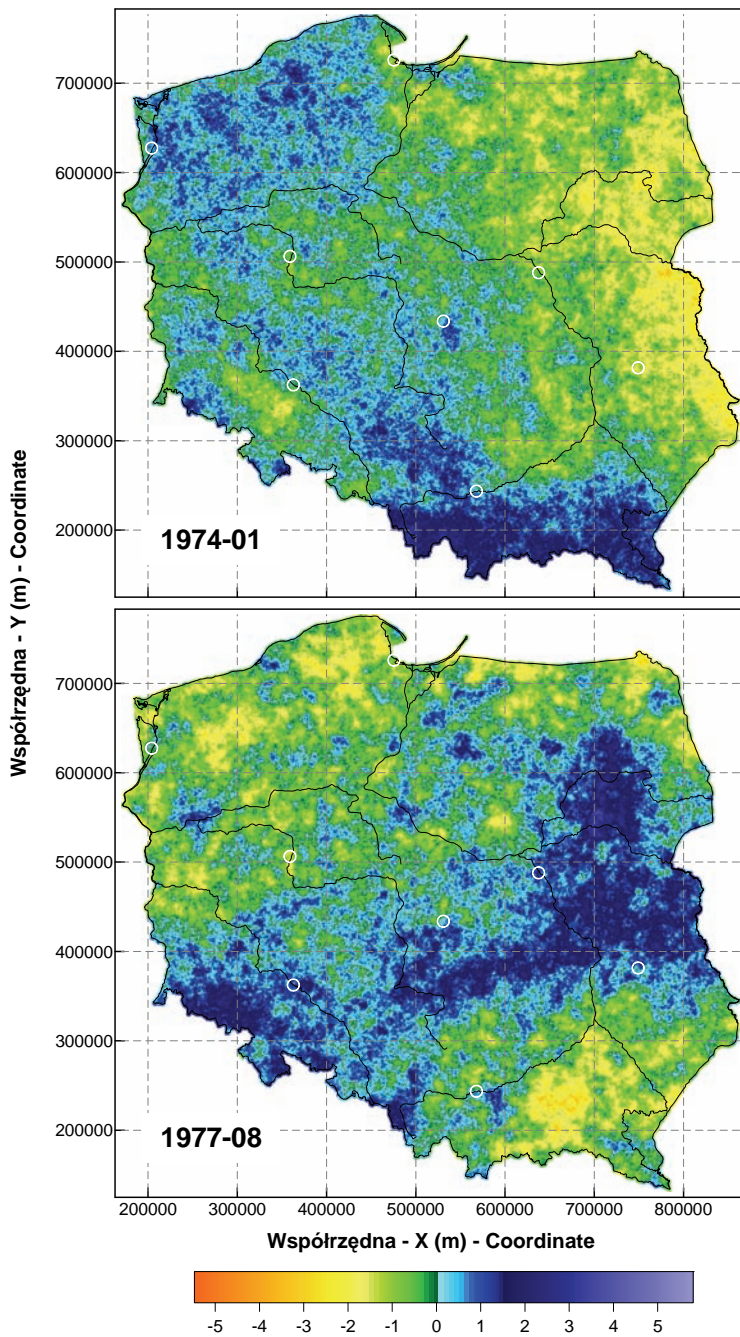
Ryc. 69. Średnie wartości semiwariancji (linie grube) dla wyróżnionych 7 klas modeli struktury przestrzennej znormalizowanych danych miesięcznych MSDO. Szrafem zaznaczono zakres jednego odchylenia standardowego, a cienkimi liniami modele wybranych miesięcy najbardziej typowych dla każdej klasy (patrz mapy na ryc. 71-75) [4]




Ryc. 70. Średnie wartości semiwariancji (linie grube) dla wyróżnionych 4 typów modeli struktury przestrzennej znormalizowanych danych miesięcznych MSDO. Szrafem zaznaczono zakres jednego odchylenia standardowego, a cienkimi liniami modele wybranych miesięcy najbardziej typowych dla każdej klasy (patrz mapy na ryc. 71-75) [4]

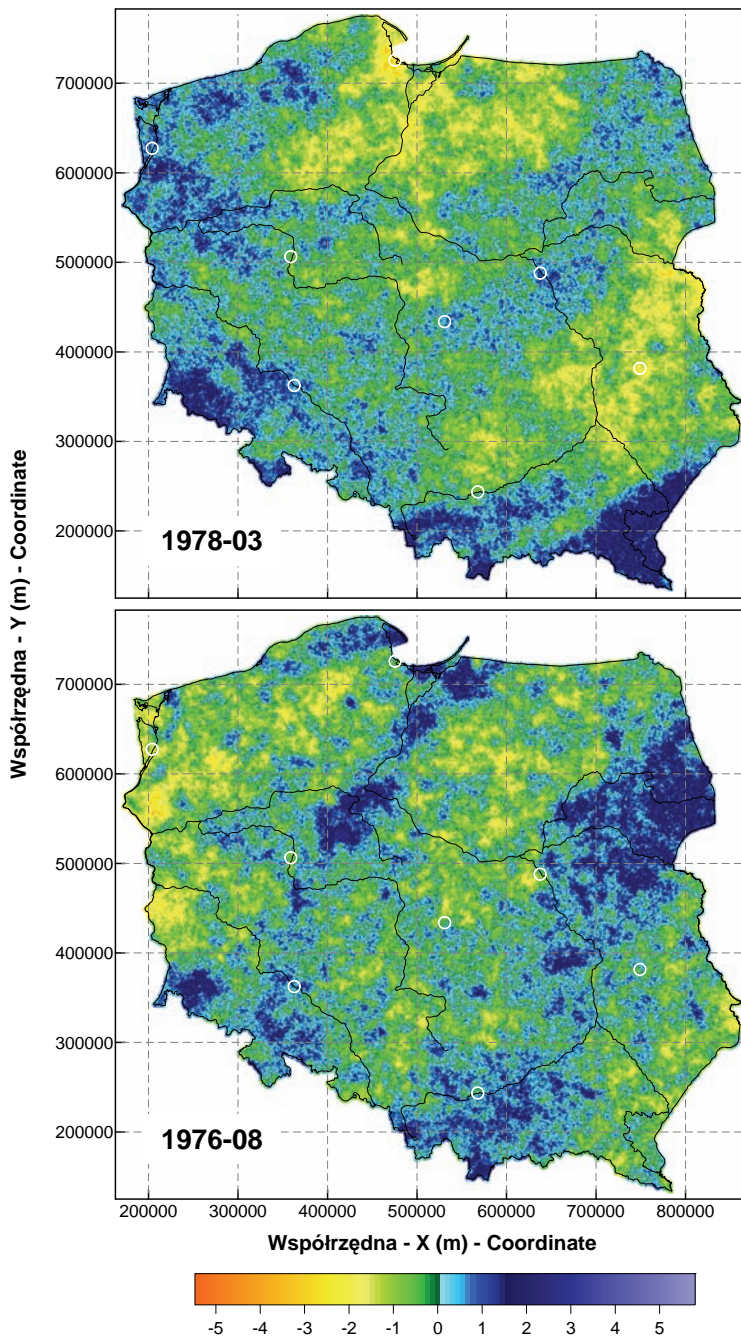



**Ryc. 71.** Pojedyncze obrazy symulacji warunkowych pola MSDO wykonanych dla danych znormalizowanych ze stycznia roku 1980 (przykład klasy 1 z 4) i sierpnia roku 1958 (przykład klasy 2 z 4 i 2 z 7) 

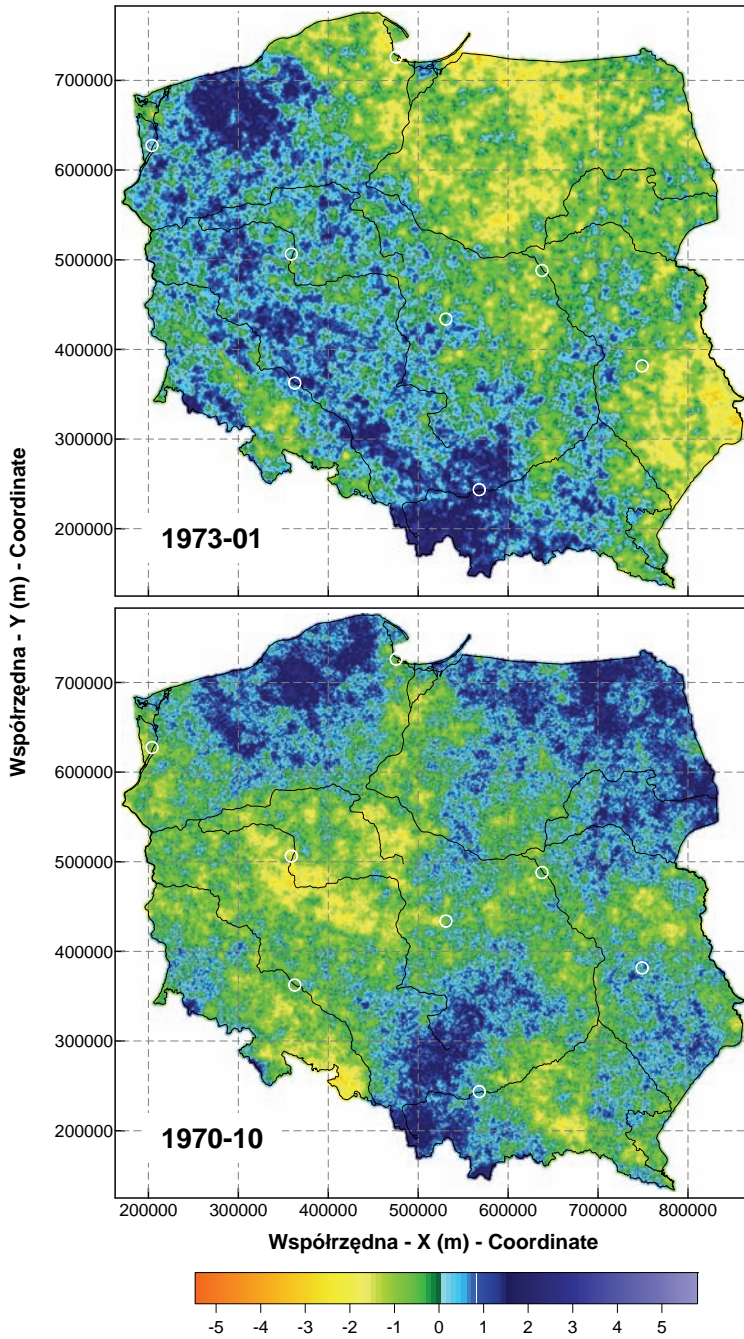


Ryc. 72. Pojedyncze obrazy symulacji warunkowych pola MSDO wykonanych dla danych znormalizowanych ze stycznia roku 1974 (przykład klasy 3 z 4) i sierpnia roku 1977 (przykład klasy 4 z 4 i 5 z 7) 

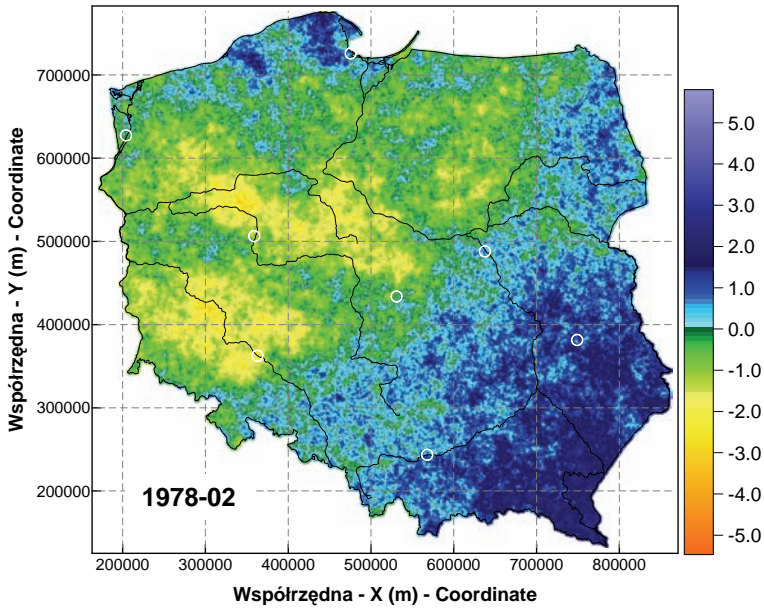





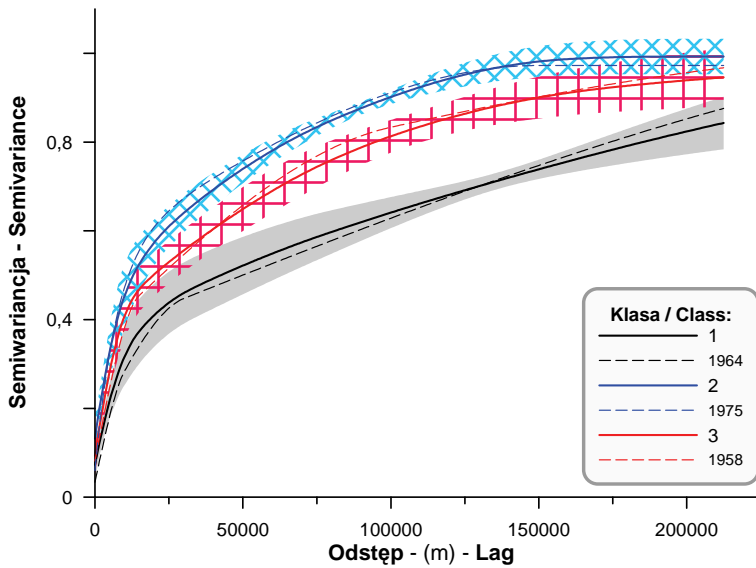
Ryc. 73. Pojedyncze obrazy symulacji warunkowych pola MSDO wykonanych dla danych znormalizowanych z marca roku 1978 (przykład klasy 1 z 7) i sierpnia roku 1976 (przykład klasy 3 z 7) 




Ryc. 74. Pojedyncze obrazy symulacji warunkowych pola MSDO wykonanych dla danych znormalizowanych ze stycznia roku 1973 (przykład klasy 4 z 7) i października roku 1970 (przykład klasy 6 z 7) 📄

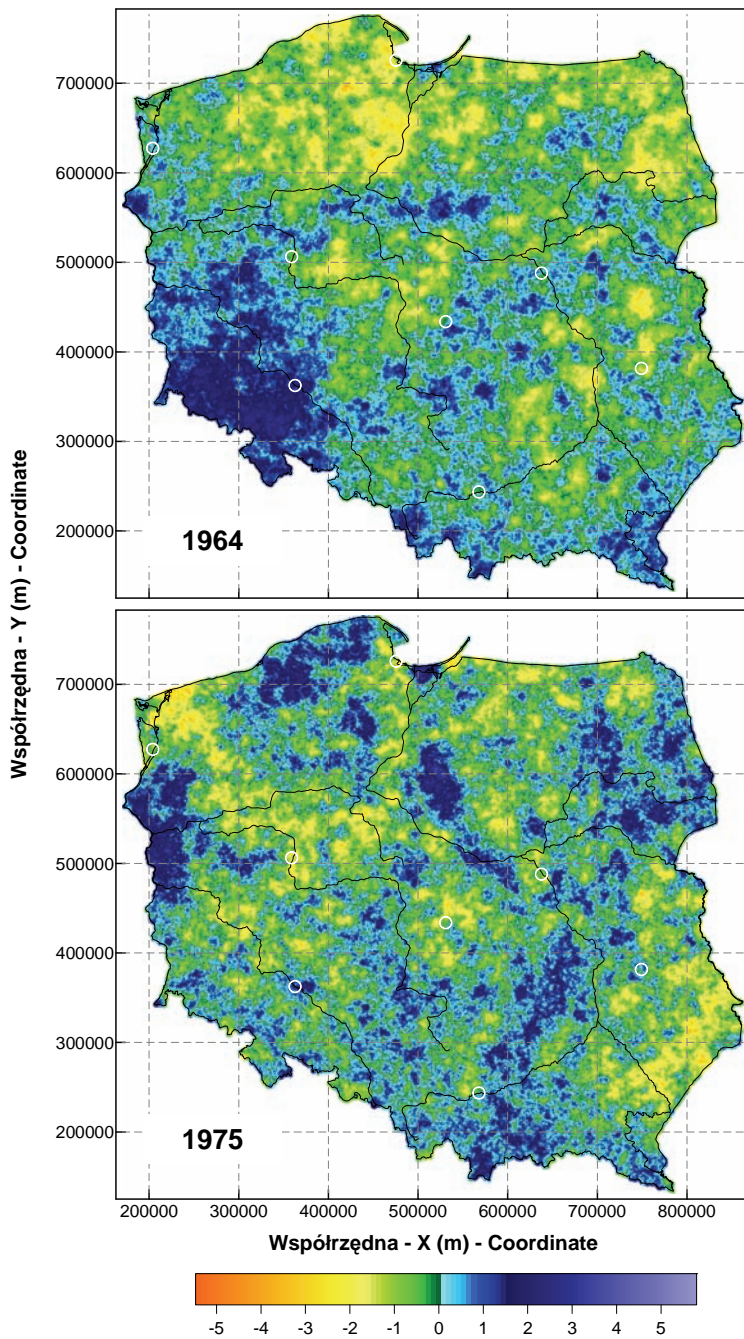



Ryc. 75. Pojedynczy obraz symulacji warunkowej pola MSDO wykonanej dla danych znormalizowanych z lutego roku 1978 (przykład klasy 7 z 7) 

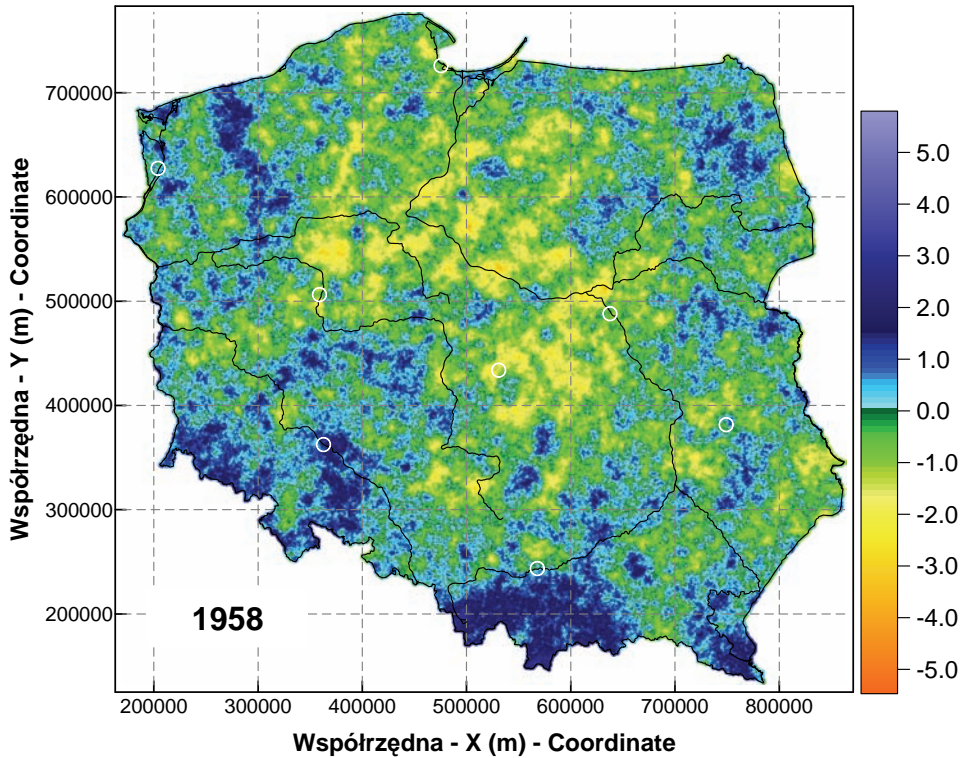



Ryc. 76. Średnie wartości semiwariancji (linie grube) dla wyróżnionych 3 typów modeli struktury przestrzennej znormalizowanych danych rocznych MSDO. Szrafem zaznaczono zakres jednego odchylenia standardowego, a cienkimi liniami modele wybranych lat najbardziej typowych dla każdej klasy (patrz mapy na ryc. 77 i 78) 





Ryc. 77. Pojedyncze obrazy symulacji warunkowych pola MSDO wykonanych dla znormalizowanych danych rocznych z roku 1964 (przykład klasy 1 z 3) i roku 1975 (przykład klasy 2 z 3) 



Ryc. 78. Pojedynczy obraz symulacji warunkowej pola MSDO wykonanej dla znormalizowanych danych rocznych z roku 1964 (przykład klasy 3 z 3) 

Grupy 1 i 4 mają charakter pośredni (ryc. 70). Pierwsza jest najbardziej liczna, bo składa się aż ze 124 przypadków. Mimo to jest, do odległości około 100 km, stosunkowo jednorodna. Pierwsza składowa w tej grupie ma charakter sferyczny, dalej przyrost semiwariancji ma charakter wykładniczy. Udział składowej nuggetowej jest tu wysoki (13,7%), a trendu niski (24,0%). Względne proporcje poszczególnych składowych są następujące:  $C_0 : C_1 : C_2 + C_3 : \text{trend} = 1,00 : 1,53 : 3,03 : 1,75$ . Pod względem obrazu przestrzennego (ryc. 71 – styczeń 1980 roku) „duże” strefy wysokich i niskich opadów są bardziej zwarte niż poprzednio, zaś w małej skali widoczna jest drobniejsza mozaika powierzchni.

Czwarta grupa wykazuje duże podobieństwo do trzeciej pod względem niskiego udziału wariancji nuggetowej (8,7%) i wykładniczego charakteru pierwszej składowej (udział 17,3%).

Przy podziale na 7 grup najbardziej zwracają uwagę typy szósty i siódmy (ryc. 69). Ten ostatni powstał z podzielenia grupy trzeciej poprzedniej klasyfikacji i zawiera jedynie 9 przypadków. Charakteryzuje go skrajnie ni-



ski udział wszystkich składowych poza trendem ( $C_0 = 4,2\%$ ,  $C_1 = 11,7\%$ ,  $C_2+C_3 = 10,5\%$ , trend = 73,6%). Przykład tak specyficznego rozkładu MSDO z lutego 1978 roku przedstawiono na rycinie 75. Polska podzielona jest na dwie rozległe strefy, wzdłuż linii biegnącej od Suwałk do Kotliny Kłodzkiej. Na wschód od tej granicy opady były wysokie, na zachód – generalnie niskie. Wyraźny jest też dość szeroki pas przejściowy, z bardzo drobną mozaiką płatów wyższych i niższych opadów.

W typie szóstym specyficzny jest jego przebieg po przekroczeniu zasięgu pierwszej struktury – bardzo stromy i przecinający wykresy średnich modeli typów 1, 2 i 3 (ryc. 69). Wszystkie inne typy, oprócz tego, mają wykresy z grubsza do siebie równoległe, świadcząc, o czym już wspomiano poprzednio, o pewnej ciągłości, stopniowości zmian. Tak odmienny charakter spadku podobieństwa danych MSDO wskazywałby zatem na wyraźnie inny genetycznie jego charakter, a nie tylko zmianę „proporcji składników”. Typ szósty, składający się z 29 przypadków, powstał z podziału grupy 1 poprzedniej klasyfikacji. Proporcje poszczególnych składowych są w nim następujące:  $C_0 = 8,9\%$ ,  $C_1 = 17,8\%$ ,  $C_2+C_3 = 41,9\%$ , trend = 31,4%. Pole MSDO z października 1970 roku – przykład typu szóstego (ryc. 74) – składa się z dużych (100-200 km), względnie zwartych powierzchni wysokich i niskich sum opadów. Wartości minimalne i maksymalne występują generalnie w centrum tych powierzchni, a ku ich granicom postępują zmiany w postaci nieregularnej mozaiki drobnych płatów.

Pierwsza klasa modeli rocznych MSDO jest najmniej liczna (ryc. 76). Stanowi ją bowiem tylko 5 przypadków, które powtarzają się co dwa lata w dekadzie lat sześćdziesiątych (lata 1962, 1964, 1966, 1968, 1970). Na rycinie 53, gdzie zamieszczono wszystkie modele roczne, tworzą one wyraźnie odróżniającą się od pozostałych wiązkę krzywych. Charakteryzuje je najniższy udział semiwariancji nuggetowej (ok. 7,2%), wykładniczy kształt pierwszej składowej „regularnej” (zasięg ok. 10-25 km, średnio 17 km) i dominacja długodystansowych źródeł zmienności przestrzennej MSDO. Semiwariancja po pierwszym załamaniu krzywej rośnie raczej liniowo. Składowa trzecia nie jest w tych modelach w ogóle obecna, a udział drugiej jest bardzo niski. Proporcje udziału poszczególnych struktur ( $C_0$ ,  $C_1$ ,  $C_2+C_3$  i trend) są następujące: 1,0 : 3,1 : 1,8 : 7,9. Dla grupy tej charakterystyczna jest także największa zmienność, szczególnie na odcinku 10-70 km.

Pozostałe dwie klasy składają się z 6 (druga: lata 1957, 1960, 1963, 1965, 1967 i 1975) i 14 przypadków (trzecia: 1956, 1958, 1959, 1961, 1969, 1971, 1972, 1973, 1974, 1976, 1977, 1978, 1979 i 1980). Najważniejszymi cechami modeli zaliczonych do klasy trzeciej są:

- stosunkowo niska wariancja nuggetowa wynosząca 9,3%;
- sferyczny kształt składowej pierwszej o zasięgu raczej krótkim (śr. ok. 16 km);

- wykładniczy charakter przebiegu krzywej semiwariancji po pierwszym jej załamaniu; poziom *plateau* osiągniany jest przy odległości ponad 200 km;
- mimo największej liczebności jest to klasa stosunkowo jednorodna; maksymalna zmienność przebiegu modeli należących do owej klasy ma miejsce przy dużych odległościach – przekraczających 150 km;
- średnie proporcje poszczególnych składowych ( $C_0$ ,  $C_1$ ,  $C_2+C_3$  i trend) są następujące: 1,0 : 3,3 : 4,6 : 1,8.

W klasie trzeciej radykalnie zmieniają się zatem proporcje udziału zmienności długodystansowej i regionalnej oraz spada nieco znaczenie składowej lokalnej.

Identyczne zestawienie wiodących cech klasy drugiej jest następujące:

- najwyższa relatywnie wariancja nuggetowa (13,2%);
- sferyczny kształt składowej pierwszej o największym zasięgu (śr. 21,2 km);
- wykładniczy charakter przebiegu krzywej semiwariancji po pierwszym jej załamaniu; poziom *plateau* osiągniany jest już przy odległości około 160-170 km;
- zróżnicowanie jest w tej klasie relatywnie najmniejsze, szczególnie na odległościach do 100 km,
- średnie proporcje poszczególnych składowych ( $C_0$ ,  $C_1$ ,  $C_2+C_3$  i trend) są następujące: 1,0 : 2,7 : 3,8 : 0,05.

Jest to zatem ta część przypadków rocznych MSDO, w których trendu w ogóle nie ma albo jest znikomo mały, największy zaś jest składnik losowy (nugget).

## 7. Sezonowa zmienność parametrów modeli struktury przestrzennej

Ocenę zmienności sezonowej parametrów modeli struktury przestrzennej miesięcznych MSDO dokonano za pomocą jednoczynnikowej analizy wariancji (ANOVA). Umożliwia ona szybkie przetestowanie hipotezy o różnicach średnich w wielu grupach, a także wskazanie – przy jej potwierdzeniu – które konkretnie grupy się różnią. Najważniejsze wyniki analizy przedstawiono na rycinach 79 i 80 oraz w tabeli 11. Wynika z nich (tab. 11), że bez żadnej wątpliwości ( $p < 0,00001$ ) istnieje sezonowe zróżnicowanie wartości wariancji nuggetowej ( $C_0$ ), oraz wariancji i zasięgu pierwszej składowej ( $C_1$  i  $A_1$ ). W odniesieniu do pozostałych parametrów analiza wariancji nie wykazała istotnego zróżnicowania średnich dla poszczególnych miesięcy.

Zmiany  $C_0$  wykazują najbardziej konsekwentny przebieg (ryc. 79). Od początku roku kalendarzowego do września następuje spadek udziału wariancji nuggetowej (z niewielkim wahaniami *in plus* w lipcu). Gradient spadku jest najsilniejszy pomiędzy marcem a kwietniem. W miesiącach wiosennych

**Tabela 11.** Podsumowanie wyników testowania istotności zmienności sezonowej parametrów modeli struktury przestrzennej znormalizowanych miesięcznych danych MSDO z użyciem jednoczynnikowej analizy wariancji. Pogrubioną czcionką zaznaczono różnice istotne na poziomie  $p < 0,05$

Parametr	SK pomiędzy grupami	df grup	ŚK pomiędzy grupami	SK reszt	df reszt	ŚK reszt	Test F	Poziom $p$
Parameter	SS effect	df effect	MS effect	SS error	df error	MS error	F test	$p$ level
<b>C<sub>0</sub></b>	<b>0,5991</b>	<b>11</b>	<b>0,05447</b>	<b>1,5165</b>	<b>288</b>	<b>0,00527</b>	<b>10,3436</b>	<b>0,0000</b>
<b>C<sub>1</sub></b>	<b>0,8001</b>	<b>11</b>	<b>0,07274</b>	<b>2,9789</b>	<b>288</b>	<b>0,01034</b>	<b>7,0320</b>	<b>0,0000</b>
C <sub>2</sub>	0,2173	11	0,01975	7,7809	288	0,02702	0,7310	0,7084
C <sub>3</sub>	0,9399	11	0,08544	16,2984	288	0,05659	1,5098	0,1270
C <sub>2</sub> +C <sub>3</sub>	0,8467	11	0,07697	16,5390	288	0,05743	1,3403	0,2015
~trend	1,0875	11	0,09886	17,7373	288	0,06159	1,6052	0,0965
<b>A<sub>1</sub></b>	<b>1,6761e<sup>9</sup></b>	<b>11</b>	<b>152368712</b>	<b>9,8438e<sup>9</sup></b>	<b>273</b>	<b>3,60577e<sup>7</sup></b>	<b>4,2257</b>	<b>0,0000</b>
A <sub>2</sub>	1,0655e <sup>10</sup>	11	968668122	1,1511e <sup>11</sup>	170	6,77118e <sup>8</sup>	1,4306	0,1631
A <sub>3</sub>	1,0958e <sup>10</sup>	11	996212587	1,9180e <sup>11</sup>	142	1,35073e <sup>9</sup>	0,7375	0,7010

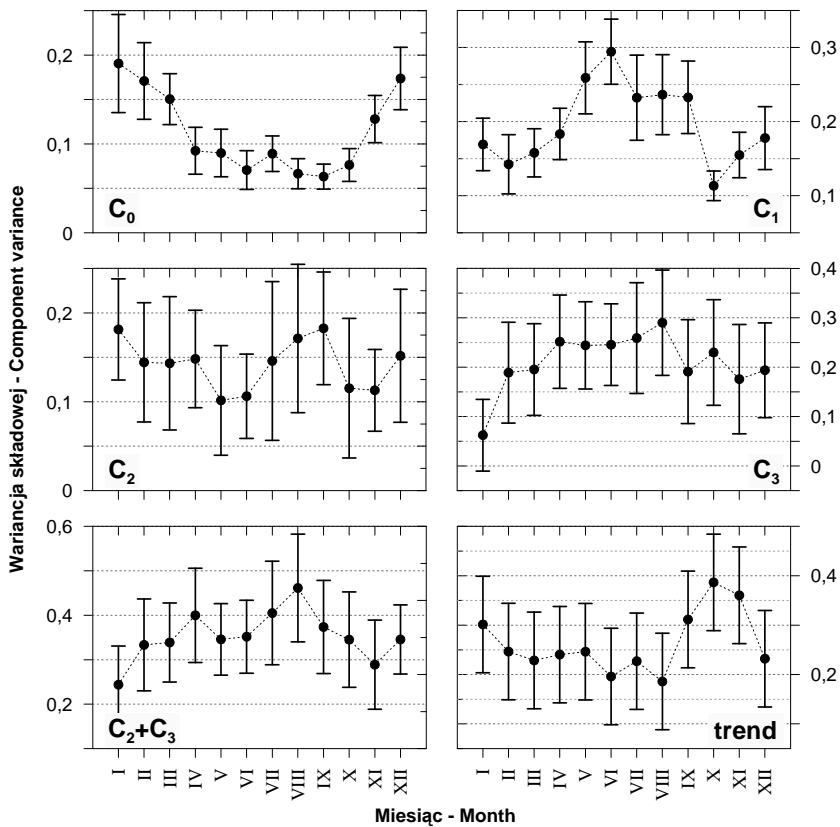
SK – suma kwadratów, df – liczba stopni swobody, ŚK – średnie kwadraty.

i letnich różnice są już niewielkie. Od października następuje szybki wzrost C<sub>0</sub>; w grudniu jego wartość jest zbliżona do styczniowego maksimum. Test *post-hoc* Tukeya wykazał, że bardzo istotne statystycznie różnice wartości wariancji nuggetowej (dla większość porównań  $p < 0,001$ ) istnieją pomiędzy styczniem i lutym a miesiącami od kwietnia do października. Marzec różni się istotnie od czerwca, sierpnia, września i października. Test ten wskazuje także na możliwość wydzielenia na poziomie ufności  $\alpha = 0,05$  dwóch homogenicznych grup miesięcy. Do pierwszej należą grudzień, styczeń, luty i marzec, do drugiej – pozostałe miesiące.

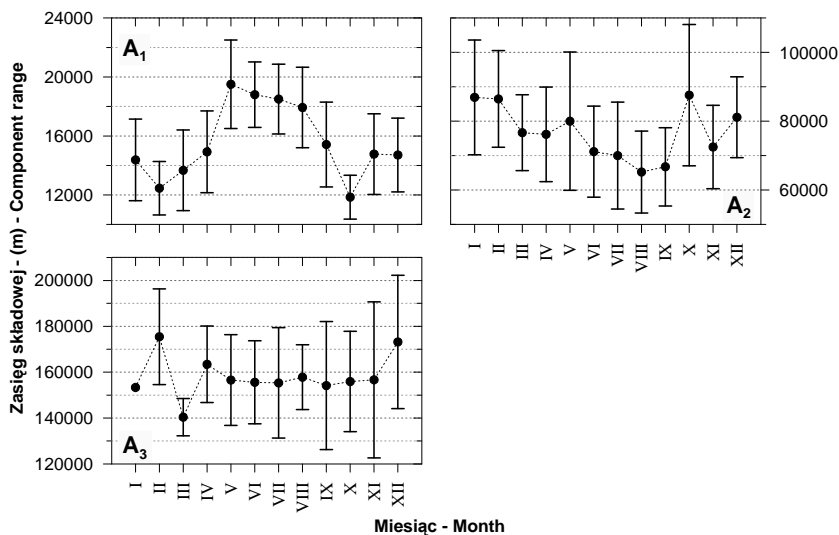
Przebieg sezonowy wariancji pierwszej składowej (C<sub>1</sub>) jest do pewnego stopnia lustrzanym odbiciem zmienności C<sub>0</sub> (ryc. 79). W okresach, kiedy C<sub>0</sub> jest wysokie C<sub>1</sub> jest niskie i odwrotnie. Są jednak dość istotne różnice w szczegółach. Minimum wariancji pierwszej składowej ma miejsce w październiku. Następnie, po niewielkim wzroście, przez kolejne 6 miesięcy utrzymuje się na zbliżonym poziomie, przy czym od lutego zaznacza się trend wzrostowy. Gwałtowny skok wartości C<sub>1</sub> ma miejsce między kwietniem i majem, a maksimum notowane jest w czerwcu. W teście Tukeya w istotny sposób ( $p < 0,05$ ) odróżniają się jedynie skrajne miesiące: maj, czerwiec i październik. Ten pierwszy różni się od lutego i marca oraz października i listopada. Wartości C<sub>1</sub> w czerwcu są odmienne od zanotowanych aż w siedmiu miesiącach: styczeń – kwiecień i październik – grudzień. Paździer-

nik z kolei, oprócz maja i czerwca, wykazuje istotne różnice także względem lipca, sierpnia i września. Żaden z zastosowanych testów *post-hoc* (Fishera, Bonferroniego, Scheffégo, Tukeya, Duncana i Newman-Keulsa) nie wskazał na możliwość wyróżnienia homogenicznych grup miesięcy.

W trakcie interpretacji analizy zmienności sezonowej wariancji nuggetowej oraz pierwszej składowej na uwagę zasługuje również silna wprost proporcjonalna zależność istniejąca pomiędzy średnimi miesięcznymi wartościami  $C_0$  i  $C_1$  a ich odchyleniami standardowymi (ryc. 81). Wykres tej relacji dla wariancji nuggetowej pokazuje względną homogeniczność okresu między kwietniem a październikiem w porównaniu z pozostałą częścią roku. Na analogicznym wykresie wykonanym dla  $C_1$  wyraźnie rysuje się ekstremalne położenie danych z października i czerwca oraz istnienie dwóch grup miesięcy: wiosenno-letnich (V-IX) i jesienno-zimowych (X-IV).



Ryc. 79. Zmienność sezonowa średnich wartości wariancji poszczególnych składowych modeli struktury przestrzennej miesięcznych znormalizowanych danych MSDO. Zaznaczono 95% zakres ufności średniej



**Ryc. 80.** Zmienność sezonowa średnich wartości zasięgu poszczególnych składowych modeli struktury przestrzennej miesięcznych znormalizowanych danych MSDO. Zaznaczono 95% zakres ufności średniej. Zakres ten dla  $A_3$  w styczniu wynosił ponad 168 km. Został on pominięty, ponieważ jego wielkość wpływała na zmniejszenie czytelności wykresu

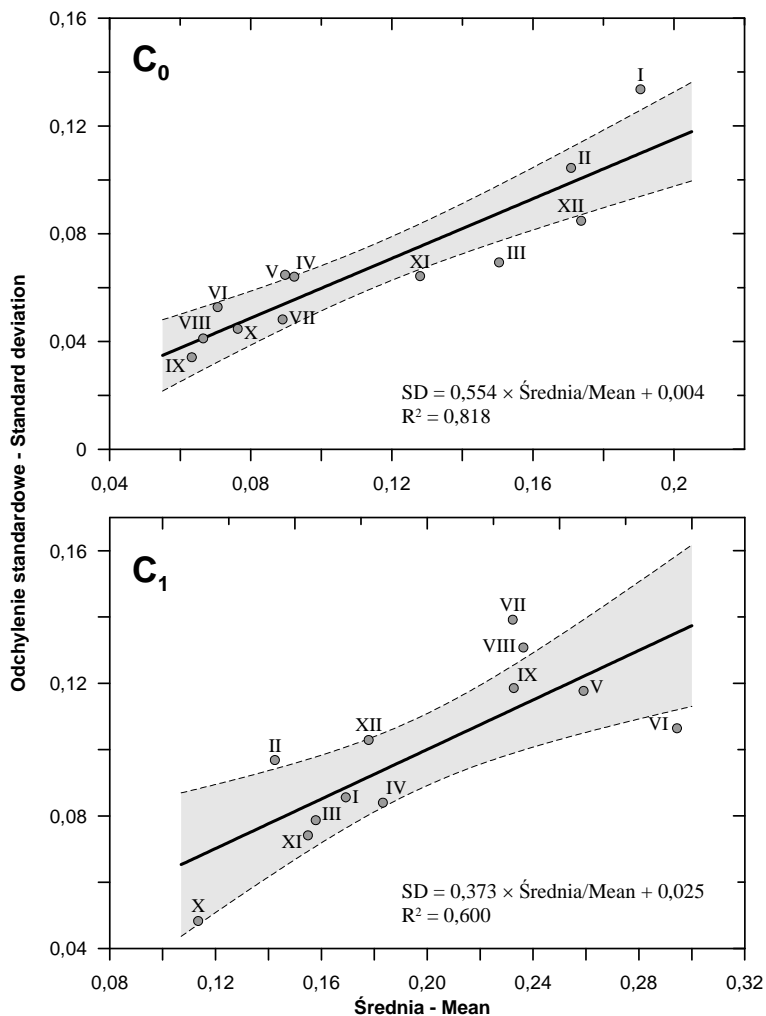
Wariancje cząstkowe pozostałych składowych ( $C_2$ ,  $C_3$ ,  $C_2 + C_3$ , trend), jak wspomniano powyżej, nie wykazują w świetle obliczeń ANOVA istotnej zmienności sezonowej. Nie znaczy to, oczywiście, że takowa w ogóle nie istnieje. Po prostu, zmienność wewnątrzgrupowa i międzygrupowa różni się na tyle nieznacznie, że hipoteza o istotnym zróżnicowaniu musiała zostać odrzucona. Jednakże, konsekwentne zmiany średnich z miesiąca na miesiąc, występujące zwłaszcza w wypadku składowej trzeciej i trendu, mogą sugerować, że pewien cykl sezonowy jednak istnieje (ryc. 79). Wariancja składowej trzeciej ( $C_3$ ) jest maksymalna i względnie stała od kwietnia do sierpnia. Później nieco spada i utrzymuje się na podobnym poziomie do końca roku. Najniższą jej wartość notuje się w styczniu. W lutym i marcu następuje szybki wzrost. W rozkładzie średnich wartości wariancji trendu najbardziej zwraca uwagę wzrost i spadek od sierpnia do grudnia, z maksimum w październiku. Od stycznia do sierpnia wartości są mało zmienne, ale z wyraźną tendencją spadkową. Minimum występuje od czerwca do sierpnia.

Wszystkie omówione wyżej przebiegi sezonowe wariancji poszczególnych składowych modeli struktury przestrzennej miesięcznych danych MSDO dają się dość prosto objaśnić w kontekście przedstawionej hipotezy o genezie związanych z nimi opadów. Zmienność roczna wariancji nuggeto-

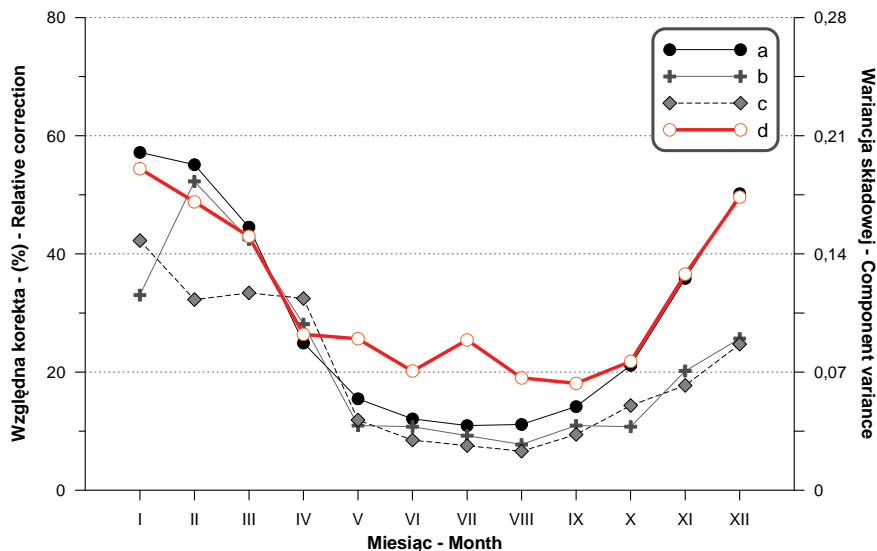
wej ( $C_0$ ) to głównie funkcja błędów pomiarowych – wyższych w miesiącach z opadami śniegu i większą prędkością wiatru (Jakubiak 1984, Lenart 1984, ryc. 79). Ewentualne zróżnicowanie sezonowe krótkodystansowej zmienności opadów (do 2,5 km) ma małe znaczenie bądź w ogóle nie występuje. Taką właśnie interpretację sugeruje zestawienie dokonane na rycinie 82. Porównano na niej określoną w niniejszym opracowaniu zmienność średnich wartości wariancji nuggetowej ( $d$ ) w trakcie roku ze współczynnikami korekcyjnymi opadów dla obszaru leżącego w zasięgu eksperymentu BALTEX (*Baltic Sea Experiment* obejmujący obszar od 9°E do 35°E i 50°N do 71°N, Raschke i in. 1998). Współczynniki korekcyjne związane są przede wszystkim z aerodynamicznymi właściwościami deszczomierzy i uzależnione od prędkości wiatru, rozkładu wielkości kropeł deszczu i struktury krystalicznej płatków śniegu. Znacznie mniejsze znaczenie mają straty na parowanie i zwilżanie. Krzywa (a) na rycinie 82 przedstawia przebieg sezonowy współczynnika korekcyjnego opracowanego na podstawie wieloletnich serii średnich wartości miesięcznych przez Legatesa i Willmota (za Ungersböck i in. 2001 oraz Rudolf i Rubel 2005) – tak zwana korekta klimatologiczna. Stosowana jest ona rutynowo do przeliczania wyników pomiarów opadów atmosferycznych gromadzonych w międzynarodowych bazach danych. Krzywe (b) i (c) – uważane za bardziej realistyczne – powstały w wyniku korekty każdego pojedynczego opadu (sumy dobowej; tzw. *on-event calculated Correction Factor*, Rudolf i Rubel 2005), ale jedynie dla dwóch lat: 1996 i 1997. Współczynniki z „pojedynczych opadów” (krzywe b i c) są generalnie konsekwentnie niższe od współczynnika „klimatycznego” (krzywa a), szczególnie w miesiącach zimowych. Może to jednak być związane ze specyfiką klimatyczną owych dwóch lat. Na wykresie 82 zakres i wzajemne położenie skal rzędnych dla wartości poszczególnych CF i  $C_0$  dobrano tak, aby uzyskać jak najlepszą możliwość porównania ich sezonowego przebiegu. Z porównania owego wynika duża zbieżność układu krzywych od października do kwietnia, szczególnie między  $C_0$  a klimatycznym CF, oraz relatywnie wyższe wartości wariancji nuggetowej w okresie wiosenno-letnim. Mniejsza względna amplituda sezonowej zmienności  $C_0$  wskazywałaby, że zmienność efektywności deszczomierzy jest ważnym, ale nie jedynym czynnikiem wpływającym na wartość błędu losowego. Ma on decydujące znaczenie w miesiącach jesienno-zimowych, a w pozostałej części roku jego rola maleje na rzecz wpływu nieciągłości i zmienności krótkodystansowej opadów. Taką interpretację sugerują wyniki analizy struktury przestrzennej pojedynczych rocznych MSDO, które zostaną zaprezentowane w oddzielnej publikacji.

Cykl roczny wariancji składowej pierwszej ( $C_1$ ) odzwierciedlałby zmiennej udział opadów konwekcyjnych – maksymalny w miesiącach o największych kontrastach termicznych: maju i czerwcu. Interpretacja zmian warian-

cji składowej trzeciej ( $C_3$ ) jest trudniejsza, bo jeśli postawiona hipoteza jest słuszna, odzwierciedlać powinna zarówno roczne zróżnicowanie aktywności cyklonalnej, jak i kierunku oraz rozmiaru anizotropii frontalnych stref opadowych. Występowanie zaś maksymalnych wartości wariancji składowej długodystansowej w miesiącach jesiennych jest prawdopodobnie efektem wyższej w tym okresie frekwencji występowania bardzo rozległych pól opadów.



**Ryc. 81.** Relacje między średnimi w poszczególnych miesiącach roku wartościami  $C_0$  i  $C_1$  modeli struktury przestrzennej znormalizowanych danych MSDO a ich odchyleniami standardowymi. Liniami przerywanymi zaznaczono 95% przedział ufności regresji



**Ryc. 82.** Sezonowa zmienność udziału wariancji nuggetowej (d) zestawiona z klimatologicznym współczynnikiem korekcyjnym (CF) opadów (a) oraz uzyskanymi w trakcie eksperymentu BALTEX współczynnikami korekcyjnymi dla lat 1996 (b) i 1997 (c). Dane dla krzywych a, b i c uzyskano z publikacji Ungersböck i in. (2001) oraz Rudolf i Rubel (2005). Dokładne objaśnienia w tekście ☐

Zmiany średnich wartości zasięgów poszczególnych składowych są również raczej regularne, ale ich interpretacja jest trudniejsza ze względu na bardzo dużą zmienność (ryc. 80). Zasięg pierwszej składowej zmienia się w ciągu roku bardzo podobnie jak jej wariancja (ryc. 79). Maksimum (19,5 km) osiąga w maju. W kolejnych miesiącach wiosenno-letnich następuje konsekwentnie lekki spadek zasięgu – do około 18 km w sierpniu. Od sierpnia do października ma miejsce gwałtowne zmniejszenie się średniej miesięcznej wartości  $A_1$  do rocznego minimum wynoszącego 11,8 km. W listopadzie, grudniu, styczniu, marcu i kwietniu średnie zasięgi  $A_1$  wahają się w granicach 13,7-14,9 km. Drugie minimum roczne występuje w lutym i wynosi 12,5 km. Statystycznie istotne różnice dotyczą głównie ekstremalnych wartości. Październik jest najbardziej specyficzny, bo odróżnia się najsilniej od największej liczby miesięcy. Jest to okres od maja do sierpnia. Maj różni się poza tym od lutego i marca, a czerwiec i lipiec tylko od lutego.

Zasięg drugiej składowej ( $A_2$ , ryc. 80) maleje od początku roku kalendarzowego do sierpnia, z niewielkim wahaniami *in plus* w maju. Spadek ten zawiera się w przedziale 87-65 km. Od września do końca roku zasięg  $A_2$  rośnie. Wyraźną anomalię w tym trendzie stanowi październik, z maksymalną



średnią wartością  $A_2$  w roku wynoszącą 87,5 km. Podobną generalną tendencję wykazuje sezonowa zmienność zasięgu trzeciej składowej ( $A_3$ , ryc. 80). W tym przypadku najbardziej zwracają uwagę anomalie ujemne w styczniu i marcu.

Należy podkreślić, że ANOVA jest metodą stosunkowo prostą i nie ma żadnych specyficznych odmian służących do badania serii czasowych, jak to miało miejsce w tym przypadku. W szczególności, obecność trendu lub cykliczności wieloletniej może spowodować, że istniejąca zmienność sezonowa nie zostanie zidentyfikowana. Sprawdzono zatem, czy w analizowanych danych wymienione wyżej składniki zmienności regularnej występują.

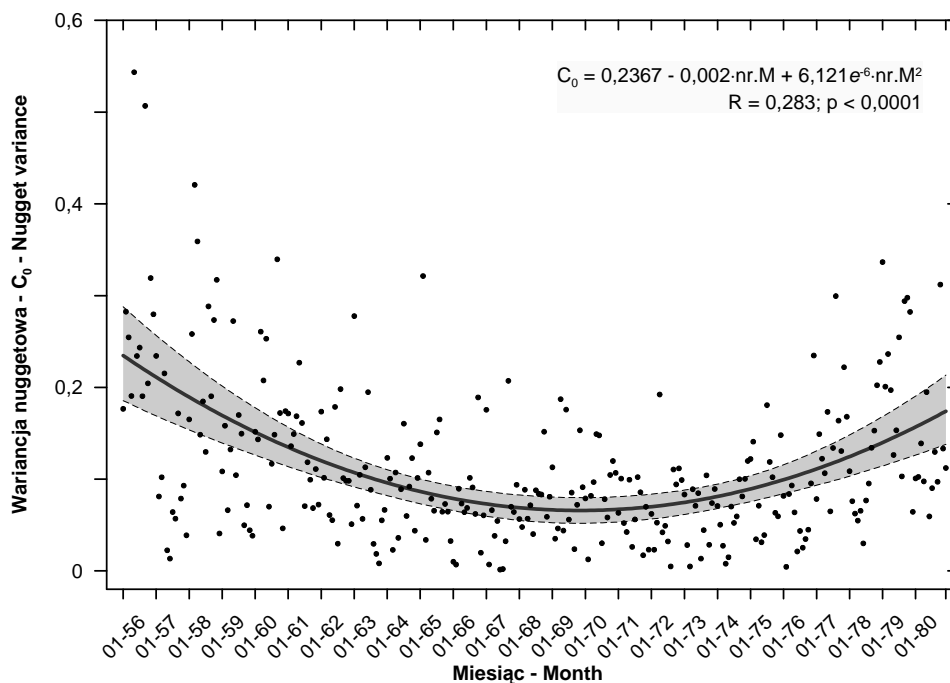
## 8. Wieloletnia zmienność parametrów modeli struktury przestrzennej

Do oceny możliwości występowania długookresowego liniowego trendu parametrów modeli struktury przestrzennej miesięcznych MSDO z wielolecia 1956-1980 użyto, jak poprzednio, nieparametrycznego testu Manna-Kendalla (MK), w jego modyfikacji uwzględniającej sezonowość (Hirsch i in. 1982, Libiseller i Grimvall 2002).

Sumaryczny test MK (ang. *combined*, tab. 12) wykazał, że z dużym prawdopodobieństwem ( $p = 0,001$ ) w analizowanym wieloleciu miał miejsce malejący trend udziału wariancji nuggetowej. Nieco mniej wiarygodny był liniowy spadek wariancji drugiej składowej ( $p = 0,026$ ). Uwagę należy również zwrócić na potencjalny wzrost jej zasięgu ( $p = 0,07$ ). We wszystkich pozostałych przypadkach hipotezę należałoby odrzucić. Testy cząstkowe – dla poszczególnych miesięcy – pokazały, że spadek  $C_0$  był znaczący przede wszystkim między grudniem a marcem, a najbardziej w styczniu ( $p = 0,001$ ) i lutym ( $p = 0,003$ ). Prawidłowość tę wyjaśnić można przede wszystkim spadkiem lokalnej zmienności opadów (por. rozdz. V) oraz, być może, spadkiem udziału opadów stałych zimą. Trudniej natomiast znaleźć uzasadnienie dla wysoko istotnej ( $p = 0,007$ ) tendencji zmniejszania się wariancji nuggetowej w sierpniu. W odniesieniu do pozostałych parametrów notowano tylko pojedyncze przypadki miesięcznych testów o  $p$  mniejszym od 0,05. Większość z nich należałoby traktować jako zupełnie losowe. Jako wyjątek można jednak potraktować spadek udziału wariancji składowej drugiej w czerwcu, ponieważ prawdopodobieństwo, że jest to układ przypadkowy jest mniejsze niż 0,004. Test Manna-Kendalla, zastosowany do parametrów rocznych modeli struktury przestrzennej MSDO, nie wykazał w żadnym kontekście istnienia istotnej liniowej tendencji.

**Tabela 12.** Ogólny wynik testu Manna-Kendalla na obecność liniowego trendu parametrów struktury przestrzennej miesięcznych kodowanych MSDO

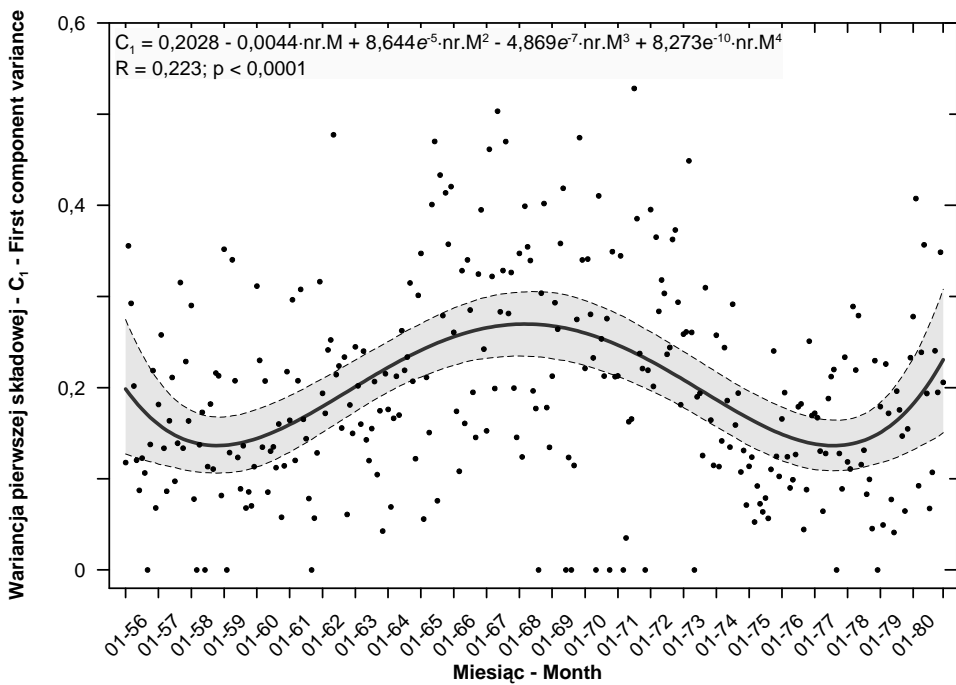
Parametr Parameter	n	Stat. testu Test stat.	Odch. Std. Std. Dev.	Stat. MK MK stat.	Poziom $p$ $p$ -value
C <sub>0</sub>	300	-902	274,350	-3,2878	0,0010
C <sub>1</sub>	300	176	144,448	1,2184	0,2231
C <sub>2</sub>	300	-382	171,377	-2,2290	0,0258
C <sub>3</sub>	300	128	125,687	1,0184	0,3085
trend	300	196	130,0513	1,5071	0,1318
C <sub>2</sub> +C <sub>3</sub>	300	-148	150,683	-0,9822	0,3260
A <sub>1</sub>	285	-62	217,160	-0,2855	0,7753
A <sub>2</sub>	182	154	85,010	1,8116	0,0701
A <sub>3</sub>	154	-92	70,062	-1,3131	0,1891



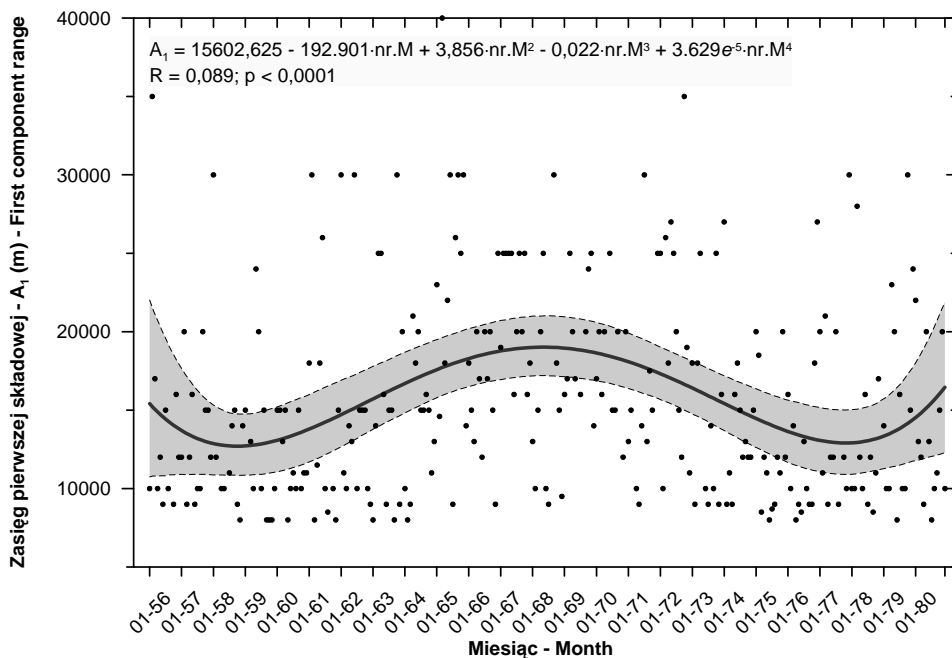
**Ryc. 83.** Trend czasowy udziału wariancji nuggetowej modeli struktury przestrzennej znormalizowanych danych miesięcznych MSDO dla wielolecia 1956-1980. Skrót „nr M” we wzorze oznacza numer kolejny miesiąca w wieloleciu

Oprócz potencjalnej liniowej tendencji sprawdzono możliwość istnienia trendu krzywoliniowego i/lub cykliczności wieloletniej. Okazało się, że można ich obecność zidentyfikować w przypadku właśnie tych trzech pa-

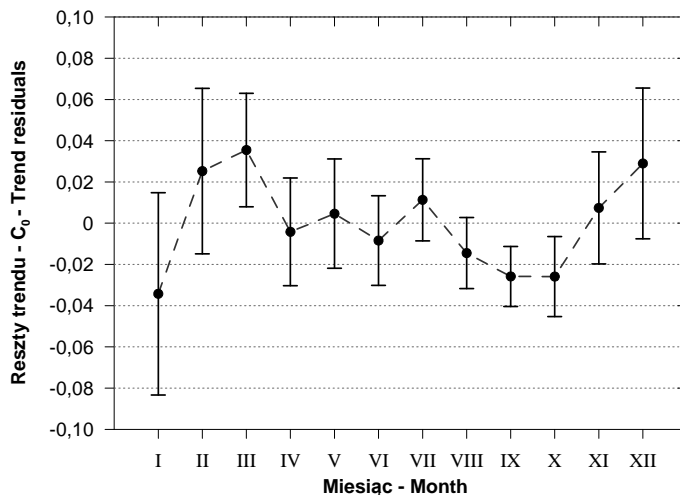
rametrów, dla których za pomocą analizy wariancji stwierdzono istotną zmienność sezonową, to jest:  $C_0$ ,  $C_1$  i  $A_1$  (ryc. 83-85). Współczynniki determinacji są bardzo niskie, bo wynoszą odpowiednio 0,283, 0,223 i 0,089, ale przy dużej liczbie analizowanych zbiorów są wysoko istotne statystycznie - w każdym przypadku  $p$  jest mniejsze od 0,0001. Trendy modelowano za pomocą funkcji wielomianowych, ale mają one wyraźnie charakter cykliczny. W przypadku  $C_0$  okres zmian jest zbyt duży, aby na podstawie posiadanej 25-letniej serii można go było zidentyfikować. Najwyższe wartości funkcja osiąga na początku i końcu okresu pomiarowego, zaś minimum w grudniu 1969 roku. Różnice między ekstremami są znaczące, bo przekraczają 250%. Trendy zmian  $C_1$  i  $A_1$  są praktycznie identyczne. W obrębie analizowanego wielolecia mieści się jeden pełen, prawie idealnie sinusoidalny, cykl zmian. Okres „fali” dla wariancji składowej pierwszej wynosi prawie 19 lat pomiędzy listopadem 1958 i wrześniem 1977 (227 m-ce), zaś jej pik wystąpił w kwietniu 1968 roku. Amplituda jest znacznie mniejsza niż poprzednio, bo wynosi około 98%. Zbieżność przebiegu funkcji obliczonej dla zasięgu tej



Ryc. 84. Trend czasowy wartości wariancji pierwszej składowej modeli struktury przestrzennej znormalizowanych danych miesięcznych MSDO dla wielolecia 1956-1980. Skrót „nr M” we wzorze oznacza numer kolejny miesiąca w wieloleciu



Ryc. 85. Trend czasowy wartości zasięgu pierwszej składowej modelu struktury przestrzennej znormalizowanych danych miesięcznych MSDO dla wielolecia 1956-1980. Skrót „nr M” we wzorze oznacza numer kolejny miesiąca w wieloleciu



Ryc. 86. Zmienność sezonowa średnich wartości reszt z trendu udziału wariancji nuggetowej modelu struktury przestrzennej miesięcznych znormalizowanych danych MSDO. Zaznaczono 95% zakres ufności średniej

samej składowej jest prawie idealna: okres wynosi 230 miesięcy (XI-58 do XII-77), a maksimum przypada na czerwiec 1968 roku (amplituda około 47%). Oczywiście, tak wysoka zgodność jest w dużym stopniu efektem istnienia korelacji pomiędzy przebiegiem  $C_1$  i  $A_1$  ( $r = 0,521$ ,  $p < 0,00001$ ,  $N = 285$ ).

Fakt istnienia i charakter wieloletniej regularnej zmienności parametrów struktury przestrzennej miesięcznych MSDO jest niewątpliwie bardzo interesujący i wymaga szczególnej uwagi. Mimo ich wysokiej istotności statystycznej nie można całkowicie wykluczyć możliwości, że jest to artefakt wynikający raczej z wystąpienia losowego zgrupowania kilkunastu anomalnych opadów, niż z rzeczywistej wieloletniej tendencji. Jest to jednakże bardzo mało prawdopodobne, tak jak i możliwość, że jakiś wpływ mogły mieć na uzyskany wynik zmiany w sieci pomiarowej, czy sposób wykonywania obliczeń (bo nie są to „surowe” dane pomiarowe, tylko efekt skomplikowanego ich przetwarzania numerycznego). Bliższe naświetlenie tego problemu wymagałoby pogłębionej analizy zmienności całego systemu klimatycznego, a zwłaszcza stosunków termiczno-wilgotnościowych na terenie środkowej Europy w drugiej połowie XX wieku. W odniesieniu do zmienności wariancji nuggetowej ( $C_0$ ) na przykład można by spróbować dokonać porównania ze wskaźnikami ilości i charakteru opadów zimowych, a także reżimu anemologicznego.

Po obliczeniu funkcji trendów  $C_0$ ,  $C_1$  i  $A_1$  sprawdzono, czy odfiltrowanie wieloletniej zmienności regularnej wprowadzi istotne zmiany do stwierdzonej poprzednio sezonowości tych parametrów struktury przestrzennej miesięcznych MSDO. Wykonano zatem analizę wariancji ponownie, ale tym razem na resztach z trendu. Różnice były znaczące. Istotne statystycznie zróżnicowanie stwierdzono tylko w wypadku wariancji nuggetowej ( $C_0$ ), a przebiegi roczne średnich wartości były nieco zmodyfikowane (ryc. 86). Najważniejsza różnica to ekstremalnie ujemne wartości reszt trendu dla stycznia, podczas gdy w grudniu, lutym i marcu są one ekstremalnie dodatnie. Miesiące, które stanowiły poprzednio homogeniczną grupę zostały rozdzielone. Styczeń charakteryzuje się ponadto największą zmiennością wartości reszt  $C_0$ . Przebieg średnich wartości reszt z trendu dla pozostałych miesięcy jest podobny jak poprzednio.

## 9. Zmienność bezwzględnych wartości wariancji składowych przestrzennych

Do charakterystyki struktury przestrzennej MSDO w niniejszym rozdziale używano do tej pory wyłącznie miar względnych<sup>43</sup> będących konsekwencją procedury normalizacji wszystkich zbiorów danych. Zalety tego

---

<sup>43</sup> Stwierdzenie to dotyczy, oczywiście, wariancji poszczególnych składowych. Ich zasięgi były wyrażane cały czas w jednostkach bezwzględnych – metrach.

podejścia były już omówione. Byłoby jednakże celowe i pożyteczne zwrócić uwagę także na zmienność w ujęciu sezonowym parametrów modeli struktury przestrzennej, a konkretnie wariacji poszczególnych składowych, wyrażonej w skali danych źródłowych – czyli milimetrach sumy opadów. Można się było spodziewać, że duże zróżnicowanie bezwzględnych wysokości MSDO w cyklu rocznym mogłoby istotnie wpływać na te relacje. Na przykład, względne udziały  $C_0$  są najwyższe w miesiącach jesienno-zimowych, a najniższe w letnich (ryc. 79). Różnice te przekraczają 100%. Ponieważ jednak amplituda średnich miesięcznych MSDO między tymi okresami również przekracza znacznie 100% (ryc. 39) można się było spodziewać, że w wartościach bezwzględnych zmienność sezonowa wariacji nuggetowej będzie niewielka, albo wręcz odwrotna. W tych przypadkach z kolei, gdzie udziały wariacji składowych nie wykazywały regularnego i istotnego statystycznie przebiegu rocznego ( $C_2$ ,  $C_3$ ,  $C_2+C_3$ , trend) ich zróżnicowanie w skali bezwzględnej może być znaczące.

Najprostsze, zdałoby się, rozwiązanie tego problemu nie mogło dać zadowalających rezultatów. Polegałoby ono na powrocie do oryginalnej skali pomiarowej poprzez wymnożenie względnych proporcji poszczególnych składowych przez wariację każdego z analizowanych zbiorów MSDO. Takie podejście byłoby skuteczne, jeśli by zbiory te miały charakter stacjonarny przestrzennie, to znaczy w obrębie zasięgu obliczeń uzyskiwano by ustaloną wariację progową identyczną lub bardzo zbliżoną do wariacji próby. Takie sytuacje w analizowanym zbiorach danych były jednak bardzo rzadkie. Zdecydowano się zatem na ponowne opracowanie modeli struktury przestrzennej, ale tym razem bezpośrednio z surowych danych pomiarowych. Tym razem jednak nakład pracy był znacząco zredukowany. Jako wzór użyto bowiem już istniejące modele danych znormalizowanych: liczbę i typ składowych oraz ich zasięgi. Zmiany wprowadzone do tych elementów modeli były bardzo niewielkie. Optymalizacji podlegały głównie wartości wariacji poszczególnych składowych. Słuszność takiej, choć bardziej pracochłonnej, metodyki potwierdza wykres relacji wartości wariacji nuggetowej ( $C_0$ ) określonej obiema opisanymi wyżej metodami (ryc. 87). Jakkolwiek istnienie między nimi stosunkowo silna, wprost proporcjonalna zależność ( $R^2 = 0,61$ ), to użycie  $C_0$  wyliczonego jako proporcja wariacji próby systematycznie zaniża rzeczywistą jej wartość (współczynnik kierunkowy krzywej równy 0,862).

Statystyki wariacji składowych modeli struktury przestrzennej miesięcznych i rocznych MSDO wyrażone w milimetrach sumy opadu zamieszczono w tabeli 13. Ryciny 88 i 89 przedstawiają również te dane, ale w wariacie, kiedy brak składowej traktowany jest jako jej zerowy udział.

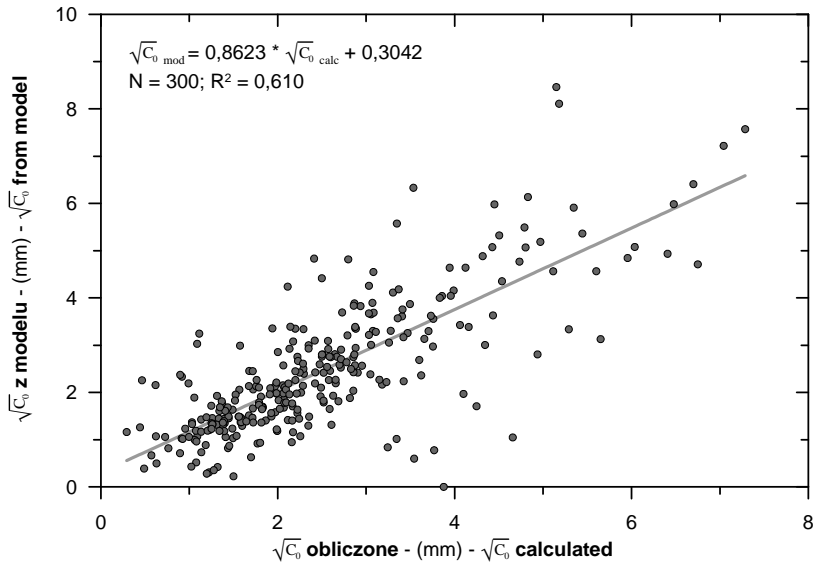
Określenie przeciętnych bezwzględnych wielkości  $C_0$  miesięcznych MSDO pokazuje, jakie różnice wyników pomiarów między sąsiadującymi sta-

nowiskami można uznać za bez wątplenia istotne. W miesiącach jesienno-zimowych (X-III) musi to być co najmniej 2,0 mm, latem (VI-VIII) – co najmniej 4,5 mm. Oznacza to, że średnie różnice MSDO w sąsiadujących stacjach są zimą niewiele większe od sumy wszystkich źródeł niepewności (por. ryc. 43). Porównując roczne MSDO jako „bezpieczny” poziom zróżnicowania, można uznać 8 mm. Podane wyżej wartości powinny się przyjmować jako minimalne cięcia izolinii przy tworzeniu map izarytmicznych dla wielolecia.

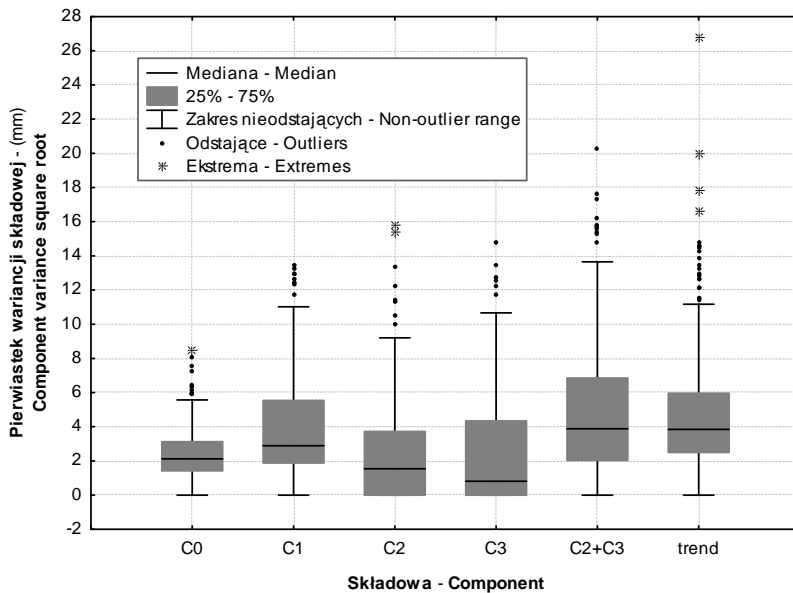
Do oceny istotności zróżnicowania sezonowego parametrów modeli struktury przestrzennej miesięcznych MSDO, opracowanych z oryginalnych danych pomiarowych użyto, jak poprzednio, analizy wariancji (tab. 14, ryc. 90). Każdą analizowaną cechę ( $C_0$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_2+C_3$ , trend) przedstawiono dla uproszczenia w oryginalnej skali pomiarowej – milimetrach sumy opadu. Wymagało to obliczenia pierwiastka kwadratowego z wartości wariancji progowej każdej składowej wszystkich modeli miesięcznych MSDO. W każdym przypadku uzyskano potwierdzenie hipotezy sezonowości (tab. 14);

**Tabela 13.** Statystyki opisowe bezwzględnych wartości pierwiastka kwadratowego z wariancji (w mm sumy opadu) poszczególnych składowych modeli struktury przestrzennej MSDO

Parametr	n	Średnia	-95% przedz. ufności śred.	+95% przedz. ufności śred.	Mediana	Minimum	Maksimum	Dolny kwartył	Górny kwartył	Odch. stand.	Skośność	Kurtoza
Parameter	n	Mean	Confidenc -95%	Confidenc +95%	Median	Minimum	Maximum	Lower Quartile	Upper Quartile	Std. Dev.	Skewness	Kurtosis
Miesięczne - Monthly												
$C_0$	299	2,50	2,34	2,67	2,18	0,22	8,46	1,47	3,29	1,427	1,280	2,013
$C_1$	290	3,99	3,67	4,31	2,91	0,62	13,51	1,89	5,56	2,750	1,342	1,399
$C_2$	177	3,94	3,53	4,35	3,24	0,26	15,76	1,89	5,37	2,749	1,639	3,511
$C_3$	153	4,65	4,20	5,09	3,84	0,29	14,77	2,60	6,37	2,786	1,138	1,418
$C_2+C_3$	267	5,11	4,69	5,54	4,27	0,26	17,64	2,44	6,75	3,502	1,360	1,750
Trend	177	6,80	6,09	7,50	5,52	0,53	29,52	3,51	8,79	4,765	1,601	3,511
Roczne - Yearly												
$C_0$	25	5,42	4,47	6,37	4,87	1,03	10,39	4,18	7,02	2,298	0,388	-0,162
$C_1$	25	9,99	9,13	10,84	10,11	6,63	14,05	8,42	11,56	2,078	0,326	-0,687
$C_2$	15	7,98	6,31	9,65	7,59	4,08	14,80	5,82	10,60	3,021	0,763	0,168
$C_3$	18	7,58	5,76	9,40	8,27	1,80	15,37	5,28	9,58	3,667	0,165	-0,296
$C_2+C_3$	24	10,67	8,56	12,79	10,10	1,80	20,39	7,92	14,20	5,010	0,177	-0,310
Trend	7	12,91	5,88	19,93	11,63	4,30	26,35	5,76	18,85	7,596	0,869	0,499



Ryc. 87. Relacja między bezwzględnymi wartościami wariancji nuggetowej zbiorów miesięcznych MSDO obliczonymi z modeli danych znormalizowanych a pochodzącymi z oryginalnych modeli „surowych” danych pomiarowych. Dokładne objaśnienia w tekście



Ryc. 88. Zmienność bezwzględnych wielkości wariancji poszczególnych składowych modeli struktury przestrzennej miesięcznych MSDO. Brak składowej traktowany był jako jej wielkość zerowa

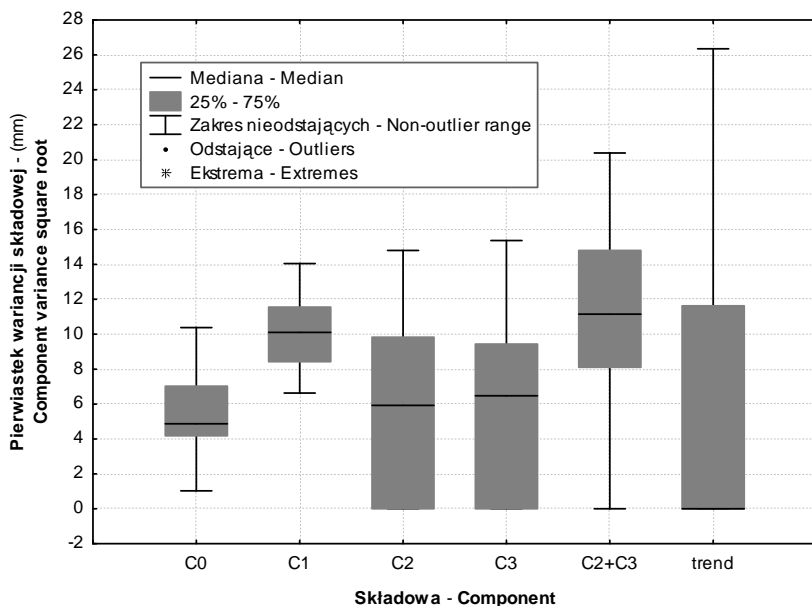


oprócz wariancji składowej drugiej ( $C_2$ ) wszystkie  $p$  były mniejsze od 0,000001. Przebiegi średnich miesięcznych wariancji każdej ze składowych (ryc. 90) są zbliżone do sezonowego cyklu średnich wysokości MSDO (ryc. 39), oraz średnich różnic MSDO w najbliższych stanowiskach (ryc. 43). Takiego

**Tabela 14.** Podsumowanie wyników testowania istotności zmienności sezonowej parametrów modeli struktury przestrzennej miesięcznych danych MSDO z użyciem jednozynnikowej analizy wariancji

Parametr	SK pomiędzy grupami	df grup	ŚK pomiędzy grupami	SK reszt	df reszt	ŚK reszt	test F	Poziom $p$
Parameter	SS effect	df effect	MS effect	SS error	df error	MS error	F test	$p$ level
$C_0$	208,924	11	18,993	393,984	288	1,3680	13,8839	0,000000
$C_1$	1635,649	11	148,695	935,403	288	3,2479	45,7816	0,000000
$C_2$	374,040	11	34,004	2189,733	288	7,6032	4,4723	0,000003
$C_3$	552,986	11	50,272	2496,827	288	8,6695	5,7986	0,000000
$C_2+C_3$	1653,303	11	150,300	2815,994	288	9,7778	15,3717	0,000000
~trend	1067,036	11	97,003	2731,745	288	9,4852	10,2268	0,000000

SK – suma kwadratów, df – liczba stopni swobody, ŚK – średnie kwadraty.



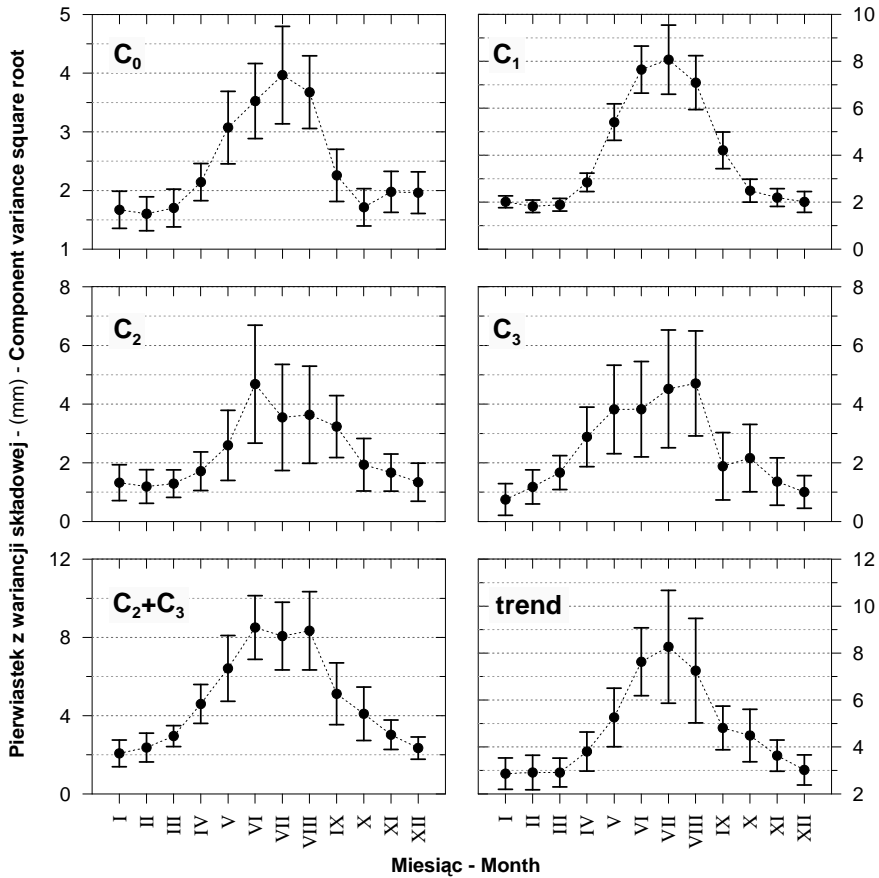
**Ryc. 89.** Zmienność bezwzględnych wielkości wariancji poszczególnych składowych modeli struktury przestrzennej rocznych MSDO. Brak składowej traktowany był jako wielkość zerowa

wyniku można było oczywiście oczekiwać, biorąc pod uwagę silne statystyczne relacje między średnimi i odchyleniami standardowymi (ryc. 42). Współczynniki determinacji relacji między średnimi MSDO w poszczególnych miesiącach a wariancjami  $C_0$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_2+C_3$  i trendu są zatem bardzo wysokie i wynoszą odpowiednio: 0,940, 0,966, 0,857, 0,858, 0,953 i 0,988. Rozrzut reszt z owej regresji ma w odniesieniu do wariancji nuggetowej oraz wariancji składowej drugiej i trendu bez wątplenia charakter losowy. Pewne regularności widoczne są w rozkładzie reszt dla składowej pierwszej i trzeciej. „Zmierzone” wartości  $C_1$  są w pierwszej połowie roku generalnie wyższe od wyliczonych z regresji względem średnich MSDO, a w drugiej – niższe. Różnice wszak nie są duże, w większości przypadków mieszczą się w przedziale  $\pm 0,5$  mm. Największy „nadmiar” ma miejsce w czerwcu, największy „deficyt” w październiku. W odniesieniu do wariancji składowej trzeciej uwagę zwracają wysokie wartości reszt między marcem a majem i niskie – między wrześniem a styczniem. Fakt, że suma średnich miesięcznych wariancji składowej drugiej i trzeciej ( $C_2+C_3$ ) wykazuje znacznie wyższy współczynnik determinacji ze średnimi sumami MSDO, niż każda z nich z osobna, jest jeszcze jedną przesłanką potwierdzającą przypuszczenie, że są one efektem tego samego, anizotropowego, zjawiska.

Największą amplitudę i jednorodność zmienności sezonowej wykazuje wariancja składowej pierwszej ( $C_1$ ) – wiązana z lokalnymi komórkami konwekcyjnymi. Złożyły się na to duże i statystycznie istotne zmiany jej względnego udziału (tab. 11, ryc. 79) oraz zróżnicowanie w ciągu roku bezwzględnych wysokości MSDO. Wielkość i zróżnicowanie  $C_1$  w miesiącach zimowych (I-III) jest ponad cztery razy mniejsza niż latem (VI-VIII). Jesienią rozmiary  $C_1$  są zbliżone do zimowych (~ 2 mm), lecz zmienność jest wyższa o około 100%. W trakcie roku występują dwa względnie homogeniczne sezony wartości tej cechy (okresy X-III i VI-VIII), rozdzielone krótkimi i niejednorodnymi okresami przejściowymi (IV-V i IX).

Najbardziej zwraca uwagę na rycinie 90 odmienność przebiegów składowych drugiej i trzeciej, zarówno każdej z osobna, jak i ich sumy ( $C_2$ ,  $C_3$ ,  $C_2+C_3$ ). We wszystkich pozostałych przypadkach ( $C_0$ ,  $C_1$ , trend), tak jak w przebiegu miesięcznych średnich i odchyłeń standardowych MSDO, maksimum występuje w lipcu. Natomiast, krzywa rocznego przebiegu średnich  $C_2$  osiąga maksimum w czerwcu, a  $C_3$  – w sierpniu. Rozrzut wartości w miesiącach letnich jest jednak na tyle duży, że różnice między nimi nie są istotne statystycznie. Z drugiej strony, przebieg średnich w kolejnych miesiącach jest na tyle konsekwentny, że trudno to uznać za przypadek.

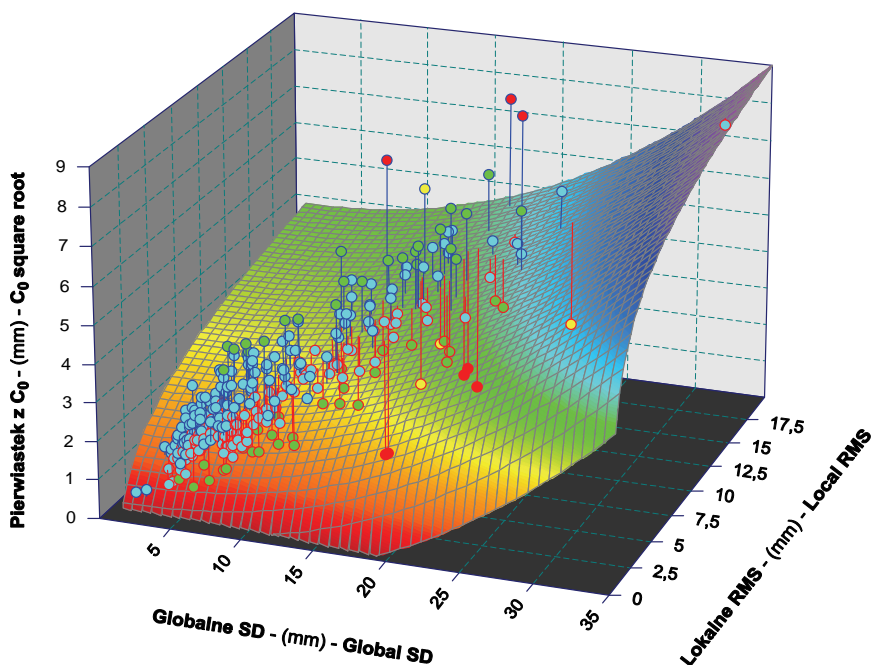
W podrozdziale 7 analizując zmienność sezonową udziału wariancji nuggetowej wykazano, że jest ona bardzo zbliżona do przebiegu współczynników korygujących (ryc. 82) błędy standardowych pomiarów opadów.




Ryc. 90. Zmienność sezonowa średnich wartości pierwiastka wariancji poszczególnych składowych modeli struktury przestrzennej miesięcznych „surowych” danych MSDO. Zaznaczono 95% zakres ufności średniej

Ich źródłem jest przede wszystkim turbulencja spowodowana wiatrem, występowanie opadów w postaci stałej, ale także straty związane z parowaniem i zwilżaniem. Na wielkość wariancji nuggetowej wpływa jednakże również zmienność krótkodystansowa opadów, na odległościach mniejszych niż najmniejszy odstęp próbkowania. W związku z tym można oczekiwać istnienia istotnych relacji pomiędzy wartościami  $C_0$  a statystykami zmienności MSDO, szczególnie lokalnymi. Hipoteza ta została sprawdzona przy zastosowaniu regresji wielokrotnej, z krokową eliminacją zmiennych. Można było dzięki temu wyeliminować zależności pozorne wynikające ze wzajemnych korelacji w zbiorze zmiennych niezależnych. Do zbioru tego włączono te statystyki globalne (średnia i odchylenie standardowe) i lokalne

(średnia, odchylenie standardowe, 3 kwartył, pierwiastek średniego kwadratu i medianę absolutnych odchyłań) MSDO, z którymi wariancja nuggetowa wykazywała korelację liniową większą niż 0,65. Uzyskany model (ryc. 91) z dostosowanym współczynnikiem determinacji równym 0,533 zawierał dwie zmienne: globalne odchylenie standardowe i pierwiastek średniego kwadratu różnic opadów w stosunku do najbliższego sąsiada, których korelacje semicząstkowe równały się odpowiednio 0,168 i 0,128<sup>44</sup>. Błąd standardowy modelu wynosił 0,969 mm. Tak więc zgodnie z przypuszczeniami, wielkość  $C_0$  jest uzależniona także od zmienności pola MSDO. Ponieważ jest to miara obliczona ze wszystkich danych, większy wpływ globalnej miary zróżnicowania nie może być zaskoczeniem.



Ryc. 91. Zależność rozmiarów wariancji nuggetowej ( $C_0$ ) miesięcznych MSDO od odchylenia standardowego sumy opadu (SD) i pierwiastka średnich kwadratów różnic opadów w stosunku do najbliższego sąsiada (RMS). Kolor punktów danych oznacza wielkość ich odchylenia od powierzchni regresji (niebieski: < 1 SD reszt, zielony: 1-2 SD, żółty: 2-3 SD, czerwony: > 3SD) 

<sup>44</sup> Korelacja semicząstkowa lub częściowa jest podobna do korelacji cząstkowej. Stanowi ona, podobnie jak współczynnik korelacji cząstkowej, miarę skorelowania dwóch zmiennych, jaka pozostaje po uwzględnieniu (tzn. wyeliminowaniu) wpływów jednej lub wielu innych predyktorów, lecz zostaje wyskalowana (tzn. odniesiona) do całkowitej zmienności zmiennej zależnej (Statsoft 2004).

## 10. Podsumowanie rozdziału

- Danymi źródłowymi do przeprowadzonej analizy było 325 modeli struktury przestrzennej, po jednym dla każdego analizowanego zbioru miesięcznych i rocznych MSDO.

- Modele wykonano na podstawie empirycznych izotropowych semiwariogramów danych znormalizowanych o zasięgu 212,5 km (85 odstępów po 2,5 km). Umożliwiło to wiarygodną ocenę struktury przestrzennej od skali lokalnej do ponadregionalnej. Pominięcie w analizie potencjalnej anizotropii miało dwa aspekty: praktyczny (czasochłonność) i merytoryczny (częściowe uniknięcie maskowania heterogeniczności pola MSDO). Dzięki normalizacji danych można było bezpośrednio porównywać wartości semiwariancji empirycznych i modelowych poszczególnych ich zbiorów.

- Wszystkie 325 modeli danych znormalizowanych miało charakter złożony, składając się z co najmniej dwóch, a maksymalnie pięciu modeli elementarnych. W trakcie modelowania wykorzystywano funkcję nuggetową, sferyczną, wykładniczą i gaussowską.

- We wszystkich modelach znormalizowanych danych MSDO, zarówno miesięcznych, jak i rocznych, konieczne było użycie struktury nuggetowej. Wynika to z nieciągłości opadów dobowych oraz asynchroniczności danych MSDO. Dodatkowym źródłem tej zmienności są systematyczne i losowe błędy określenia objętości opadu oraz ich zróżnicowanie w małej skali.

- Bardzo pospolitą strukturą elementarną, zarówno przy miesięcznych, jak i rocznych zbiorach danych MSDO, była sferyczna. Oznacza to liniowy spadek podobieństwa wraz z odległością i zmienność przestrzenną o postaci przeplatających się, zbliżonych rozmiarami płatów wysokich i niskich wartości opadów. Proporcje przypadków użycia funkcji sferycznej są podobne dla poszczególnych składowych przy miesięcznych i rocznych zbiorach danych MSDO, i wynoszą 65-100%.

- W około 36% przypadków miesięcznych, jak i rocznych zbiorów danych MSDO do modelowania składowej pierwszej używano funkcji wykładniczej. Oznacza to, że na krótkich dystansach występuje często mozaika nieregularnych powierzchni wysokich i niskich opadów, a zasięg (średnica) owych płatów zmienia się w szerokim zakresie w sposób losowy. W większych skalach przestrzennych MSDO mają raczej zasięgi stałe (powtarzalne), a nie losowe. Stwierdzono, że struktura wykładnicza była częściej stosowana do modelowania struktury przestrzennej MSDO w miesiącach letnich.

- Najrzadziej przy modelowaniu struktury przestrzennej znormalizowanych danych MSDO używano funkcji gaussowskiej. Reprezentowała ona proces losowy zbliżony do ruchów Browna z przestrzennie skorelo-

wanymi odchyleniami. Stosowano ją do modelowania składowej o zasięgu znacznie większym niż maksymalny zakres przestrzenny analizy i wklęsłym kształcie.

- Zestawienie średnich modeli semiwariogramów dla poszczególnych miesięcy wykazało, że pomimo dużej ich zmienności istnieją wyraźne układy sezonowe. Duże znaczenie ma występowanie pojedynczych ekstremalnych przypadków, głównie miesięcy posusznych.

- Zbiór modeli struktury przestrzennej miesięcznych i rocznych MSDO został pogrupowany na podstawie dwóch kryteriów: (1) liczby i charakteru elementarnych składowych, z których zbudowany był model, (2) wielkości autokorelacji jako funkcji odległości.

- Efektem pierwszego podziału było 6 klas: (1) nugget + 1 składowa + trend, (2) nugget + 2 składowe + trend, (3) nugget + 3 składowe + trend, (4) nugget + 2 składowe, (5) nugget + 3 składowe. Klasa szósta (nugget + 1 składowa) zawierała tylko jeden przypadek. Roczne MSDO wykazują struktury przestrzenne jedynie typu drugiego, czwartego lub piątego. Najczęściej występuje typ drugi: w 45% MSDO miesięcznych i 44% – rocznych. Udział struktur typu 1, 4 i 5 dla zbioru danych miesięcznych podobny jest od ok. 15 do 18%. Struktura MSDO rocznych typu 4 występuje w 20% przypadków, typu 5 – w 36%.

- Występowanie pięciu podstawowych typów struktury przestrzennej dla danych miesięcznych nie wykazuje istotnej zmienności sezonowej.

- Podział modeli ze względu na wielkość autokorelacji jako funkcję odległości dał dwa warianty (4 i 7 klas) dla serii miesięcznych i 3 klasy dla rocznych. Klasyfikacja ta odzwierciedla przede wszystkim ciągłość zmian struktury pola MSDO, świadcząc, że jest ono „mieszaniną”, w której proporcje poszczególnych składników zmieniają się stopniowo. Testowanie sezonowości wykazało, że występuje ona przy bardziej szczegółowym podziale modeli serii miesięcznych MSDO (na 7 grup).

- Najbardziej powszechnie występująca jest autokorelacja miesięcznych MSDO o zasięgu od poniżej 10 do 40 km (średnio ok. 15,5 km, SD = 6,4 km). Drugie załamanie krzywej autokorelacji występuje przy odległościach wahających się od 25 do 180 km (średnio ok. 76,6 km, SD = 26,4 km.). Zasięg trzeciej struktury podobieństwa MSDO wynosi przeciętnie ok. 158 km (SD = 36,4 km) i waha się w przedziale od 90 do 250 km. Ocena zasięgu trendu występującego w wyróżnionych 1, 2 i 3 typie struktury przestrzennej MSDO może być jedynie szacunkowa. Prawdopodobnie są to odległości rzędu 250-350 km.

- Zasięgi składowych modeli rocznych MSDO są zbliżone, lecz systematycznie dłuższe: 17,5, 84,6 i 173,8 km.

- Przypuszczalnie składowa druga i trzecia dają statystyczny obraz działania tego samego anizotropowego zjawiska – intensywnych opadów fron-

talnych. W uzyskanych modelach izotropowych zatem raz pojawia się wyraźnie autokorelacja o zasięgu zgodnym z głównym kierunkiem anizotropii, raz prostopadła do niej; w innym zaś przypadku – obie wartości. Efekt zależy od liczby i konfiguracji stanowisk objętych opadem.

- Wieloletnie proporcje udziału zmienności poszczególnych składowych (wariancje  $C_0$ ,  $C_1$ ,  $C_2 + C_3$ , trendu) tworzących pole miesięcznych MSDO na obszarze Polski wynoszą jak 1 : 1,73 : 3,11 : 2,98.

- Szereg proporcji udziału zmienności składowych (wariancje  $C_0$ ,  $C_1$ ,  $C_2 + C_3$ , trendu) tworzących pole rocznych MSDO na obszarze Polski różni się większą rolą składowej pierwszej, a mniejszą trendu 1 : 3,11 : 3,92 : 2,12.

- Pole MSDO na obszarze Polski jest zatem najczęściej efektem sumowania się skutków działania trzech typów procesów operujących w różnych skalach przestrzennych: lokalnej (<10-20 km), regionalnej (50-150 km) i ponadregionalnej (> 200 km). Wyróżnione skale przestrzenne wiążą się prawdopodobnie z konwekcyjną/orograficzną, synoptyczną (frontalną) i „klimatologiczną” genezą wysokich opadów (fale planetarne?). Ich udział jest bardzo zmienny. Generalnie dominują jednak wysokie sumy dobowe opadów o rozciągłości przestrzennej od 50 do 150 km, związanych z migracją frontów atmosferycznych (35-38%).

- Bezwzględne rozmiary wariancji poszczególnych składowych wykazują silną sezonową zmienność, uzależnioną od rocznej cykliczności MSDO. Podobny charakter ma przebieg sezonowy zasięgu pierwszej składowej. Zasięgi pozostałych składowych nie wykazują żadnej, istotnej statystycznie, cykliczności.

- Zmienność sezonowa udziału wariancji nuggetowej jest bardzo zbliżona do średniego wieloletniego przebiegu współczynników korygujących błędy standardowych pomiarów opadów. Ich źródłem jest przede wszystkim turbulencja spowodowana wiatrem, występowanie opadów w postaci stałej, ale także straty związane z parowaniem i zwilżaniem. Stwierdzono również, że bezwzględna wielkość  $C_0$  jest uzależniona także od zmienności pola MSDO, zarówno globalnej, jak i lokalnej.

- Również względny udział wariancji pierwszej składowej wykazuje wysoce istotną zmienność sezonową z maksimum w czerwcu oraz maju i minimum w październiku. Taki cykl roczny odzwierciedla prawdopodobnie zmienny udział opadów konwekcyjnych – maksymalny w miesiącach o największych kontrastach termicznych.

- Zidentyfikowano istnienie, istotnej statystycznie, zmienności wieloletniej niektórych parametrów struktury przestrzennej MSDO. Był to przede wszystkim malejący trend wariancji nuggetowej. Nieco mniej wiarygodny był liniowy spadek wariancji drugiej składowej.

- Spadek  $C_0$  był znaczący przede wszystkim między grudniem a marcem, a najbardziej w styczniu i lutym. Przyczyną było najprawdopodobniej

zmniejszanie w tym okresie roku lokalnej zmienności opadów (por. rozdz. V) oraz być może spadek udziału opadów stałych.

- Stwierdzono również możliwość istnienia wieloletniej cyklicznej zmienności  $C_0$ ,  $C_1$  i  $A_1$ . W wypadku  $C_0$  okres zmian jest zbyt duży, aby na podstawie posiadanej 25-letniej serii można go było zidentyfikować. Trendy zmian  $C_1$  i  $A_1$  są identyczne. W obrębie analizowanego wielolecia mieści się jeden sinusoidalny cykl zmian o okresie 19 lat.



# VII

---

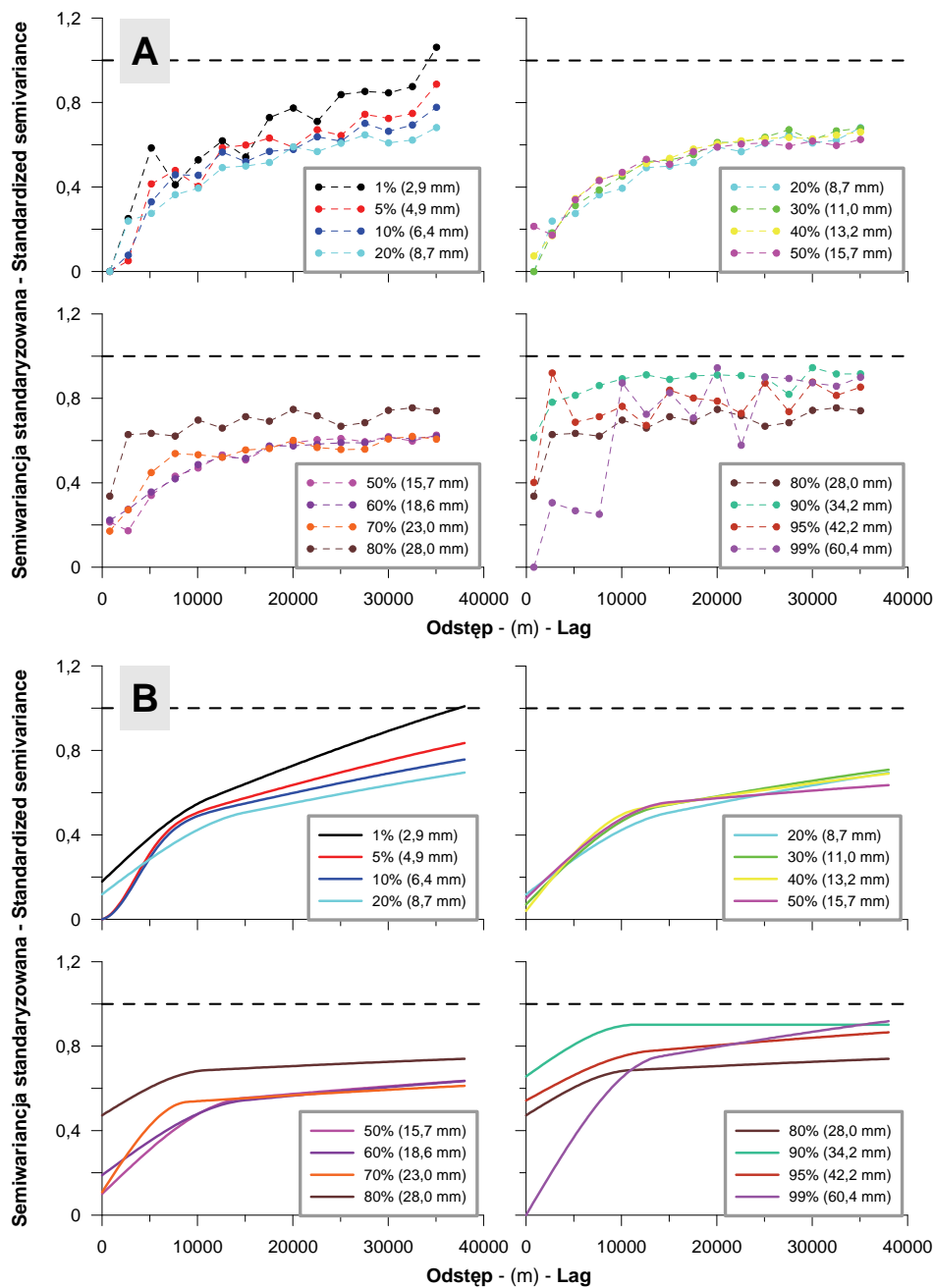
## **Analiza struktury przestrzennej MSDO na podstawie semiwariogramów danych kodowanych**

### **1. Wprowadzenie**

Drugim zbiorem informacji, umożliwiającym przeprowadzenie analizy struktury przestrzennej MSDO z nieco innego punktu widzenia, były semiwariogramy danych kodowanych<sup>44</sup> (patrz podrozdz. III.2.3 i III.2.4, ryc. 7-12). Oprócz potencjalnej zmienności sezonowej można również w tym wypadku rozpatrywać zróżnicowanie struktury w odniesieniu do względnej i bezwzględnej wartości dobowej sumy opadu. Z drugiej jednak strony liczba semiwariogramów danych kodowanych ( $325 \times 13 = 4225$ ) oraz ich pierwotne przeznaczenie – estymacja pola prawdopodobieństwa MSDO z pomocą IK – spowodowała, że zasięg analizy ograniczono do 36,25 km (14 odstępów po 2,5 km + 1 odstęp połówkowy o szerokości 1,25 km). Dane te umożliwiają zatem charakterystykę jedynie lokalnej struktury opadów. Miało to także dobre strony. Mniejsza liczba punktów danych empirycznych ułatwiała lepsze dopasowanie modelu i skutkowało uzyskaniem bardziej klarownych relacji. Należy jednak pamiętać, że sens obliczonych wartości semiwariancji jest nieco inny niż poprzednio (por. podrozdz. III.2.3). Dają one bowiem prawdopodobieństwo, że zmierzone sumy opadów na stanowiskach leżących od siebie w określonej odległości będą po „przeciwnych” stronach określonej wartości progowej (inaczej częstość różnic międzyklasowych

---

<sup>44</sup> Dane te (semiwariogramy empiryczne danych kodowanych oraz ich modele) są zapisane na dołączonym dysku DVD.



Ryc. 92. Przykład standaryzowanych semiwariogramów empirycznych (A) i ich modeli (B) dla sekwencji wartości progowych (percentyle: 1, 5, 10, 20, ..., 90, 95, 99) MSDO z lipca roku 1963. Dla ułatwienia porównań użyto tych samych kolorów, a na kolejnych wykresach powtarzany jest ostatni semiwariogram z poprzedniego wykresu

jako funkcja odległości). Ponieważ semiwariogramy empiryczne danych kodowanych były obliczane z „surowych” wektorów binarnych, żeby je (i ich modele również) móc ze sobą porównywać, należało jeszcze przeprowadzić operację standaryzacji. Polegała ona na podzieleniu obliczonych wartości semiwariancji empirycznych i modelowanych przez wariancję próby liczoną dla każdego analizowanego zbioru danych kodowanych. Na rycinie 92 przedstawiono przykład takich standaryzowanych semiwariogramów empirycznych i ich modeli dla jednego zbioru miesięcznych MSDO z lipca 1963. Ilustruje on dobrze zasadnicze prawidłowości zmienności struktury przestrzennej w różnych klasach wysokości sum opadów:

- względnie „regularną” strukturę opadów ze środkowego (percentyle 20-70%: od 8,7 do 23,0 mm) zakresu zmienności MSDO w przeciwieństwie do jego wartości skrajnych, szczególnie przy granicy górnej (percentyle 95 i 99%: powyżej 42,4 mm);

- relatywnie niskie i mało zmienne wartości semiwariancji nuggetowej dla niskich i średnich klas wysokości sum opadów (do 70% percentyla, do 23,0 mm) kontrastujące z sytuacją najczęściej występującą w klasach górnych;

- zróżnicowane proporcje udziału wariancji struktury pierwszej w stosunku do całkowitej jej zmienności wskazujące na zróżnicowany zasięg przestrzenny opadów w różnych klasach ich wysokości:

- przy górnej granicy rozkładu danych stanowi ona często jedyny element struktury przestrzennej; wariancja progowa osiągnięta jest na bardzo krótkim dystansie, rzędu kilku, kilkunastu kilometrów;

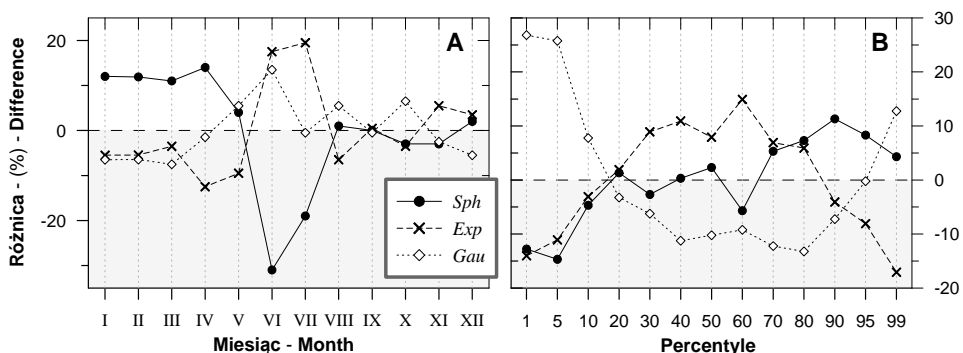
- przy dolnej granicy rozkładu danych gradient spadku autokorelacji składowej drugiej (i ewentualnie kolejnych) jest najwyższy; zasięg opadów prawdopodobnie nie przekracza znacząco zasięgu analizy, wynosząc przeciętnie kilkadziesiąt kilometrów;

- w środkowym zakresie gradient spadku autokorelacji składowej drugiej (i ewentualnie kolejnych) jest najniższy, co sugeruje możliwość występowania opadów bardzo rozległych, o zasięgu ponad 100 km.

W trakcie dalszej interpretacji wyników analizy modeli struktury przestrzennej danych kodowanych używane będzie dla uproszczenia sformułowanie: składowa druga i jako symbol jej wariancji –  $C_2$ . Ma ono tym razem jednakże nieco inne znaczenie niż w poprzednim rozdziale. O ile wówczas był to konkretny, dokładnie określony element modelu, to tym razem oznacza to całą ekstrapolowaną zmienność opadów w danej klasie ich wysokości poza zasięgiem pierwszej składowej. Odpowiada to zatem w przybliżeniu sumie „ $C_2 + C_3 + \text{trend}$ ” z analizy modeli danych znormalizowanych (por. rozdz. VI). Wartość  $C_2$  modeli danych kodowanych określano dla arbitralnie przyjętego stałego zasięgu 75 km, który pozwalał z zadowalającą dokładnością odwzorowywać liniowy najczęściej charakter wzrostu semiwariancji na odcinku poza zasięgiem pierwszej składowej.

## 2. Częstość elementarnych modeli i czynniki ją kształtujące

Do modelowania semiwariancji danych kodowanych używano czterech modeli elementarnych: nuggetowego, sferycznego, wykładniczego i gaussowskiego. Składową pierwszą miesięcznych MSDO w 89,2% przypadków (na 3900 modeli) przedstawiano używając modelu sferycznego, w 5,7% – wykładniczego i 5,1% – gaussowskiego. Ich zastosowanie wykazuje istotną regularność zarówno w odniesieniu do sezonu, jak i względnej wysokości opadu (percentyla, ryc. 93). W obu przypadkach różnica wartości oczekiwanych i obserwowanych jest w teście Chi-kwadrat istotna na poziomie  $p < 0,000001$ . W układzie sezonowym dotyczy to jedynie okresu do początku roku kalendarzowego do lipca (ryc. 93A). Najpierw do marca frekwencja używania modelu sferycznego przewyższa o ponad 10% tę oczekiwaną, wynikającą z rozkładu równomiernego. Wiąże się to z 5-7% wielkości „deficytem” wykorzystywania pozostałych dwóch modeli elementarnych. Sytuacja ta ulega dużej zmianie w każdym z kolejnych miesięcy. W kwietniu zaczyna rosnać udział modelu gaussowskiego, osiągając maksimum „nadmiaru” wynoszące około 13,5% w czerwcu. Od lipca rzeczywista frekwencja tego modelu oscyluje wokół oczekiwanej. Wykorzystanie struktury wykładniczej było najniższe w kwietniu i maju (12,5 i 9,5% poniżej oczekiwań), a następnie osiąga swoje maksimum w czerwcu i lipcu (17,5 i 19,5 % powyżej oczekiwań). W tych samych miesiącach występuje oczywiście silna „depresja” w krzywej używania dominującego modelu sferycznego (-31 i -19%). W pozostałych miesiącach (sierpień – grudzień) odchylenia od wartości oczekiwanej



Ryc. 93. Różnica pomiędzy obserwowanymi a oczekiwanymi frekwencjami wykorzystania modeli elementarnych (sferycznego, wykładniczego i gaussowskiego) do modelowania składowej pierwszej struktury przestrzennej danych kodowanych w relacji do pory roku (A) i względnej wysokości opadu (B)

są niewielkie. Interpretacja opisanego powyżej sezonowego zróżnicowania wymaga przywołania informacji przytoczonych w podrozdziale III.2.5.2 na temat specyfiki rozkładu przestrzennego reprezentowanego przez każdy z używanych tu modeli. Wynikałoby z tego, że w okresie zimowo-wiosennym w opadach o charakterze lokalnym dominują powtarzalne (bądź regularne) zasięgi. Powierzchnie komórek opadowych są zatem względnie zbliżone. W czerwcu i lipcu sytuacja zmienia się zasadniczo. Spektrum rozmiarów komórek opadowych ulega poszerzeniu, nabierając charakteru losowego.

Bardziej regularny jest rozkład frekwencji modeli elementarnych, użytych przy składowej pierwszej modeli kodowanych danych MSDO, w odniesieniu do względnej wysokości opadu (percentyla, ryc. 93B). Funkcję sferyczną relatywnie najrzadziej używano przy niskich progach (percentyle 1 i 5, ok. 15% poniżej oczekiwań). Od 10 do 60 percentyla jego frekwencja była zbliżona do oczekiwanej. Przy wyższych opadach (powyżej 70%) częstość używania modelu sferycznego była konsekwentnie wyższa od oczekiwanej (4-11%). Generalne przebiegi różnic frekwencji dla pozostałych dwóch wykorzystywanych modeli elementarnych są względem siebie lustrzanym odbiciem: wykładniczy - wypukły, gaussowski - wklęsły. W szczegółach jednakże widoczne są różnice. Model wykładniczy najrzadziej wykorzystywano przy skrajnych wartościach progowych (1 i 99 percentyl). Deficyt frekwencji był podobny - 14 i 17%. Od tych wartości krzywa częstości rosła i między 30 a 80 percentylem wahała się w zakresie od +5 do +15% - z maksimum przy 60 percentylu. Rozkład frekwencji wykorzystania modelu gaussowskiego był asymetryczny. Najczęściej, bo aż 26 i 25% powyżej „normy”, był używany przy 1 i 5 percentylu. Na drugim końcu skali względnej wysokości opadów również był częściej stosowany niż wynikałoby to ze średniej. Ale w tym wypadku nadmiar wynosił jedynie 12,5%. Największy deficyt używania modelu gaussowskiego (9-13%) miał miejsce między 40 a 80 percentylem. Także w tym momencie trzeba odnieść się do opisanego w podrozdziale VI.3 charakterystyki każdego z używanych w niniejszym opracowaniu elementarnych modeli struktury przestrzennej. Wydaje się, że w tym wypadku główną osią zróżnicowania nie jest powtarzalność/zróżnicowanie zasięgów komórek opadowych, jak to miało miejsce w ujęciu sezonowym. Ważniejsze zdaje się zróżnicowanie zjawiska na bardzo krótkich dystansach, które odróżnia modele sferyczne i wykładnicze z jednej strony, od modelu gaussowskiego - z drugiej. Dwa pierwsze charakteryzuje bowiem szybki, liniowy spadek autokorelacji, trzeci - względnie wolny, paraboliczny. Zróżnicowanie to szczególnie jest widoczne przy bardzo niskich opadach - pierwszym i piątym percentylu. Zatem, takie MSDO charakteryzowałyby się częstszą frekwencją występowania wyjątkowo niewielkiej zmienności na krótkich dystansach. Ich udział bardzo spada przy opadach w zakresie 30-80 percentyla. Relacje między modelami przy bardzo dużych względnych MSDO są

prawdopodobnie, ze względu na bardzo chaotyczny charakter semiwariogramów empirycznych, przypadkowe. Ich zestawienie bowiem – model sferyczny z gaussowskim „kontra” wykładniczy – nie sposób logicznie wytłumaczyć.

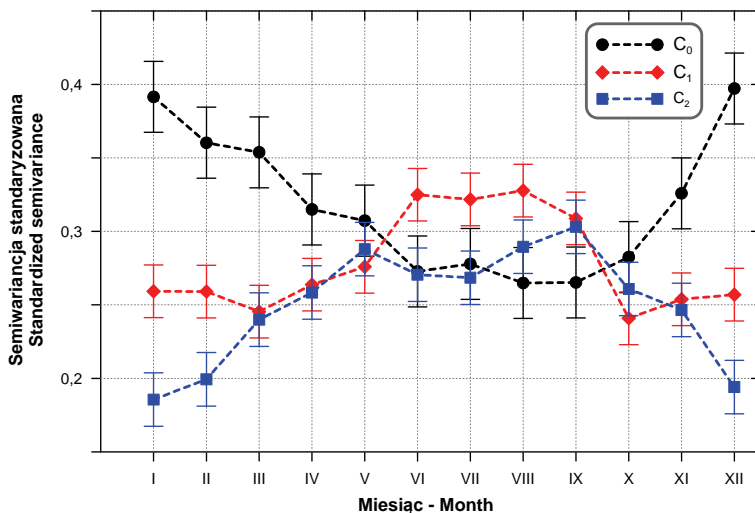
Proporcje wykorzystanych modeli elementarnych względem składowej pierwszej rocznych kodowanych MSDO były bardzo zbliżone. Również zdecydowanie najczęściej użyty był model sferyczny (295 przypadków na 325, 90,8%), a pozostałe 9,2% (30 przypadków) po połowie przypadło na modele wykładniczy i gaussowski. Relatywnie nieduża próba spowodowała, że w tym zbiorze nie ma widocznego istotnego zróżnicowania w zależności od względnej wysokości opadów.


### 3. Zróżnicowanie sezonowe

Do oceny istotności zróżnicowania standaryzowanych parametrów modeli danych kodowanych w zależności od pory roku i względnej wysokości opadów wykorzystano, jak poprzednio, analizę wariancji, ale tym razem w postaci wieloczynnikowej (MANOVA). Ogólne wyniki wskazują na bardzo wysoką istotność statystyczną różnic dla zmienności sezonowej (M-c) i względnej wysokości opadu (P) niezależnie wykorzystanego testu (tab. 15).

**Tabela 15.** Generalne wyniki wieloczynnikowej analizy wariancji (MANOVA) zależności standaryzowanych parametrów modeli danych kodowanych od pory roku (miesiąca, M-c), względnej wysokości opadu (percentyl, P) i ich interakcji (M-c\*P)

	Typ testu Test type	Wartość Value	Stat. F F stat.	df grup Effect df	df reszt Error df	Poziom $p$ $p$ level
Wyraz wolny Intercept	Wilksa	0,03702	24 324,10	4	3 740,00	0,000000
	Pillaisa	0,96298	24 324,10	4	3 740,00	0,000000
	Hotellinga	26,01508	24 324,10	4	3 740,00	0,000000
	Roya	26,01508	24 324,10	4	3 740,00	0,000000
Miesiąc (M-c) Month	Wilksa	0,88960	10,10	44	14 310,26	0,000000
	Pillaisa	0,11354	9,94	44	14972,00	0,000000
	Hotellinga	0,12058	10,25	44	14 954,00	0,000000
	Roya	0,08034	27,34	11	3 743,00	0,000000
Percentyl (P) Percentyle	Wilksa	0,47371	64,25	48	14 408,91	0,000000
	Pillaisa	0,57536	52,40	48	14 972,00	0,000000
	Hotellinga	1,00829	78,53	48	14 954,00	0,000000
	Roya	0,89579	279,41	12	3 743,00	0,000000
M-c*P	Wilksa	0,87446	0,97	528	14 958,09	0,699033
	Pillaisa	0,13176	0,97	528	14 972,00	0,702618
	Hotellinga	0,13659	0,97	528	14 954,00	0,695429
	Roya	0,04696	1,33	132	3 743,00	0,007541



**Ryc. 94.** Zmienność sezonowa średnich wartości wariancji poszczególnych składowych modeli struktury przestrzennej miesięcznych kodowanych danych MSDO. Zaznaczono 95% zakres ufności średniej 

W wypadku czynnika kombinacji pora roku/wysokość opadu (M-c\*P) tylko w teście Roya poziom istotności wskazywał istnienie różnic ( $p = 0,0075$ ). We wszystkich pozostałych testach  $p$  było znacznie większe od 0,5. Szczegółowe wyniki MANOVA względem poszczególnych parametrów pokazują nieco bardziej zróżnicowany obraz (tab. 16). Potwierdza się całkowity brak wpływu na analizowane parametry czynnika kombinacji pora roku/wysokość opadu (M-c\*P). Zasięg pierwszej struktury ( $A_1$ ) nie wykazuje zróżnicowania sezonowego ( $p > 0,6$ ), a jedynie zmienność zależną od względnej wysokości MSDO ( $p > 0,004$ ). W pozostałych wariantach układu „czynnik zmienności/parametry modeli” (M-c:  $C_0$ ,  $C_1$  i  $C_2$ ; P:  $C_0$ ,  $C_1$  i  $C_2$ ) istotność różnic jest bardzo wysoka –  $p$  w każdym przypadku jest mniejsze od  $1,0 \times 10^{-6}$ .

W zakresie, w którym można dokonać porównań z analizą przeprowadzoną uprzednio dla modeli danych znormalizowanych (tab. 16, M-c:  $A_1$ ,  $C_0$  i  $C_1$ ) zmienność czasowa jest wyraźniej zarysowana, ale w zasadzie identyczna. Co warto podkreślić – wyniki te zostały uzyskane niezależnie od siebie. Zbieżność ta jest szczególnie widoczna na wykresach przebiegu czasowego  $C_0$  (ryc. 79 i 94). Potwierdza się nie tylko generalny układ zmienności sezonowej, ale także pewne drugorzędne jego cechy, jak odchylenia od ogólnej tendencji widoczne w kwietniu i lipcu. Jest to tym bardziej godne uwagi, ponieważ skala tych odchyżeń jest znacznie mniejsza niż 95-procentowy przedział ufności średniej. Mniejszy relatywnie zakres owego przedziału, niż było to w wypadku danych znormalizowanych, spowodował, że więcej jest istotnych różnic przy porównywaniu poszczególnych

miesiący parami (testy *post hoc*) i są one znacznie silniejsze. Styczeń i grudzień są odmienne od ośmiu miesięcy między kwietniem a listopadem. Dla przedziału od maja do października prawdopodobieństwo błędu hipotezy o różnicach jest mniejsze od 0,0001. Wartości  $C_0$  dla lutego i marca różnią się od tych, które obliczono dla miesięcy od czerwca do października. Najbardziej specyficznym miesiącem jest listopad, który różni się zarówno od stycznia i grudnia, jak i sierpnia oraz września.

**Tabela 16.** Wyniki wieloczynnikowej analizy wariancji (MANOVA) zależności poszczególnych parametrów modeli danych kodowanych (A1,  $C_0$ ,  $C_1$  i  $C_2$ ) od pory roku (miesiąca, M-c), względnej wysokości opadu (percentyl, P) i ich interakcji (M-c\*P). Wariancje cząstkowe ( $C_0$ ,  $C_1$  i  $C_2$ ) zostały przed obliczeniami zestandardyzowane

		df	SK pomiędzy grupami	ŚK pomiędzy grupami	Test F	Poziom $p$
		df	SS effect	MS effect	F stat.	$p$ level
<b>A<sub>1</sub></b>						
W. wolny	Intercept	1	8,142541e <sup>11</sup>	8,142541e <sup>11</sup>	7601,380	0,000000
Miesiąc (M-c)	Month (M-c)	11	9,426540e <sup>8</sup>	8,569582e <sup>7</sup>	0,800	0,640322
Percentyl (P)	Percentyle (P)	12	3,075658e <sup>9</sup>	2,563049e <sup>8</sup>	2,393	0,004453
M-c*P	M-c*P	132	1,497168e <sup>10</sup>	1,134218e <sup>8</sup>	1,059	0,308801
Reszty	Error	3743	4,009474e <sup>11</sup>	1,071193e <sup>8</sup>		
Suma	Total	3898	4,199419e <sup>11</sup>			
<b>C<sub>0</sub></b>						
W. wolny	Intercept	1	393,8922	393,8922	8007,327	0,000000
Miesiąc (M-c)	Month (M-c)	11	8,2707	0,7519	15,285	0,000000
Percentyl (P)	Percentyle (P)	12	47,0863	3,9239	79,767	0,000000
M-c*P	M-c*P	132	6,0912	0,0461	0,938	0,679437
Reszty	Error	3743	184,1237	0,0492		
Suma	Total	3898	245,5758			
<b>C<sub>1</sub></b>						
W. wolny	Intercept	1	301,7155	301,7155	11147,30	0,000000
Miesiąc (M-c)	Month (M-c)	11	3,8742	0,3522	13,01	0,000000
Percentyl (P)	Percentyle (P)	12	26,5359	2,2113	81,70	0,000000
M-c*P	M-c*P	132	3,0793	0,0233	0,86	0,868731
Reszty	Error	3743	101,3089	0,0271		
Suma	Total	3898	134,8118			
<b>C<sub>2</sub></b>						
W. wolny	Intercept	1	244,3719	244,3719	8719,756	0,000000
Miesiąc (M-c)	Month (M-c)	11	5,4339	0,4940	17,627	0,000000
Percentyl (P)	Percentyle (P)	12	12,9491	1,0791	38,504	0,000000
M-c*P	M-c*P	132	2,9052	0,0220	0,785	0,965591
Reszty	Error	3743	104,8979	0,0280		
Suma	Total	3898	126,1921			



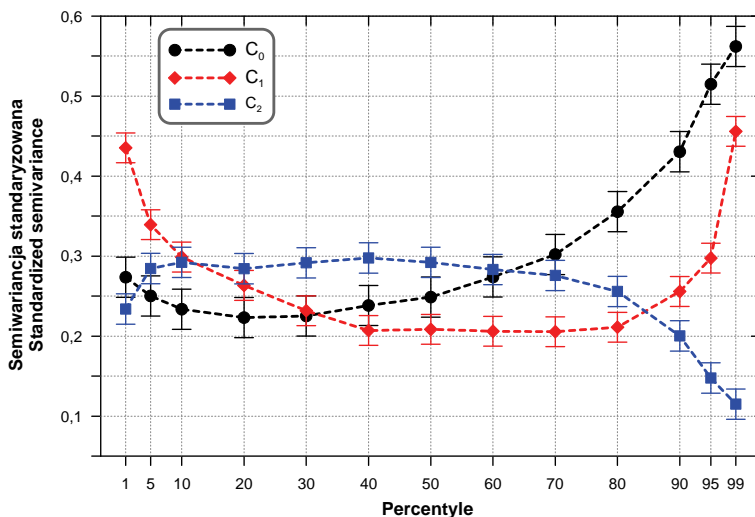
Sezonowa zmienność wariancji składowej pierwszej ( $C_1$ ) jest również generalnie taka sama, choć pojawiają się niewielkie różnice w istotnych detalach (ryc. 79 i 94). Dotyczy to zwłaszcza przebiegu maksimum krzywej. Z modeli danych znormalizowanych wynikało, że obejmuje on pięć miesięcy od maja do września, z wyraźnym pikiem w czerwcu. Obraz uzyskany z analizy modeli danych kodowanych pokazuje, że maj w tym aspekcie wyraźnie odróżnia się od miesięcy letnich, a bardzo słabo zarysowane maksimum występuje w sierpniu. Układ wśród miesięcy jesiennych jest identyczny, ale zimą najniższą wartością  $C_1$  cechuje się marzec, podczas gdy poprzednio był to luty. Należy jednakże podkreślić, że oprócz odmienności pozycji maja, wszystkie inne zasygnalizowane wyżej różnice mieszczą się w 95-procentowym przedziale ufności średniej. W porównaniach miesięcy parami bardzo wyraźnie rysuje się dwudzielność roku: jeden okres trwa od października do maja, drugi obejmuje pozostałe cztery miesiące. Maj, mimo że wykazuje duże podobieństwo do pozostałych miesięcy w grupie, ma najwięcej cech przejściowych.

Jak wspomniano już uprzednio, jako składową drugą ( $C_2$ ) modeli danych kodowanych określono ekstrapolowaną sumaryczną wariancję dla odstępów większego niż zasięg pierwszej struktury. Jej zmienność sezonową można zatem porównywać jedynie z sumą składowych  $C_2 + C_3$  modeli danych znormalizowanych lub nawet z dodaniem do niej także trendu. Hipoteza o zgodności tych właśnie rozkładów sezonowych znajduje potwierdzenie w rzeczywistości. Generalny przebieg średnich miesięcznych wartości  $C_2$  modeli danych kodowanych jest bowiem zbliżony do rozkładu sum składowej 2 i 3 przedstawionych na rycinie 79. Różnice między nimi są wszakże dość widoczne i bardziej znaczące niż w dwóch omawianych powyżej sytuacjach (dla  $C_0$  i  $C_1$ ). Minimum krzywej  $C_2$  danych kodowanych występuje konsekwentnie na identycznym poziomie na przełomie roku – w grudniu i styczniu. Poprzednio wyraźne najniższe wartości notowano tylko w styczniu, a kolejny miesiąc w szeregu rosnącym – listopad – charakteryzował się już wyższymi poziomem  $C_3$  i sumy  $C_2$  z  $C_3$ . Wzrost wartości  $C_2$  w pierwszej połowie roku zatrzymuje się w maju, podczas gdy poprzednio był to kwiecień. Maksimum też jest przesunięte o miesiąc: z sierpnia na wrzesień. Różnice te wydają się być związane ze zmiennością sezonową procesów działających w największej analizowanej skali przestrzennej – określanych jako trend. Przy porównaniu średnich wartości  $C_2$  w poszczególnych miesiącach parami wyraźnie zaznacza się względna homogeniczność okresu zimowego – miesiące od grudnia do lutego różnią się istotnie od wszystkich innych pomiędzy kwietniem a listopadem (grudzień i styczeń również z marcem). Przejściowy charakter mają marzec i listopad, które różnią się zarówno od grudnia i stycznia, jak i sierpnia oraz września.

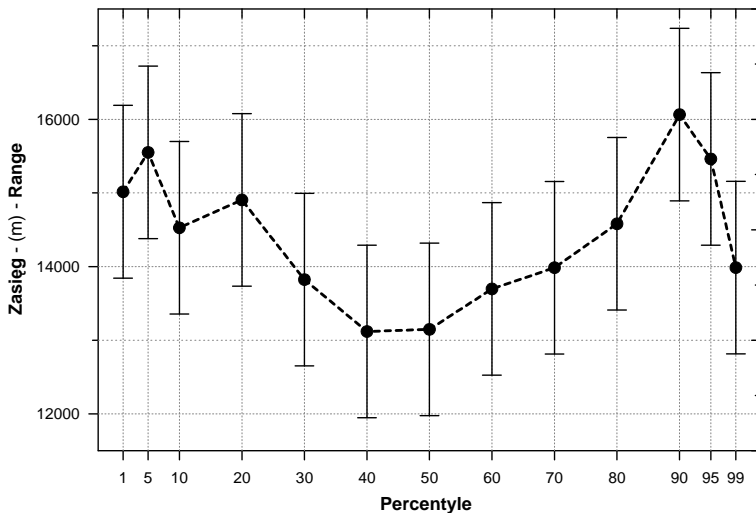
Zasadnicza odmienność reguł sezonowej zmienności parametrów struktury przestrzennej MSDO, ustalonych na podstawie danych znormalizowanych względem obecnie omawianych, dotyczy zasięgu pierwszej struktury  $A_1$ . Poprzednio stwierdzono wysoką istotność takich różnic ( $p < 0,0001$ ) i regularny przebieg zbliżony do występującego przy wariancji  $C_1$ . Tym razem bez wątplenia takiej regularności nie zanotowano ( $p = 0,64$ ). I nie wynika to jedynie z dużej szerokości 95% przedziału ufności średniej. Zmiany jej bowiem z miesiąca na miesiąc nie wykazują żadnego porządku. Różnica ta jest zastanawiająca i trudna do racjonalnego wytłumaczenia. Być może, procedura obliczania i modelowania semiwariancji odrębnie dla poszczególnych klas opadów wprowadziła dodatkowy element zmienności niwelujący efekt sezonowy.

#### 4. Zróźnicowanie zależne od względnej wysokości opadu

Najważniejszym wynikiem tego etapu prac było jednakże stwierdzenie charakteru zależności parametrów modeli struktury przestrzennej MSDO od względnej i bezwzględnej wysokości opadu (ryc. 95-111). Szczególnie te pierwsze są znacznie „wyraźniejsze” niż poprzednio, a zakres niepewności jest mniejszy.



Ryc. 95. Zmienność średnich wartości wariancji poszczególnych składowych modeli struktury przestrzennej kodowanych danych miesięcznych MSDO zależna od względnej wysokości opadu. Zaznaczono 95% zakres ufności średniej



**Ryc. 96.** Zmienność średnich wartości zasięgu pierwszej składowej modeli struktury przestrzennej kodowanych danych miesięcznych MSDO zależna od względnej wysokości opadu. Zaznaczono 95% zakres ufności średniej

Wariancja nuggetowa ( $C_0$ ) jest najniższa, bo wynosi około 0,22, przy opadach odpowiadających 20 percentylowi skumulowanej funkcji rozkładu miesięcznych MSDO (ryc. 95). Jej udział rośnie zarówno przy niższych, jak wyższych sumach dobowych. Wzrost ten przy najniższych opadach jest jednak stosunkowo niewielki i nie można wykluczyć, że stanowi częściowo artefakt związany z większą niepewnością tworzenia modeli dla skrajnych wartości progowych (por. podrozdz. III.2.5.8). Za taką interpretacją przemawia także wyższe odchylenie standardowe, a co za tym idzie – błąd standardowy określenia średniej. Z drugiej strony jednak, konsekwentne i stopniowe zmiany raczej wykluczają przypadkowość. Wzrost  $C_0$  dla wyższych względnych wartości miesięcznych MSOD ma charakter wykładniczy i jest bardzo płynny. Trudno wskazać na jakieś wyraźne załamania krzywej mogące świadczyć o zmianach o charakterze jakościowym. Tempo przyrostu maleje jednak nieznacznie pomiędzy 95 i 99 percentylem, co mogłoby wskazywać na logistyczny charakter zależności. Maksimum  $C_0$  ma wartość około 0,56. Wyniki testu *post hoc* Tukey wskazują, że biorąc pod uwagę kryterium wariancji, wartości  $C_0$  do 70 percentyla nie różnią się istotnie. Od 80% gradient zmian jest już tak duży, że kolejne średnie  $C_0$  różnią się od wszystkich pozostałych. Siedemdziesiąty percentyl jest odmienny zarówno od najniższych (10-40%), jak i najwyższych (90-99%).

Wariancja pierwszej składowej ( $C_1$ ) jest najniższa i względnie stała w przedziale od 40 do 80 percentyla, wynosząc około 0,21 (ryc. 95). Jej udział

rośnie zarówno przy niskich, jak i wysokich opadach. Gradient oraz rozmiar zmian różnią się jednak nieco. Wzrost  $C_1$  przy przejściu od 40 do 10 percentyla jest w zasadzie liniowy do poziomu około 0,3. Przy kolejnych dwóch najniższych wartościach progowych gradient przyrostu  $C_1$  jest coraz większy. Dla najniższych miesięcznych MSDO (1% krzywej skumulowanego rozkładu) wariancja pierwszej składowej wynosi ponad 0,43. Wzrost udziału  $C_1$  przy opadach wyższych od 80 percentyla ma charakter paraboliczny. Maksymalna wartość przy 99% wynosi 0,456. Jak można się domyślać, test *post hoc* wskazał na istnienie homogenicznej grupy w zakresie między 40 a 80 percentylem. Również rozkład wartości  $C_1$  dla 30 percentyla wykazuje duże do niej podobieństwo. Gradient zmian  $C_1$ , zarówno przy kolejnych niższych jak i wyższych progach MSDO, był tak duży, że różnią się one istotnie zazwyczaj w stosunku do opisaney wyżej „grupy”, jak i do wartości sąsiednich.

Zmienność obu wyżej wymienionych składowych w relacji do względnej wysokości opadów była w jednym aspekcie podobna – i  $C_0$ , i  $C_1$  najniższe udziały miały przy „średnich” MSDO, wykazując duży wzrost przy skrajnych wartościach progowych (w obu przypadkach maksimum było związane z 99% percentylem). Krzywa udziału  $C_2$  jest do pewnego stopnia ich „lustrzanym odbiciem” (ryc. 95). Maksymalne proporcje – rzędu 0,28-0,29 – występują między 5 a 60 percentylem; przy mniejszych i większych wartościach progowych są niższe. Szczególnie znaczący jest spadek udziału  $C_2$  poczynając od 80 percentyla rozkładu MSDO i wynosi 0,26-0,12. W przedziale od 5 do 60 percentyla średnie wartości  $C_2$  są do siebie podobne, a różnią się istotnie od tych z najniższej i trzech najwyższych klas MSDO. Siedemdziesiąty i osiemdziesiąty percentyl nie wykazują już istotnej odmienności w stosunku do pierwszego, dziewięćdziesiąty różni się od wszystkich oprócz pierwszego, a dwa najwyższe są podobne tylko do siebie nawzajem.

Bardziej złożony obraz daje relacja pomiędzy zasięgiem pierwszej składowej ( $A_1$ ) a względną wysokością miesięcznych MSDO (ryc. 96). Minimalny zasięg, średnio rzędu 13,1 km, dotyczy MSDO nieco mniejszych i zbliżonych do mediany, zaś maksymalny – dziewięćdziesiątego (16,0 km) i piątego percentyla (15,5 km). Skrajnie niskie (1%) i wysokie (95 i 99%) MSDO charakteryzują się nieco mniejszym zasięgiem. Rozrzut wartości jest jednakże tak duży, że statystycznie istotne są jedynie różnice średnich zasięgu  $A_1$  pomiędzy 40 i 50 a 90% rozkładu MSDO.

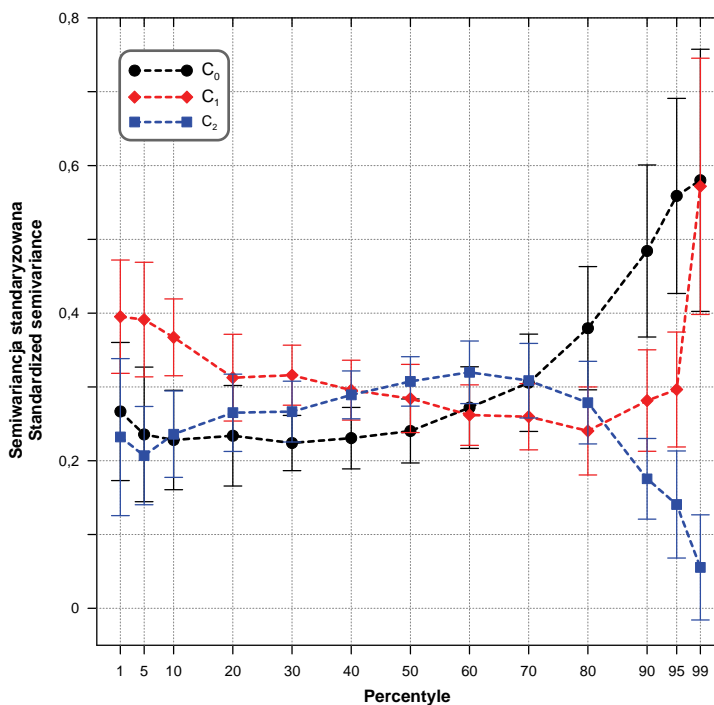
Podobną analizę przeprowadzono charakteryzując zmienność w odniesieniu do względnej wysokości opadów parametrów modeli rocznych kodowanych danych MSDO (tab. 17, ryc. 97 i 98). Generalne prawidłowości są zbliżone do tych ustalonych dla zbioru miesięcznych MSDO, lecz mniejsza ilość danych skutkowałą większym zakresem niepewności, w związku z tym mniejszą ilością statystycznie istotnych różnic. Najważniejsza wynikająca z tego odmienność to negatywny wynik ANOVA dla zróżnicowania za-

sięgu pierwszej struktury  $A_1$ . Opis wyników będzie skupiony głównie na podkreśleniu różnic między wynikami uzyskanymi dla miesięcznych i rocznych MSDO.

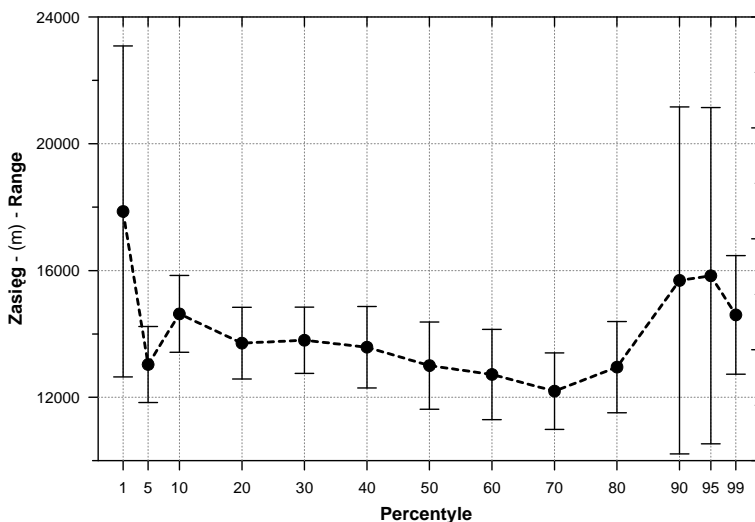
**Tabela 17.** Podsumowanie wyników testowania istotności zmienności parametrów modeli danych kodowanych struktury przestrzennej rocznych MSDO w zależności od względnej wysokości opadu z użyciem jednoczynnikowej analizy wariancji. Pogrubioną czcionką zaznaczono różnice istotne na poziomie  $p < 0,05$

Parametr	SK pomiędzy grupami	df grup	ŚK pomiędzy grupami	SK reszt	df reszt	ŚK reszt	test F	Poziom $p$
Parameter	SS effect	df effect	MS effect	SS error	df terror	MS error	F test	$p$ level
$A_1$	748 644 886	12	62 387 074	1,454579e <sup>10</sup>	312	46 621 132	1,338172	0,195512
$C_0$	5	12	0	15,24538	312	0	8,804782	0,000000
$C_1$	2	12	0	10,09755	312	0	5,967920	0,000000
$C_2$	2	12	0	6,563103	312	0	6,931703	0,000000

SK – suma kwadratów, df – liczba stopni swobody, ŚK – średnie kwadraty.



**Ryc. 97.** Zmienność średnich wartości wariancji poszczególnych składowych modeli struktury przestrzennej kodowanych danych rocznych MSDO zależna od względnej wysokości opadu. Zaznaczono 95% zakres ufności średniej



Ryc. 98. Zmienność średnich wartości zasięgu pierwszej składowej modeli struktury przestrzennej kodowanych danych rocznych MSDO zależna od względnej wysokości opadu. Zaznaczono 95% zakres ufności średniej

Najmniej jest ich w odniesieniu do wariancji nuggetowej ( $C_0$ ), a istniejące są najprawdopodobniej zupełnie przypadkowe. Dotyczy to zarówno samego przebiegu średnich dla poszczególnych progów MSDO, jak i ich bezwzględnych wartości.

Bardziej specyficznie wygląda obraz zmian udziału wariancji pierwszej struktury ( $C_1$ , ryc. 97). Zmniejsza się on bowiem dość regularnie od wartości około 0,4 przy pierwszym percentylu do 0,24 przy osiemdziesiątym. Nie ma tu zatem charakterystycznego dla miesięcznych danych MSDO „szerokiego” minimum między 40 a 80 percentylem. Znacznie większy jest też gradient przyrostu udziału  $C_1$  pomiędzy dwoma najwyższymi wartościami progowymi. Zakres niepewności jest jednak zbyt duży, żeby uznać to za niewątpliwą prawidłowość.

Również zmiany udziału wariancji struktury drugiej ( $C_2$ , ryc. 97) są bardziej „wyraziste” niż było to przy danych miesięcznych. Wzrost jej wartości jest bowiem konsekwentny od około 0,21 przy piątym percentylu do wyraźnego maksimum na poziomie 0,32 przy percentylu sześćdziesiątym. Następujący od tego miejsca spadek ma charakter paraboliczny i jest znacznie „głębszy” - minimum przy najwyższym progu wynosi zaledwie 0,06.

Zróznicowanie zasięgu pierwszej struktury ( $A_1$ ) rocznych MSDO w zależności od względnej wysokości opadu (ryc. 98) również wykazuje wy-

raźne odmierności względem relacji ustalonych dla danych miesięcznych (ryc. 96). Minimalny średni zasięg, wynoszący około 12,1 km, zanotowano przy 70% rozkładu rocznych MSDO, a nie jak poprzednio przy opadach zbliżonych do mediany. Minimum to jest także wyraźne, bo aż o 1 km, niższe. Inne różnice występujące przy skrajnych progach są, ze względu na dość chaotyczne skoki średniej  $A_1$  i duży zakres jej niepewności, raczej przypadkowe.

## 5. Zróżnicowanie zależne od bezwzględnej wysokości opadu – miejsce

W przeprowadzonej powyżej analizie porównywano charakterystyki statystyczne parametrów modeli struktury przestrzennej kodowanych danych miesięcznych i rocznych MSDO w odniesieniu do względnej wysokości opadów, wyrażonej w postaci określonego progu (percentyla) skumulowanej funkcji rozkładu. Oznaczało to, że w jednej klasie łączone były nieraz opady o bardzo różniących się sumach. Z drugiej zaś strony, różne, nawet nie „sąsiadujące” bezpośrednio percentyle rozkładu poszczególnych zbiorów danych MSDO mogły mieć identyczne wysokości opadu. Wynikało to zarówno z sezonowej, jak i wywołanej wszystkimi innymi czynnikami zmienności czasowej i przestrzennej MSDO (por. rozdz. V).

Jak duże było to zróżnicowanie pokazuje rycina 99. Zbiór percentyli pierwszych miesięcznych MSDO obejmował, na przykład, opady dobowe w przedziale 0-13,9 mm. Zakres ten zajął się nawet z wartościami notowanymi dla 99 percentyla (7,4-167,7 mm). Jakkolwiek uzyskane i omówione wyżej prawidłowości wskazują, że owa tak zmienna względna wysokość opadów ma bardzo duży wpływ na ich strukturę przestrzenną, postanowiono sprawdzić, czy podobnie będzie w odniesieniu do ich wartości bezwzględnych. Każdy z analizowanych 4225 percentyli ((300 miesięcy + 25 lat)  $\times$  13 progów) był w przygotowanej bazie danych opisany także wartością odpowiadającą mu sumy opadów. Posortowanie ich zatem w porządku uzależnionym od bezwzględnej wielkości opadu nie stanowiło żadnego problemu.

Relacje parametrów struktury przestrzennej MSDO w odniesieniu do bezwzględnej sumy opadów badano na dwa sposoby: parametryczny i nieparametryczny. W pierwszym podejściu określano je za pomocą funkcji regresji krzywoliniowej. Ze względu jednak na bardzo duży rozrzut danych, uzyskane wyniki nie były w pełni jednoznaczne. Dlatego, dla ich weryfikacji zastosowano także, jak poprzednio, analizę wariancji w odniesieniu do klas wysokości sum opadów. Były one ustalone osobno dla zbioru miesięcznych

i rocznych MSDO, i z pewnością bardziej „jednorodnie” niż pogrupowane względem percentyli. Ze względu na mocno zróżnicowaną liczebność przypadków MSDO w poszczególnych przedziałach ich wysokości, zastosowano zmienną szerokość klas. W wypadku danych miesięcznych były to przedziały co 2 mm od 0 do 20 mm, następnie co 5 mm do 50 mm. Wszystkie miesięczne MSDO o wysokości powyżej 50 mm połączono w jedną klasę. W tym ujęciu nie badano ewentualnej zmienności sezonowej. Roczne dane podzielono na 12 klas: do 20 mm, 20-25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-100 i powyżej 100 mm. Wyniki obliczeń ANOVA parametrów modeli struktury przestrzennej MSDO w odniesieniu do wyżej przedstawionego podziału na klasy sum opadów przedstawiono w tabelach 18 i 19. W obu analizowanych układach, miesięcznych i rocznych, uzyskano potwierdzenie bardzo wysokiej istotności różnic międzyklasowych ( $p < 0,000001$ )

**Tabela 18.** Podsumowanie wyników testowania istotności zmienności parametrów modeli danych kodowanych struktury przestrzennej miesięcznych MSDO w zależności od bezwzględnej wysokości opadu z użyciem jednoczynnikowej analizy wariancji. Pogrubioną czcionką zaznaczono różnice istotne na poziomie  $p < 0,05$

Parametr	SK pomiędzy grupami	df grup	ŚK pomiędzy grupami	SK reszt	df reszt	ŚK reszt	test F	Poziom $p$
Parameter	SS effect	df effect	MS effect	SS error	df error	MS error	F test	$p$ level
A <sub>1</sub>	1,654282 <sup>e9</sup>	16	103 392 626	4,182876 <sup>e11</sup>	3882	107 750 537	0,95956	0,499153
C <sub>0</sub>	<b>17,11920</b>	<b>16</b>	<b>1</b>	<b>228,5576</b>	<b>3883</b>	<b>0</b>	<b>18,17755</b>	<b>0,000000</b>
C <sub>1</sub>	<b>7,834559</b>	<b>16</b>	<b>0</b>	<b>127,0546</b>	<b>3883</b>	<b>0</b>	<b>14,96482</b>	<b>0,000000</b>
C <sub>2</sub>	<b>2,329983</b>	<b>16</b>	<b>0</b>	<b>123,9248</b>	<b>3883</b>	<b>0</b>	<b>4,56291</b>	<b>0,000000</b>

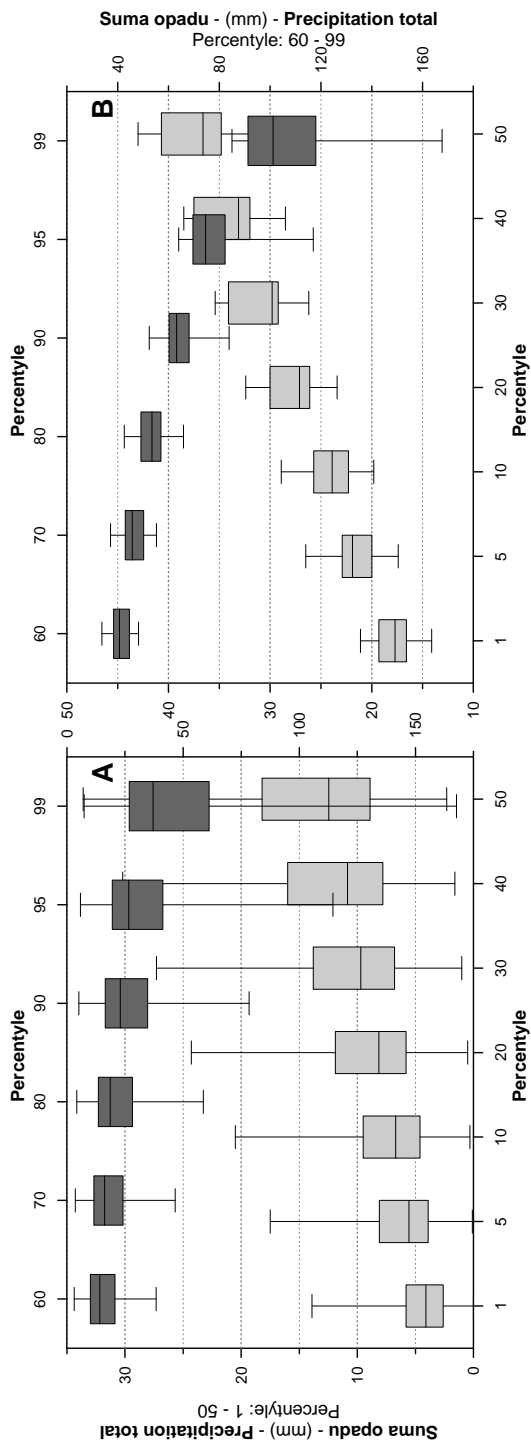
SK – suma kwadratów, df – liczba stopni swobody, ŚK – średnie kwadraty.

**Tabela 19.** Podsumowanie wyników testowania istotności zmienności parametrów modeli danych kodowanych struktury przestrzennej rocznych MSDO w zależności od bezwzględnej wysokości opadu z użyciem jednoczynnikowej analizy wariancji. Pogrubioną czcionką zaznaczono różnice istotne na poziomie  $p < 0,05$

Parametr	SK pomiędzy grupami	df grup	ŚK pomiędzy grupami	SK reszt	df reszt	ŚK reszt	test F	Poziom $p$
Parameter	SS effect	df effect	MS effect	SS error	df error	MS error	F test	$p$ level
A <sub>1</sub>	783 833 153	11	71 257 559	1,451060 <sup>e10</sup>	313	46 359 760	1,537056	0,117015
C <sub>0</sub>	<b>5</b>	<b>11</b>	<b>0</b>	<b>15,53581</b>	<b>313</b>	<b>0</b>	<b>8,923921</b>	<b>0,000000</b>
C <sub>1</sub>	<b>2</b>	<b>11</b>	<b>0</b>	<b>10,29770</b>	<b>313</b>	<b>0</b>	<b>5,851328</b>	<b>0,000000</b>
C <sub>2</sub>	<b>1</b>	<b>11</b>	<b>0</b>	<b>6,828523</b>	<b>313</b>	<b>0</b>	<b>6,185223</b>	<b>0,000000</b>

SK – suma kwadratów, df – liczba stopni swobody, ŚK – średnie kwadraty.





**Ryc. 99.** Rozkład statystyczny (minimum, pierwszy kwartyl, mediana, trzeci kwartyl, maksimum) bezwzględnych wysokości MSDO dla poszczególnych progów dla danych miesięcznych (A) i rocznych (B). Dla zwiększenia czytelności wykresów oś prawa (dla percentyli 60-99) jest skalowana odwrotnie

średnich wartości wariancji poszczególnych składowych ( $C_0$ ,  $C_1$  i  $C_2$ ). W odniesieniu do zasięgu pierwszej składowej ( $A_1$ ) istotnego zróżnicowania nie stwierdzono.

Do wyznaczenia funkcji matematycznych opisujących relacje między parametrami struktury przestrzennej MSDO a ich bezwzględną wysokością użyto wersji ewaluacyjnej programu TableCurve 2D 5.0.1 ([www.systat.com/products/TableCurve2D/](http://www.systat.com/products/TableCurve2D/)). Testuje ona dopasowanie wskazanych danych do ponad 3660 funkcji. Uzyskane modele mogą być następnie porządkowane według szeregu ich cech (ilości parametrów, współczynnika determinacji, błędu standardowego estymacji itp.), co ułatwia selekcję najbardziej z nich optymalnego, z punktu widzenia przyjętych kryteriów (prostota i/lub jakość dopasowania). W trakcie analizy możliwa jest też graficzna selekcja i tymczasowa eliminacja danych odstających i ponowne „przeliczenie” modeli. Program TableCurve ma także możliwość nieparametrycznego wygładzania (filtrowania) danych za pomocą funkcji sklepanych (*spline*) i fourierowskich. W niniejszym opracowaniu ograniczono zakres poszukiwania do modeli składających się maksymalnie z czterech parametrów (kryterium prostoty) o najwyższej jakości dopasowania. Często jednakże nie wybierano funkcji o najwyższym współczynniku determinacji. Niektóre z nich bowiem przy krańcach zakresu danych wykazywały chaotyczne bądź nierealistyczne przebiegi.

We wszystkich analizowanych kontekstach nierzadko zdarzało się, że na jedną wartość  $X$  (sumę dobową opadu) „przypadała” więcej niż jedna wartość  $Y$  (parametrów modeli struktury przestrzennej). Wariancję takich  $Y$  można traktować jako „czysty” czy losowy błąd. Jeśli w zbiorze danych istnieją takie replikacje zmiennej zależnej, to możliwe jest oddzielenie części wariancji związanej z „czystym” błędem od tej skojarzonej z dopasowanym modelem. Wykonywana jest wówczas rozszerzona analiza wariancji, obejmująca między innymi określenie maksymalnego możliwego współczynnika determinacji, który jest niezależny od wybranego modelu. Oprócz tego liczona jest statystyka *Lack Fit*, będąca różnicą całkowitej sumy kwadratów odchyłeń (reszt modelu) i tej jej części, która jest związana z wielokrotnymi  $Y$  przypadającymi na jednego  $X$  (czyli „czystym” błędem). Jeśli wartość  $F$  tej statystyki jest istotna ( $\geq 1$ ), wówczas wybrany model może być nieodpowiedni. W takich przypadkach pomocna jest szczegółowa analiza reszt w celu wykrycia systematycznych trendów, które mogą potwierdzić nieadekwatność modelu. W wypadku analizowanych danych wynika to przede wszystkim z niejednorodności genetycznej i ze zmienności sezonowej MSDO, i szukanie innego modelu niczego tu nie zmieni. Widać to szczególnie dobrze na wykresach dla rocznych zbiorów danych (ryc. 108-111), gdzie dość wyraźnie odróżniają się od siebie pewne układy punktów. Statystyki

LF (*Lack Fit*) zamieszczono w tabelach zbiorczych (20-22). Tam gdzie dawały one nierealistyczne wyniki, pola tabel są puste.

Na kolejnych rycinach (100-103) przedstawiono rozrzut wartości parametrów modeli wszystkich danych miesięcznych w relacji do wysokości opadów wraz z przebiegiem wybranej funkcji krzywoliniowej. Oprócz samej linii funkcji zaznaczono jej 95-procentowy przedział ufności. Poza tym, użyto funkcji typu Loess do zbudowania modelu lokalnych regresji (wygładzanie danych). Procedura ta łączy regresję lokalnie ważonych najmniejszych kwadratów (ang. *locally-weighted least-squares*), która ma charakter nieciągły, z ciągłością i właściwościami wygładzającymi funkcji sklepanych (B-spline). Algorytm ten czasami pozwala ujawnić ukryte struktury o niskiej częstotliwości w danych zawierających dużo „szumu”. Podstawowym parametrem tego algorytmu jest szerokość okna wygładzania, wyrażona w procentach

**Tabela 20.** Podsumowanie wyników modelowania relacji między parametrami modeli struktury przestrzennej miesięcznych kodowanych danych MSDO a bezwzględną wysokością opadów

Parametr Parameter	n	Stopień Order	R	R-max	SEE	F	<i>p</i>	F-LF	<i>p</i> -LF
C <sub>0</sub>	3900	4	0,0666	0,2858	0,243	94,020	0,00000	1,71977	0,00000
C <sub>1</sub>	3900	4	0,0545	0,2761	0,181	76,242	0,00000	1,71433	0,00000
A <sub>1</sub>	3803	4	0,0120		3793,149	16,738	0,00000		
C <sub>2</sub>	3896	4	0,0255	0,1967	0,174	35,337	0,00000	1,19555	0,00192

n - liczba uwzględnionych danych, Stopień - liczba parametrów modelu, R - współczynnik determinacji modelu, R-max - maksymalny możliwy współczynnik determinacji niezależnie od typu modelu, SEE - błąd standardowy estymacji, F - statystyka F modelu, *p* - poziom istotności modelu, F-LF - statystyka F braku dopasowania (*Lack Fit*), *p*-LF - poziom istotności braku dopasowania. Dokładne objaśnienia w tekście.

**Tabela 21.** Podsumowanie wyników modelowania relacji między parametrami modeli struktury przestrzennej rocznych kodowanych danych MSDO a bezwzględną wysokością opadów

Parametr Parameter	n	Stopień Order	R	R-max	SEE	F	<i>p</i>	F-LF	<i>p</i> -LF
C <sub>0</sub>	324	4	0,2539	0,8431	0,215	38,094	0,00000	1,09248	0,33337
C <sub>1</sub>	321	4	0,0969	0,8463	0,161	12,824	0,00000	1,43751	0,03421
A <sub>1</sub>	320	4	0,0405	0,2729	3379,023	5,843	0,00068	0,091021	1,00000
C <sub>2</sub>	322	4	0,2336	0,8259	0,130	34,051	0,00000	1,01554	0,48070

n - liczba uwzględnionych danych, Stopień - liczba parametrów modelu, R - współczynnik determinacji modelu, R-max - maksymalny możliwy współczynnik determinacji niezależnie od typu modelu, SEE - błąd standardowy estymacji, F - statystyka F modelu, *p* - poziom istotności modelu, F-LF - statystyka F braku dopasowania, *p*-LF - poziom istotności braku dopasowania. Dokładne objaśnienia w tekście.

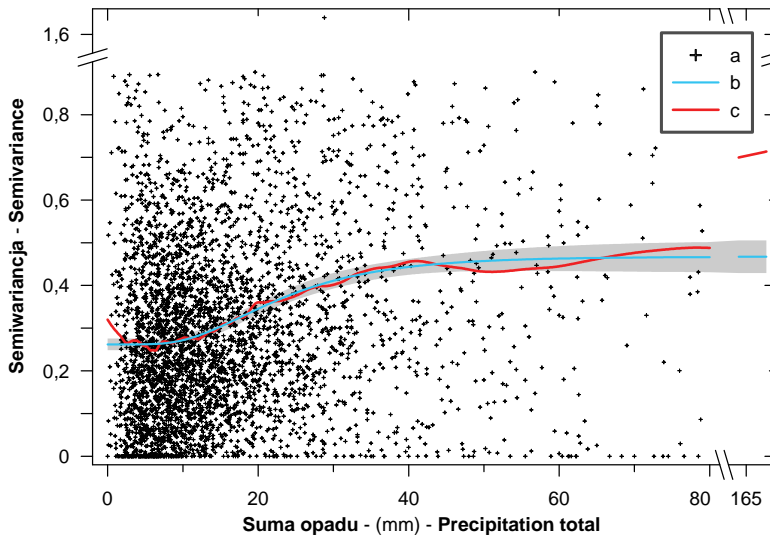
punktów całego zbioru danych. Szerokość ta jest zatem zmienna zależnie od lokalnej „gęstości” danych.

Zmienność wariancji nuggetowej w relacji do wysokości miesięcznych MSDO jest stosunkowo prosta (ryc. 100). Również rozbieżności między przebiegiem modelu globalnej a lokalnej regresji są relatywnie niewielkie. Do około 7 mm poziom  $C_0$  się nie zmienia i wynosi 0,26. Model Loess wskazuje jednakże, że przy skrajnie niskich MSDO błąd losowy jest prawdopodobnie znacząco wyższy. Przy wyższych sumach opadów następuje wzrost wariancji nuggetowej. Największy gradient przyrostu ma miejsce w przedziale 15-25 mm. Dalej wzrost  $C_0$  jest już coraz wolniejszy i osiąga swoje maksimum – około 0,46 – przy opadach przekraczających 50 mm na dobę. Rozrzut wartości jest jednakże olbrzymi. W całym zakresie wysokości MSDO potencjalna zmienność  $C_0$  jest w zasadzie maksymalna.

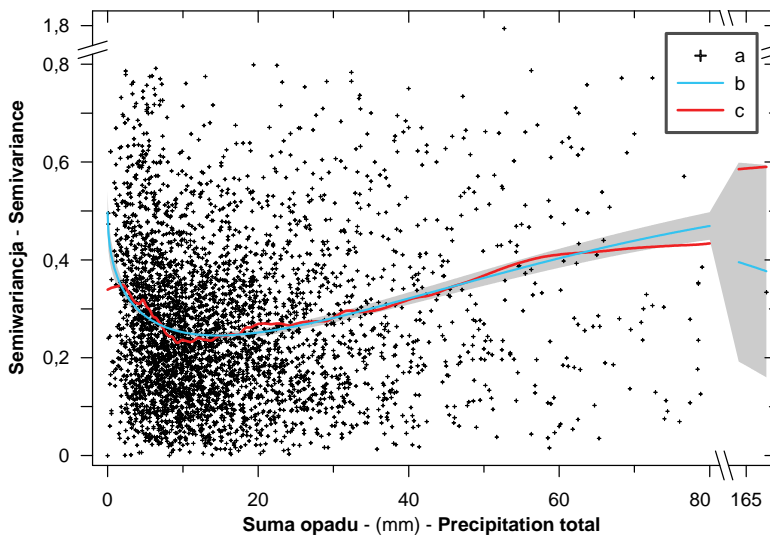
Skala zmian modelu regresji wariancji pierwszej składowej w relacji do wysokości miesięcznych MSDO jest podobna, lecz o innym jego przebiegu (ryc. 101). Przy najniższych opadach jest on relatywnie wysoki (ok. 0,36). Przy nieco wyższych – w zakresie 10 do 15 mm – funkcja osiąga minimum wynoszące około 0,25. Dalej następuje jej powolny liniowy wzrost, którego gradient zmniejsza się nieco po przekroczeniu sum dobowych wynoszących 50 mm. Maksimum, równe około 0,52 funkcja regresji osiąga przy opadach o wysokości około 110 mm. Po przekroczeniu jednakże poziomu 70 mm widoczne są duże rozbieżności między modelem globalnym a lokalnym. Zwraca uwagę również, że chmura punktów na wykresie, która przy niższych opadach jest względnie jednorodna, po przekroczeniu progu około 60 mm na dobę wykazuje wyraźną dwudzielność. Być może związana jest ona z wyraźniejszym zróżnicowaniem struktury przestrzennej najwyższych MSDO w zależności od ich genezy.

Model regresji  $C_2$  względem wysokości miesięcznych MSDO jest najprostszy (ryc. 102). Ponownie inne są tu prawidłowości zmian w klasie opadów najniższych – do 10 mm – względem całej reszty. Najpierw zatem funkcja nieco rośnie, aby od wymienionej wyżej granicy łagodnie, liniowo spadać. Wyraźniejsze, konsekwentne, różnice między modelem globalnym a lokalnym widoczne są od poziomu 45 mm sum dobowych opadu.

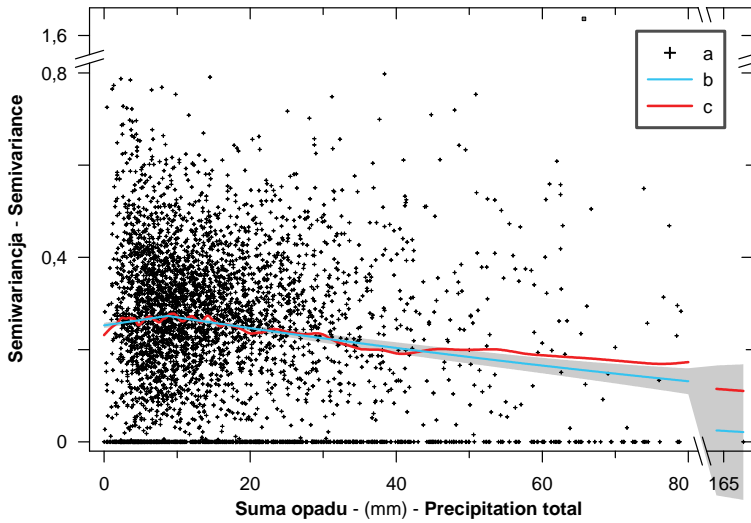
Bardzo niewielka jest amplituda zmian wartości regresji  $A_1$  względem wysokości miesięcznych MSDO (ryc. 103). Wynosi ona bowiem jedynie 1,2 km. Maksimum funkcji wynoszące około 13,5 km występuje zarówno przy najniższych sumach opadów (do 2 mm), jak i po przekroczeniu 30 mm. Najmniejszy zasięg pierwsza składowa ma przy opadach dobowych, rzędu 15 mm. Model lokalny fluktuuje wokół globalnego, nie wykazując większych odchyłań.



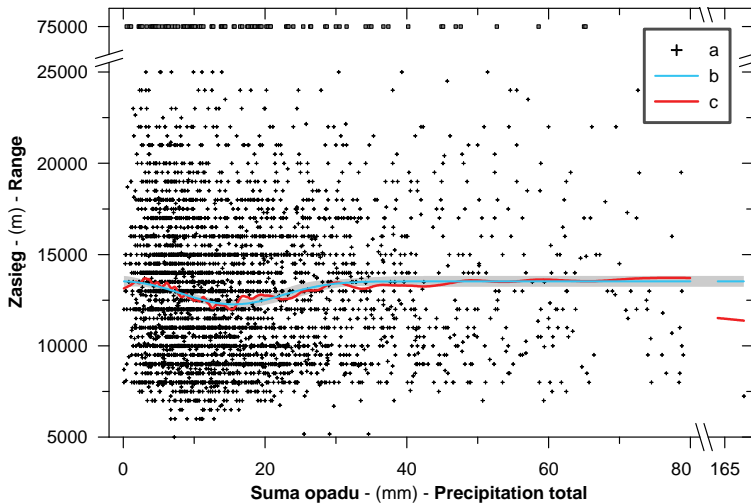
**Ryc. 100.** Zależna od bezwzględnej wysokości opadu zmienność wariancji nuggetowej modeli struktury przestrzennej miesięcznych kodowanych danych MSDO: a - dane, b - krzywa regresji wielomianowej z zaznaczeniem 95% przedziału ufności, c - rozkład danych wygładzony za pomocą funkcji sklepanych (Loess, B-spline rzędu 3, 12% wygładzania) ☐



**Ryc. 101.** Zależna od bezwzględnej wysokości opadu zmienność wariancji pierwszej składowej modeli struktury przestrzennej miesięcznych kodowanych danych MSDO: a - dane, b - krzywa regresji wielomianowej z zaznaczeniem 95% przedziału ufności, c - rozkład danych wygładzony za pomocą funkcji sklepanych (Loess, B-spline rzędu 3, 10% wygładzania) ☐



**Ryc. 102.** Zależna od bezwzględnej wysokości opadu zmienność wariancji drugiej składowej modeli struktury przestrzennej miesięcznych kodowanych danych MSDO: a – dane, b – krzywa regresji wielomianowej z zaznaczeniem 95% przedziału ufności, c – rozkład danych wygładzony za pomocą funkcji sklejących (Loess, B-spline rzędu 3, 10% wygładzania). Z obliczeń usunięto 4 najbardziej odstające wyniki. Jeden z nich widoczny na wykresie zaznaczono innym symbolem niż dane pozostałe ☐



**Ryc. 103.** Zależna od bezwzględnej wysokości opadu zmienność zasięgu pierwszej składowej modeli struktury przestrzennej miesięcznych kodowanych danych MSDO: a – dane, b – krzywa regresji wielomianowej z zaznaczeniem 95% przedziału ufności, c – rozkład danych wygładzony za pomocą funkcji sklejących (Loess, B-spline rzędu 3, 8% wygładzania). Z obliczeń usunięto 97 najbardziej odstających wyników. Widoczna na wykresie ich część została zaznaczona innym symbolem niż dane pozostałe ☐

## 6. Sezonowe zróżnicowanie zależne od bezwzględnej wysokości opadu

We wszystkich omawianych powyżej kontekstach parametry struktury przestrzennej miesięcznych MSDO wykazują bardzo duży rozrzut wartości w relacji do bezwzględnej sumy opadów. Mimo to wszystkie cztery modele regresji są wysoko istotne statystycznie ( $p < 0,00001$ , tab. 20). Występowanie wielu Y względem tych samych wartości X powoduje jednak, że maksymalne możliwe współczynniki determinacji niezależne od typu modelu są bardzo niskie - rzędu 0,19-0,28. Także wysokie wartości statystyk *F Lack Fit* wskazują, że dobór modelu jest przy tego typu danych obciążony dużą dozą niepewności. Część obserwowanej zmienności parametrów struktury przestrzennej miesięcznych MSDO jest efektem, co wykazano poprzednio, zróżnicowania sezonowego. Jaki jest rozmiar tego zróżnicowania i jaka jego charakterystyka jakościowa można było łatwo zbadać używając analogicznej metodyki. Na kolejnych rycinach (ryc. 104-107) oraz w tabeli 22 przedstawiono zatem wyniki modelowania zmienności  $C_0$ ,  $C_1$ ,  $C_2$  i  $A_1$  dla poszczególnych miesięcy. Dla zwiększenia czytelności bardzo już skomplikowanych wykresów zamieszczono na nich jedynie linie regresji globalnych, a szerokość 95-procentowego pasa ufności funkcji przedstawiono tylko dla stycznia i lipca.

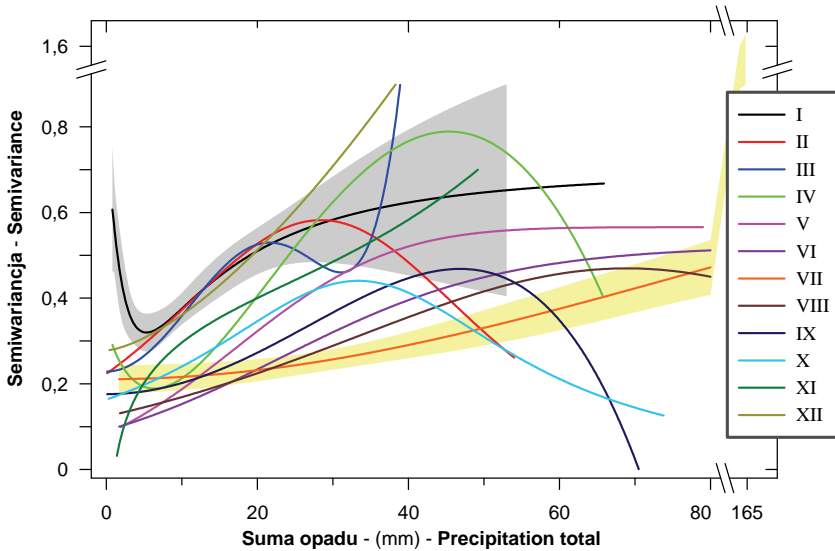
Relacja wariacji nuggetowej do wysokości opadów jest bez wątpienia zróżnicowana sezonowo (ryc. 104). Jest to szczególnie wiarygodne dla przedziału wysokości opadów od 5 do 30/40 mm. W każdym, w zasadzie, miesiącu MSDO o takiej wysokości notowano dostatecznie często, żeby uzyskana funkcja regresji była wyznaczona z dużą precyzją. Różnice przy niższych i wyższych opadach są mniej pewne. Wszystkie miesięczne równania regresji  $C_0$  są wysoko istotne statystycznie ( $p < 0,00001$ ), lecz dla większości z nich istotna jest również statystyka *Lack Fit*. Najbardziej charakterystyczna jest zwarta „wiązka” linii miesięcy zimowych od grudnia do marca przy opadach w zakresie 5 do przeszło 20 mm. Dla tego przedziału typowy jest wysoki gradient przyrostu  $C_0$  wraz ze zwiększaniem się opadów. Przy wyższych sumach dobowych krzywe regresji dla tych miesięcy zaczynają się odróżniać, ale znaczna część tej zmienności mieści się w zakresie niepewności funkcji. W kolejnych miesiącach wiosennych i letnich, do lipca, widoczna jest generalna tendencja do zmniejszania się wariacji nuggetowej i mniejszego gradientu jej przyrostu w relacji do sum opadów. Od sierpnia do listopada tendencja jest odwrotna. Do reguły tej „nie pasuje” jedynie kwiecień z najszybszymi przyrostami  $C_0$  w przedziale opadów 10-40 mm na dobę. Dla większości miesięcy charakterystyczny jest monotoniczny wzrost wariacji nuggetowej ze wzrostem sumy opadów. Są jednak wyjątki od tej reguły, kiedy po przekroczeniu pewnej granicy następuje spadek wielkości  $C_0$ . Naj-

bardziej wiarygodne wydają się takie przypadki w lutym i październiku, ponieważ przebieg krzywej występuje przy stosunkowo niskich sumach opadów (28 i 32 mm). W odniesieniu do lutego, a także w pewnym stopniu marca, można by wysunąć przypuszczenie, że jest to efektem zmiennej proporcji opadów stałych i ciekłych w MSDO, w zależności od ich wysokości. Mechanizm odpowiedzialny za podobny efekt spadku wariancji nuggetowej przy wysokich opadach w październiku jest z pewnością inny. Warto zwrócić uwagę, że podobne przebiegi funkcji, ale już mniej pewne, bo występujące przy jeszcze wyższych opadach, jest widoczne we wrześniu, a także sierpniu. Być może znaczenie ma tu prędkość wiatru, wpływająca na błąd pomiaru, statystycznie różna przy wysokich opadach o odmiennej genezie.

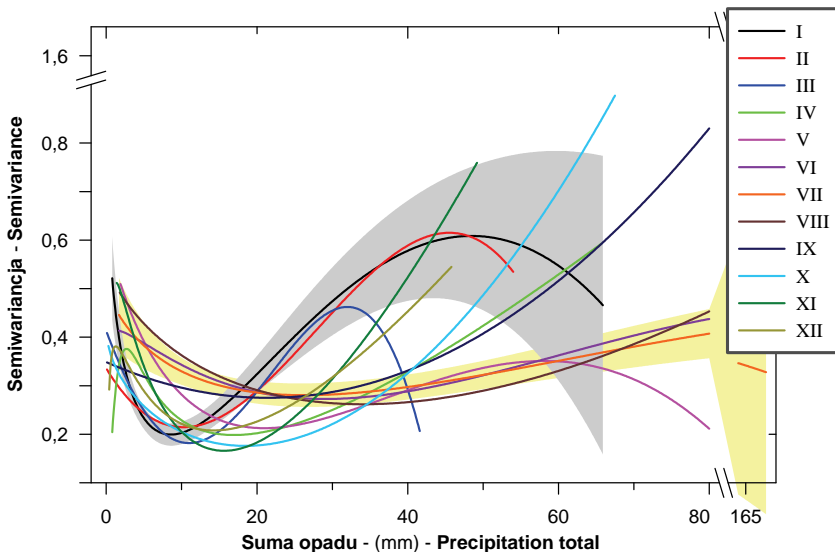
Zmienność sezonowa wariancji pierwszej składowej jest jeszcze silniej zarysowana (ryc. 105). Tym razem wyraźnie odróżniają się trzy grupy miesięcy: styczeń - marzec, październik - grudzień i maj - kwiecień oraz czerwiec - wrzesień. Druga z nich jest nieco niejednorodna. Przy niskich MSDO (do 10-15 mm) przebieg funkcji  $C_1$  w pierwszej i drugiej grupie jest podobny: silny spadek wartości. Dalej zaznaczają się już różnice. W miesiącach zimowych minimum  $C_1$  występuje przy sumach dobowych, rzędu 8-12 mm. Przy wyższych opadach wariancja pierwszej składowej bardzo szybko rośnie. Jesienią i wiosną minimum funkcji jest nieco „głębsze” i przesunięte w stronę nieco wyższych opadów, a gradient występującego dalej wzrostu  $C_1$  jest raczej niższy. Trzecia wymieniona cecha nie dotyczy listopada i grudnia, które pod tym względem są podobne do miesięcy zimowych. Latem zróżnicowanie wielkości wariancji pierwszej składowej jest w porównaniu z pozostałą częścią roku niewielkie - i prawdopodobnie w dużym stopniu nieistotne statystycznie. Równania regresji czerwca, lipca i września są bowiem istotne na poziomie mniejszym od 0,001, podczas gdy dla pozostałych miesięcy  $p < 0,00001$ . Najważniejszy jest jednak wyraźnie większy udział  $C_1$  latem przy niskich opadach (do 15 mm) i niższy przy sumach przekraczających 30-40 mm. Jest to prawdopodobnie skutkiem zmiennej sezonowo proporcji w miesięcznych MSDO krótkodystansowych opadów konwekcyjnych.

Wariancja drugiej składowej jest zróżnicowana sezonowo przede wszystkim przy średnich i wysokich MSDO - powyżej 10 mm na dobę (ryc. 106). Przebieg funkcji w 9 przypadkach na 12 jest podobny: paraboliczny wzrost do maksimum, a później liniowy lub wykładniczy spadek. Różnią je gradienty zmian i położenie maksimum. Trzy wyjątki charakteryzujące się konsekwentnym spadkiem  $C_2$  wraz z wysokością opadu to wrzesień, październik i listopad. Najniższe maksimum wariancji drugiej składowej (ok. 0,22) występujące przy relatywnie niskich opadach (3-10 mm) jest charakterystyczne dla miesięcy zimowych (grudzień - luty). Także dość wysoki





Ryc. 104. Zależna od bezwzględnej wysokości opadu zmienność wariancji nuggetowej modeli struktury przestrzennej kodowanych danych MSDO dla poszczególnych miesięcy zilustrowana za pomocą funkcji wielomianowych. Dla stycznia i lipca zaznaczono 95% przedział ufności funkcji



Ryc. 105. Zależna od bezwzględnej wysokości opadu zmienność wariancji pierwszej składowej modeli struktury przestrzennej kodowanych danych MSDO dla poszczególnych miesięcy zilustrowana za pomocą funkcji wielomianowych. Dla stycznia i lipca zaznaczono 95% przedział ufności funkcji

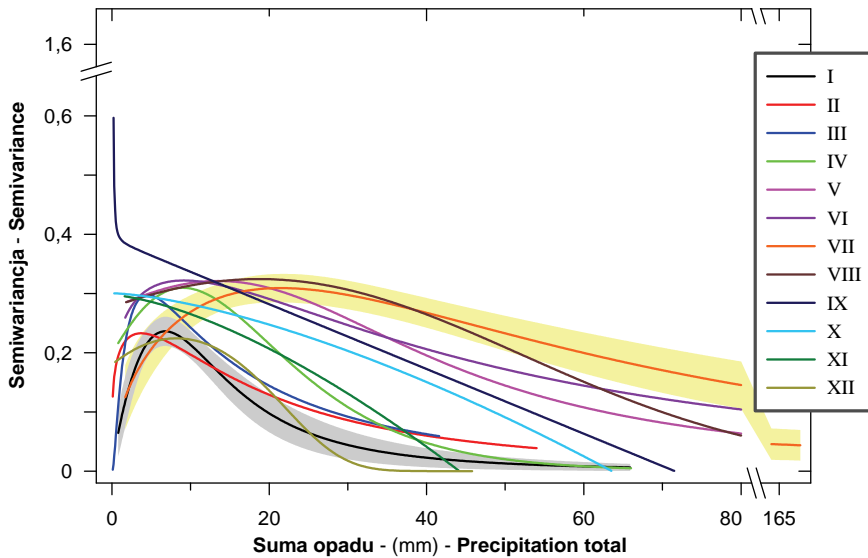
**Tabela 22.** Podsumowanie wyników modelowania relacji między parametrami modeli struktury przestrzennej kodowanych danych MSDO dla poszczególnych miesięcy a bezwzględną wysokością opadów ☒

Parametr Parameter	N	Stopień Order	R	R-max	SEE	F	<i>p</i>	F-LF	<i>p</i> -LF
<b>Styczeń - January</b>									
C <sub>0</sub>	325	4	0,0678	0,6147	0,251	9,213	0,00001	1,451	0,00930
C <sub>1</sub>	323	4	0,1403	0,5203	0,159	18,909	0,00000	0,816	0,90123
A <sub>1</sub>	313	3	0,0257		3856,244	5,625	0,00398		
C <sub>2</sub>		3	0,0993	0,5173	0,147	19,429	0,00000	0,881	0,78883
<b>Luty - February</b>									
C <sub>0</sub>	324	4	0,1457	0,6273	0,219	19,751	0,00000	1,1581	0,17819
C <sub>1</sub>	321	4	0,0603	0,5236	0,182	8,203	0,00003	0,873	0,80333
A <sub>1</sub>	309	3	0,0212		3535,048	4,855	0,00840		
C <sub>2</sub>	323	3	0,0508	0,5218	0,156	10,147	0,00005	0,885	0,77943
<b>Marzec - March</b>									
C <sub>0</sub>	325	4	0,1995	0,7113	0,220	28,340	0,00000	1,564	0,00255
C <sub>1</sub>	325	4	0,1431	0,6201	0,155	19,424	0,00000	1,104	0,26631
A <sub>1</sub>	314	3	0,0896		3354,071	16,951	0,00000	-0,634	0,00000
C <sub>2</sub>	324	3	0,1067	0,5964	0,161	20,864	0,00000	1,073	0,32962
<b>Kwiecień - April</b>									
C <sub>0</sub>	325	4	0,3557	0,8050	0,210	61,139	0,00000	1,601	0,00204
C <sub>1</sub>	321	4	0,1375	0,6419	0,139	18,394	0,00000	0,971	0,57642
A <sub>1</sub>	316	3	0,0326	-0,8727	3528,558	6,837	0,00124	-0,347	0,00000
C <sub>2</sub>	322	3	0,1863	0,4913	0,146	38,355	0,00000	0,407	1,00000
<b>Maj - May</b>									
C <sub>0</sub>	322	3	0,2962	0,8342	0,206	69,277	0,00000	1,339	0,05284
C <sub>1</sub>	323	4	0,1574	0,8028	0,141	21,448	0,00000	1,345	0,04997
A <sub>1</sub>	316	3	0,0109		3680,230	3,243	0,04037	-0,225	0,00000
C <sub>2</sub>	323	3	0,1012	0,7655	0,160	19,684	0,00000	1,146	0,22701
<b>Czerwiec - June</b>									
C <sub>0</sub>	325	3	0,2518	0,8732	0,202	56,184	0,00000	1,794	0,00096
C <sub>1</sub>	323	4	0,0461	0,8196	0,177	6,536	0,00027	1,584	0,00695
A <sub>1</sub>	323	3	0,0018	0,6926	3792,236	1,793	0,16816	0,815	0,88348
C <sub>2</sub>	322	3	0,1532	0,7517	0,146	30,630	0,00000	0,864	0,80268
<b>Lipiec - July</b>									
C <sub>0</sub>	322	4	0,1760	0,8113	0,206	24,257	0,00000	1,005	0,50333
C <sub>1</sub>	324	4	0,0492	0,8038	0,163	6,930	0,00016	1,139	0,25732
A <sub>1</sub>	319	3	0,0035	0,8284	3643,931	2,066	0,12835	1,427	0,03706
C <sub>2</sub>	322	3	0,0918	0,8102	0,152	17,776	0,00000	1,150	0,24036

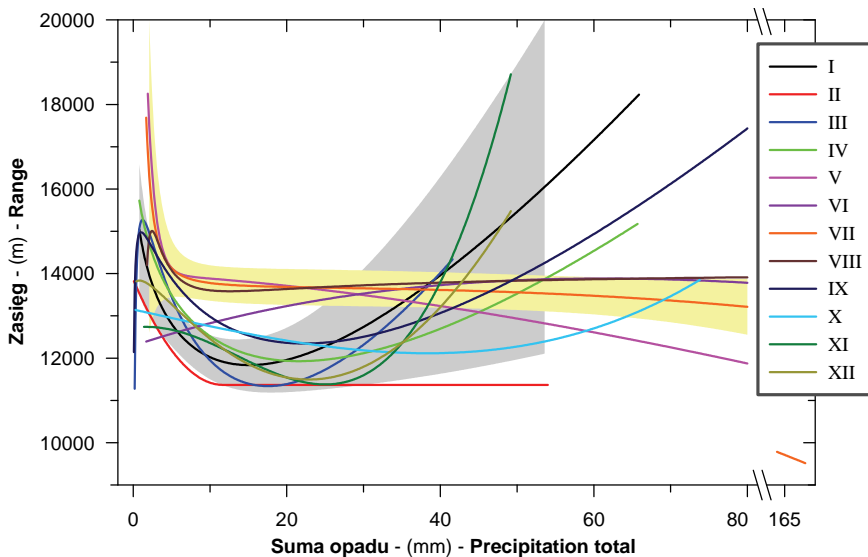
Sierpień - August									
C <sub>0</sub>	323	3	0,1625	0,8617	0,216	32,847	0,00000	1,399	0,04826
C <sub>1</sub>	325	3	0,1674	0,8446	0,169	34,187	0,00000	1,218	0,16309
A <sub>1</sub>	323	3	0,0000	0,7485	3884,089	0,420	0,65733	0,830	0,84660
C <sub>2</sub>	321	3	0,1527	0,8375	0,158	30,427	0,00000	1,197	0,18586
Wrzesień - September									
C <sub>0</sub>	324	3	0,1417	0,7820	0,201	28,260	0,00000	1,351	0,04295
C <sub>1</sub>	321	3	0,0337	0,7491	0,160	7,096	0,00097	1,291	0,07376
A <sub>1</sub>	319	3	0,0251	0,7295	3877,011	5,617	0,00401	1,190	0,16280
C <sub>2</sub>	321	3	0,2235	0,7489	0,160	47,700	0,00000	0,947	0,63409
Październik - October									
C <sub>0</sub>	325	3	0,0794	0,7176	0,247	15,517	0,00000	1,222	0,11745
C <sub>1</sub>	324	3	0,2238	0,7238	0,156	48,208	0,00000	0,983	0,54809
A <sub>1</sub>	314	3	0,0000		3755,738	1,011	0,36489	-0,383	0,00000
C <sub>2</sub>	323	2	0,0856	0,6603	0,160	32,243	0,00000	0,922	0,69374
Listopad - November									
C <sub>0</sub>	322	3	0,1660	0,6377	0,219	33,562	0,00000	0,952	0,62452
C <sub>1</sub>	322	3	0,2651	0,6807	0,145	59,594	0,00000	0,976	0,56291
A <sub>1</sub>	316	3	0,0199	-0,5382	3636,159	4,712	0,00963	-0,276	0,00000
C <sub>2</sub>	323	2	0,0943	0,6180	0,160	35,636	0,00000	1,023	0,44721
Grudzień - December									
C <sub>0</sub>	322	2	0,2255	0,6215	0,218	95,752	0,00000	0,889	0,77123
C <sub>1</sub>	325	3	0,0667	0,6288	0,181	13,118	0,00000	1,268	0,06829
A <sub>1</sub>	316	3	0,0285		3719,956	10,792	0,00002	-1,269	0,00000
C <sub>2</sub>	323	3	0,1281	0,6455	0,139	25,224	0,00000	1,221	0,10632

n - liczba uwzględnionych danych, Stopień - stopień wielomianu, R - współczynnik determinacji modelu, R-max - maksymalny możliwy współczynnik determinacji niezależnie od typu modelu, SEE - błąd standardowy estymacji, F - statystyka F modelu, p - poziom istotności modelu, F-LF - statystyka F braku dopasowania, p-LF - poziom istotności braku dopasowania. Dokładne objaśnienia w tekście.

w tym okresie roku gradient spadku przy wyższych opadach - w grudniu przy opadach nieco powyżej 30 mm udział C<sub>2</sub> jest już zerowy. W kolejnych miesiącach - marcu, kwietniu i maju - maksima wariancji składowej drugiej rosna do poziomu około 0,32 i przesuwają się w kierunku wyższych opadów - wynoszących około 15 mm. Miesiące letnie odróżniają się od nich przede wszystkim znacznie mniejszym gradientem spadku C<sub>2</sub> przy wysokich MSDO. Na przykład, w lipcu udział C<sub>2</sub> jest przy opadach rzędu 60 mm na dobę dwukrotnie wyższy niż w maju, mimo że maksima w obu tych miesiącach są zbliżone. Największa amplituda w tym zakresie, bo przekraczająca 300%, jest między styczniem a sierpniem dla MSDO wynoszących około 20 mm. Wszystkie równania regresji C<sub>2</sub> dla poszczególnych miesięcy względem wysokości opadów są bardzo wysoko istotne statystycznie (p < 0,00001). Mniej jest też w wypadku tego parametru istotnych statystyk *Lack Fit*.



Ryc. 106. Zależna od bezwzględnej wysokości opadu zmienność wariancji drugiej składowej modeli struktury przestrzennej kodowanych danych MSDO dla poszczególnych miesięcy zilustrowana za pomocą funkcji wielomianowych. Dla stycznia i lipca zaznaczono 95% przedział ufności funkcji



Ryc. 107. Zależna od bezwzględnej wysokości opadu zmienność zasięgu pierwszej składowej modeli struktury przestrzennej kodowanych danych MSDO dla poszczególnych miesięcy zilustrowana za pomocą funkcji wielomianowych. Dla stycznia i lipca zaznaczono 95% przedział ufności funkcji

Najbardziej zróżnicowany obraz, jeśli chodzi o jakość modeli regresyjnych, dotyczy zmienności w poszczególnych miesiącach zasięgu pierwszej składowej ( $A_1$ ), w relacji do bezwzględnej wysokości opadów (ryc. 107). Dla marca i grudnia są one bowiem wysoko istotne ( $p < 0,0001$ ), latem i w październiku żadnej zależności nie stwierdzono. W pozostałych miesiącach obraz jest mocno zmienny, generalnie jednak między listopadem a kwietniem zróżnicowanie  $A_1$  jest statystycznie potwierdzone, a w ciepłym półroczu – odrzucone. Od reguły tej wyjątkiem jest jedynie wrzesień. Abstrahując od mało wiarygodnych fragmentów wykresów funkcji przy skrajnie niskich opadach, na rycinie (ryc. 107) widoczne są dwa główne typy relacji. Pierwszy to układ paraboliczny z minimum wynoszącym około 11,5 km przy MSDO rzędu 15-30 mm i dłuższymi zasięgami  $A_1$  przy niższych i wyższych opadach. Do grupy tej należą miesiące od listopada do kwietnia. Drugi typ obejmuje okres od maja do sierpnia. Przeciętny zasięg pierwszej składowej jest bardzo mało zróżnicowany i wynosi 13-14 km. Wrzesień i październik mają charakter przejściowy.

## 7. Zróżnicowanie zależne od bezwzględnej wysokości opadu - lata

Do analizy relacji parametrów modeli struktury przestrzennej rocznych MSDO w odniesieniu do bezwzględnej wysokości opadów użyto identycznych metod (ryc. 108-111, tab. 21). Dane te są, można powiedzieć, bardziej „jednorodne” niż cały zbiór dla miesięcy, bo uzyskane modele regresji są lepiej do nich dopasowane. Ogólne prawidłowości się raczej powtarzają, choć odmienne są nieco proporcje.

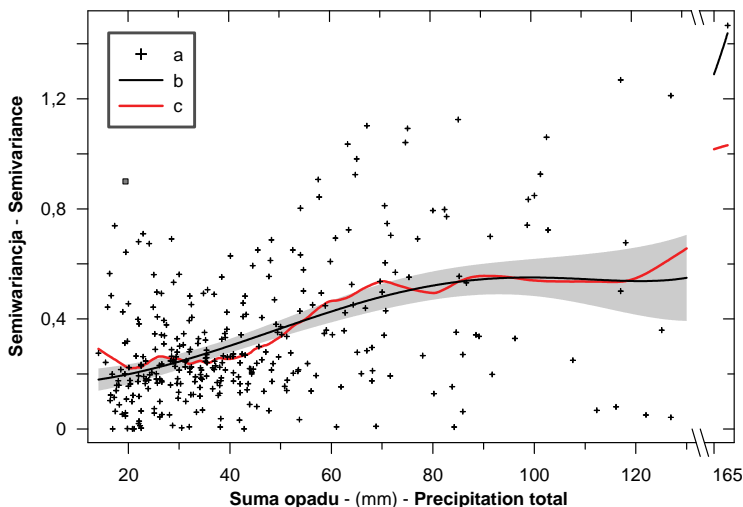
„Wyjściowy” poziom semiwariancji nuggetowej ( $C_0$ , ryc. 108) jest wyraźnie niższy niż w wypadku danych miesięcznych (ryc. 100). Wynosi bowiem jedynie 0,18, podczas gdy poprzednio było to 0,26. Przyczyną jest najprawdopodobniej brak opadów półrocza zimowego w zbiorze rocznych MSDO. Model globalny pokazuje konsekwentny wzrost wartości  $C_0$  aż do opadów rzędu 90-100 mm na dobę, a później stabilizację na poziomie około 0,54. Przy miesięcznych MSDO to *plateau* pojawiało się już przy sumach około 50 mm. Brak jest w globalnym modelu dla danych rocznych również początkowego odcinka względnie stałych wartości  $C_0$  przy niskich opadach. Zaznacza się jedynie nieco mniejszy gradient wzrostu w przedziale do 30 mm sum dobowych opadów. Na uwagę zasługuje jednak fakt sporych rozbieżności między modelami lokalnym i globalnym. Ten pierwszy wskazuje ponownie na wyraźnie większe błędy przy skrajnie niskich opadach oraz na względnie stały poziom  $C_0$  do opadów rzędu 40 mm na dobę. Również „stały” poziom wariancji nuggetowej jest w modelu Loess osiągniany przy wy-

rażnie niższej wartości MSDO, bo wynoszącej około 70 mm. Wartość ta jest istotna także z innego względu. Przy tym poziomie opadów chmura punktów na wykresie dzieli się bowiem wyraźnie na dwoje. Może to sugerować, że wysokie opady o różnej genezie odróżniają się mocno wielkością błędu pomiarowego, najprawdopodobniej ze względu na prędkość wiatru. Trzeba wyraźnie zaznaczyć, że omówione wyżej większe, niż to było w przypadku danych miesięcznych, rozbieżności między modelem lokalnym a globalnym nie wynikają bynajmniej z prostszej formy matematycznej tego drugiego.

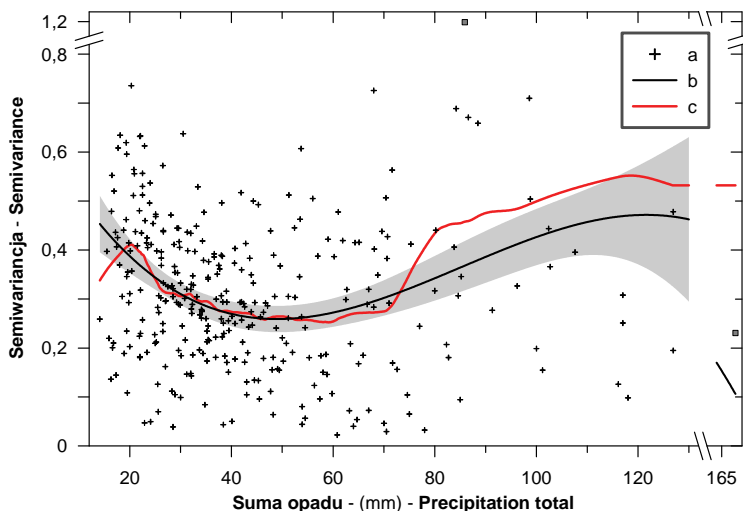
Generalny przebieg zmienności wariancji składowej pierwszej ( $C_1$ ) modeli rocznych MSDO względem wysokości opadów jest bardzo zbliżony do opisywanego dla danych miesięcznych (por. ryc. 109 i 101). Podobna jest także amplituda funkcji regresji. Najważniejsza różnica dotyczy lokalizacji jej minimum. Poprzednio mieściło się ono przy opadach rzędu 10-15 mm – przy rocznych MSDO występuje po przekroczeniu granicy około 50 mm na dobę. Zwraca uwagę także duża rozbieżność modelu lokalnego i globalnego przy opadach większych od 70 mm. Może to być jednak artefakt wynikający z małej reprezentatywności próby najwyższych sum dobowych.

Zdecydowanie bardziej „wyrazista”, niż było to przy danych miesięcznych, jest wywołana wysokością opadów zmienność wariancji składowej drugiej ( $C_2$ , ryc. 110). Jest to efektem znacznie większej amplitudy funkcji regresji, podczas gdy jej generalny układ pozostaje podobny.  $C_2$  rośnie od wartości 0,16 przy najniższych rocznych MSDO do poziomu około 0,31 przy opadach przekraczających 40 mm na dobę, a później wykładniczo maleje prawie do zera przy opadach najwyższych. Maksimum funkcji jest zatem względem danych miesięcznych przesunięte w kierunku znacznie wyższych sum dobowych. Na uwagę zasługuje fakt, że przy najniższych rocznych MSDO (do 22 mm na dobę), model lokalny sugeruje spadek, a nie wzrost udziału  $C_2$ . Danych w tym zakresie wielkości rocznych MSDO jest jednak stosunkowo mało i wykazują one tak duży rozrzut, że może być to jedynie artefakt numeryczny.

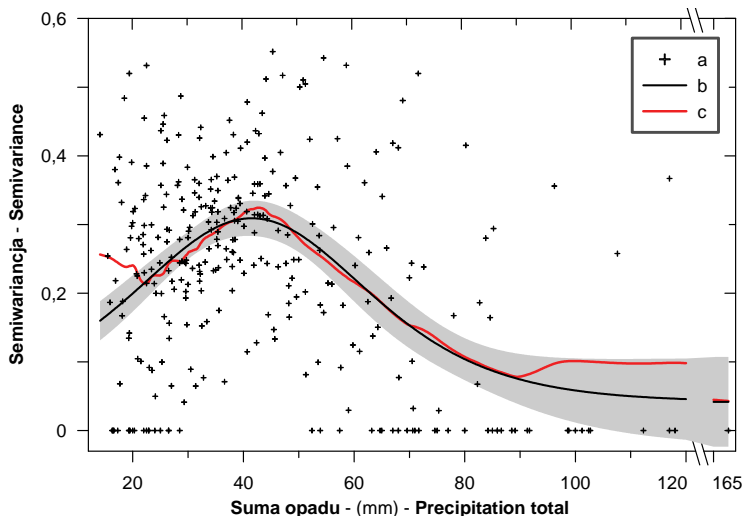
Podobnie, jak to miało miejsce przy danych miesięcznych, mało „wyrazna” jest zmienność zasięgu pierwszej składowej ( $A_1$ , ryc. 111). Największy zasięg (średnio ok. 15,6 km) mają MSDO o najniższych sumach. Maleje on wraz ze wzrostem opadów do około 12,6 km przy sumach dobowych przekraczających 50 mm. Przy wyższych MSDO funkcja zaczyna rosnąć i osiąga poziom około 15 km, kiedy opady są większe od 120 mm. Podobnie jednak, jak było to w wypadku wariancji nuggetowej od poziomu 70 mm, zaznacza się także dwudzielność chmury punktów na wykresie. Ogólny rozrzut punktów jest jednak tak duży, że może być to zupełnie przypadkowe.



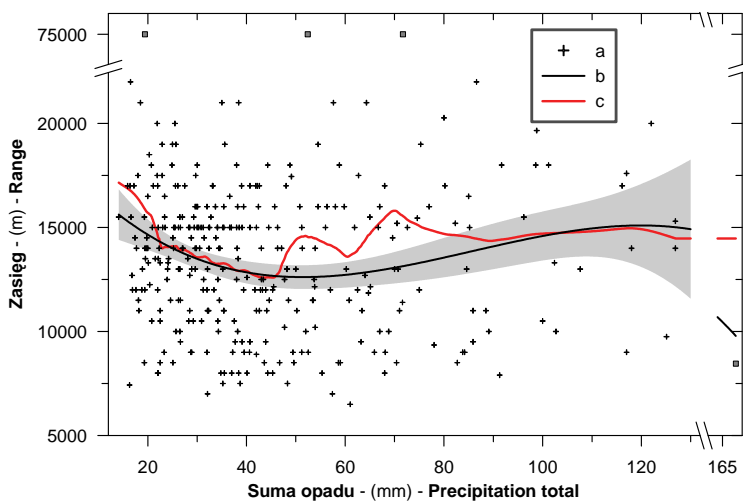
**Ryc. 108.** Zależna od bezwzględnej wysokości opadu zmienność wariancji nuggetowej modeli struktury przestrzennej rocznych kodowanych danych MSDO: a - dane, b - krzywa regresji wielomianowej z zaznaczeniem 95% przedziału ufności, c - rozkład danych wygładzony za pomocą funkcji sklepanych (Loess, B-spline rzędu 3, 20% wygładzania). Z obliczeń usunięto 1 najbardziej odstający wynik zaznaczony innym symbolem niż dane pozostałe  $\square$



**Ryc. 109.** Zależna od bezwzględnej wysokości opadu zmienność wariancji pierwszej składowej modeli struktury przestrzennej rocznych kodowanych danych MSDO: a - dane, b - krzywa regresji wielomianowej z zaznaczeniem 95% przedziału ufności, c - rozkład danych wygładzony za pomocą funkcji sklepanych (Loess, B-spline rzędu 3, 20% wygładzania). Z obliczeń usunięto 4 najbardziej odstające wyniki. Dwa z nich widoczne na wykresie zaznaczono innym symbolem niż dane pozostałe  $\square$



**Ryc. 110.** Zależna od bezwzględnej wysokości opadu zmienność wariancji pierwszej składowej modeli struktury przestrzennej rocznych kodowanych danych MSDO: a - dane, b - krzywa regresji wielomianowej z zaznaczeniem 95% przedziału ufności, c - rozkład danych wygładzony za pomocą funkcji sklepanych (Loess, B-spline rzędu 3, 20% wygładzania). Z obliczeń usunięto 3 najbardziej odstające wyniki. Żaden z nich nie został uwidoczniony na wykresie ☐



**Ryc. 111.** Zależna od bezwzględnej wysokości opadu zmienność zasięgu pierwszej składowej modeli struktury przestrzennej rocznych kodowanych danych MSDO: a - dane, b - krzywa regresji wielomianowej z zaznaczeniem 95% przedziału ufności, c - rozkład danych wygładzony za pomocą funkcji sklepanych (Loess, B-spline rzędu 3, 20% wygładzania). Z obliczeń usunięto 5 najbardziej odstających wyników. Cztery z nich widoczne na wykresie zaznaczono innym symbolem niż dane pozostałe ☐



## 8. Zmienność wieloletnia

Jak poprzednio, do oceny możliwości wystąpienia wieloletniego trendu analizowanych parametrów struktury przestrzennej MSDO użyto zmodyfikowanego testu Manna-Kendalla (Hirsch i in.1982, Libiseller i Grimvall 2002). Może on uwzględniać zróżnicowanie w danych pochodzące aż z trzech źródeł: sezonu, punktu pomiarowego i stanowiska (ang. *season, plot, site*). Przy obliczeniach testu dla miesięcznych kodowanych MSDO oprócz ujęcia cykliczności (miesiąc), dołączono hipotezę o potencjalnie również odmiennych trendach dla opadów o różnej wielkości (percentyl). Umownie przypisano temu czynnikowi rolę różnych punktów pomiarowych (*plot*).

Na podstawie ogólnego testu (tzw. *combined test*, tab. 23) dla parametrów modeli miesięcznych jedynie dla wariancji składowej pierwszej ( $C_1$ ) zidentyfikowano istotny statystycznie ( $p = 0,03$ ) trend malejący. Pozostałe parametry nie wykazują w analizowanym 25-leciu kierunkowych tendencji o charakterze liniowym. Jest to zatem wynik odmienny od uzyskanego dla danych znormalizowanych (tab. 12), gdzie wysoko istotny trend malejący ( $p = 0,001$ ) stwierdzono dla wariancji nuggetowej, a istotny ( $p = 0,026$ ) dla  $C_2$ . Być może dodatkowa zmienność wprowadzona poprzez podział na klasy wysokości opadu skutecznie „zamaskowała” te relacje.

Oprócz globalnego testu uzyskano również wyniki 624 testów cząstkowych (4 parametry  $\times$  12 miesięcy  $\times$  13 wartości progowych). Zamieszczenie ich wszystkich w postaci tabelarycznej jest oczywiście niemożliwe, ze względu na objętość. Tylko w 36 przypadkach uzyskano statystyki testu

**Tabela 23.** Ogólny wynik testu Manna-Kendalla na obecność liniowego trendu parametrów struktury przestrzennej miesięcznych kodowanych MSDO

Parametr Parameter	n n	Stat. testu Test stat.	Odch. Std. Std. Dev.	Stat. MK MK stat.	Poziom $p$ $p$ -value
$C_0$	3900	-470	1266,516	-0,3711	0,7106
$C_1$	<b>3900</b>	<b>-1864</b>	<b>862,181</b>	<b>-2,1620</b>	<b>0,0306</b>
$C_2$	3900	-98	1075,226	-0,0911	0,9274
$A_1$	3900	2333	1570,444	1,4856	0,1374

**Tabela 24.** Ogólny wynik testu Manna-Kendalla na obecność liniowego trendu parametrów struktury przestrzennej rocznych kodowanych MSDO

Parametr Parameter	n n	Stat. testu Test stat.	Odch. Std. Std. Dev.	Stat. MK MK stat.	Poziom $p$ $p$ -value
$C_0$	325	165	207,9303	0,7935	0,4275
$C_1$	325	-352	202,1781	-1,7410	0,0817
$C_2$	325	-21	208,3995	-0,1008	0,9197
$A_1$	325	59	251,3172	0,2348	0,8144

istotne na poziomie  $p < 0,05$ . Zakładając czysto losowy rozrzut wyników, można by takich wartości oczekiwać w 31 przypadkach. Jak można się było zatem spodziewać, różnica ta (36 w stosunku do 31) okazała się w teście Chi-kwadrat nieistotna. Podobne rezultaty uzyskano w odniesieniu do trzech z czterech analizowanych parametrów:  $A_1$ ,  $C_1$  i  $C_2$ . Częstość wyników istotnych na poziomie  $p < 0,05$  w odniesieniu do  $C_0$  – 18 na 156 – przekraczała o ponad 100% losową normę, a test Chi-kwadrat wykazał, że możliwość uzyskania przypadkowo takiego wyniku wynosi nie więcej niż 4 na 100. Zwraca uwagę również fakt, że w 11 z owych 18 sytuacji, kiedy trend wariancji nuggetowej jest istotny, dotyczy to skrajnych percentyli (1, 5 i 10 oraz 90, 95 i 99), a wśród nich 9 wskazuje na tendencję rosnącą. Są to jednak wszystko wyniki „balansujące” na granicy względnej statystycznej istotności i nie powinno się do nich przykładać przesadnego znaczenia. Stanowią one raczej inspirację do dalszych studiów z wykorzystaniem bardziej kompletnego zbioru danych sum dobowych opadów z dłuższego wielolecia.

Dla danych rocznych MSDO test ogólny nie wykazał obecności istotnego statystycznie na poziomie  $\alpha = 0,05$  trendu w odniesieniu do żadnego parametru (tab. 24). Proporcje wielkości statystyki M-K są dla nich jednakże podobne do tych, jaki stwierdzono względem danych miesięcznych. Najbliżej granicy istotności była zatem wariancja pierwszej składowej ( $p = 0,0817$ ).

## 9. Podsumowanie rozdziału

- Analizę zmienności struktury miesięcznych i rocznych MSDO w zależności od wysokości opadów przeprowadzono za pomocą semiwariogramów danych kodowanych. Dla każdego zbioru danych, na podstawie dystrybuanty empirycznej, wyznaczono 13 wartości progowych: 1, 5, 10, 20, ..., 90, 95 i 99 percentyl. Obliczenia wykonywano dla 15 odstępów po 2,5 km każdy. W sumie obliczono i wymodelowano 4225 semiwariogramów (325 zbiorów danych  $\times$  13 wartości progowych).

- Do modelowania semiwariancji danych kodowanych używano czterech modeli elementarnych: nuggetowego, sferycznego, wykładniczego i gaussowskiego. Składową pierwszą miesięcznych MSDO w 89,2% przypadków przedstawiano używając modelu sferycznego, w 5,7% – wykładniczego i 5,1% – gaussowskiego.

- Zastosowanie modeli elementarnych użytych do dopasowania składowej pierwszej miesięcznych MSDO wykazuje istotną regularność zarówno w odniesieniu do sezonu, jak i względnej wysokości opadu.

W okresie zimowo-wiosennym w opadach o charakterze lokalnym dominują powtarzalne zasięgi. Powierzchnie komórek opadowych są zatem względnie zbliżone. W czerwcu i lipcu spektrum ich rozmiarów ulega poszerzeniu, częściej wykazując charakter losowy.

- Niskie MSDO charakteryzują się częstszą frekwencją występowania wyjątkowo niewielkiej zmienności (modele o parabolicznym początku) na krótkich dystansach. Udział takich sum dobowych bardzo spada przy opadach powyżej 30 procentyla.

- W zbiorze modeli rocznych MSDO powyższych relacji nie stwierdzono, choć proporcje wykorzystania elementarnych struktur są zbliżone.

- Do oceny istotności zróżnicowania standaryzowanych parametrów modeli danych kodowanych wykorzystano wieloczynnikową analizę wariancji (MANOVA).

- Stwierdzono wysoką istotność różnic osobno w układzie sezonowym i względem względnej wysokości opadu. Kombinacja czynników „pora roku + względna wysokość opadu” na zmienność parametrów modeli wpływu nie miała. Zasięg pierwszej struktury ( $A_1$ ) nie wykazuje zróżnicowania sezonowego, a jedynie zmienność zależną od względnej wysokości MSDO.

- Cykl zmian sezonowych wariancji nuggetowej ( $C_0$ ) i wariancji pierwszej składowej ( $C_1$ ) modeli danych kodowanych był bardzo zbliżony do stwierdzonego dla całych, nie podzielonych na klasy zbiorów miesięcznych MSDO.

- Składowa druga ( $C_2$ ) ze względu na swój złożony charakter wykazywała bardziej specyficzny charakter zróżnicowania w trakcie roku. Różnice te wydają się być związane ze zmiennością sezonową procesów działających w największej analizowanej skali przestrzennej – określanych jako trend. Wyraźnie zaznacza się względna homogeniczność okresu zimowego. Przejściowy charakter wykazują marzec i listopad, które różnią się zarówno od grudnia i stycznia, jak i sierpnia oraz września.

- Nie stwierdzono istotnej statystycznie sezonowej zmienności zasięgu pierwszej składowej struktury modeli danych kodowanych ( $A_1$ ). Była to jedyna wyraźna niezgodność z wynikami uzyskanymi poprzednio w trakcie analizy modeli danych znormalizowanych.

- Najważniejszym wynikiem tego etapu prac było stwierdzenie charakteru zależności parametrów modeli struktury przestrzennej MSDO od względnej i bezwzględnej wysokości opadu. Oprócz zasięgu wszystkie pozostałe analizowane parametry zarówno danych miesięcznych, jak i rocznych wykazywały bardzo istotne zróżnicowanie w klasach wysokości opadu. Większe znaczenie miała tu względna wysokość opadu związana z jej usytuowaniem w ramach dystrybuanty empirycznej.

- Udział wariancji nuggetowej ( $C_0$ ) modeli danych miesięcznych był najmniejszy przy opadach najniższych, wyraźnie rosnąc od poziomu 30 procentyla.

- Udział wariancji składowej pierwszej ( $C_1$ ) był najniższy i stały przy opadach lokujących się między 40 a 80% krzywej rozkładu empirycznego. Ekstremalnie niskie i wysokie MSDO charakteryzowały się podobnym udziałem  $C_1$ .

- Wyraźnie odmienna w stosunku do względnej wysokości opadów była relacja udziału składowej drugiej ( $C_2$ ). Była ona największa przy niskich i średnich opadach, zmniejszając się wyraźnie powyżej 60 percentyla.

- Amplitudy zróżnicowania w każdym z powyższych przypadków były znacznie większe od 100%. Relacje w zbiorze modeli rocznych MSDO różniły się od omówionych powyżej tylko nieznacznie.

- Istotne statystycznie zróżnicowanie zasięgu pierwszej składowej MSDO stwierdzono jedynie w odniesieniu do względnej sumy opadu i tylko w przypadku danych miesięcznych. Jest on najkrótszy (średnio ok. 13 km) przy sumach zbliżonych do mediany, najdłuższy zaś (średnio 15,5-16,0 km) przy opadach najniższych i bardzo wysokich (1 i 5 oraz 90 i 95 percentyl). Ekstremalnie wydajne opady (99 percentyl) mają zasięg wyraźnie krótszy.

- Relacje parametrów struktury przestrzennej MSDO w odniesieniu do bezwzględnej sumy opadów badano na dwa sposoby: parametryczny i nieparametryczny. W pierwszym podejściu określano je za pomocą funkcji regresji krzywoliniowej. Ze względu jednak na bardzo duży rozrzut danych, uzyskane wyniki nie były w pełni jednoznaczne. Dlatego, dla ich weryfikacji zastosowano także, jak poprzednio, analizę wariancji w odniesieniu do klas wysokości sum opadów.

- W obu analizowanych układach, miesięcznych i rocznych, uzyskano potwierdzenie bardzo wysokiej istotności różnic międzyklasowych ( $p < < 0,000001$ ) średnich wartości wariancji poszczególnych składowych ( $C_0$ ,  $C_1$  i  $C_2$ ). W odniesieniu do zasięgu pierwszej składowej ( $A_1$ ) istotnego zróżnicowania nie stwierdzono.

- Przy sumach dobowych opadów poniżej 7 mm średni poziom wariancji nuggetowej ( $C_0$ ) jest względnie stały. Przy opadach wyższych, np. przekraczających 50 mm na dobę, następuje jej wzrost do maksimum. Największy jego gradient ma miejsce w przedziale 15-25 mm. Zmienność  $C_0$  jest maksymalna w całym zakresie wysokości MSDO.

- Poziom wariancji składowej pierwszej ( $C_1$ ) jest przy najniższych opadach wysoki. Przy nieco wyższych – w zakresie 10-15 mm – funkcja osiąga minimum. Dalej ma miejsce jej liniowy wzrost, którego gradient zmniejsza się po przekroczeniu sum dobowych wynoszących 50 mm. Maksimum funkcja regresji  $C_1$  względem sum miesięcznych MSDO osiąga przy opadach o wysokości około 110 mm. Zasugerowano, na podstawie charakter rozrzutu punktów danych, zróżnicowanie struktury przestrzennej najwyższych MSDO w zależności od ich genezy.

- Funkcja regresji  $C_2$  względem wysokości miesięcznych MSDO w klasie opadów najniższych – do 10 mm – nieco rośnie, aby od wymienionej wyżej granicy łągodnie, liniowo spadać.

- Amplituda zmian wartości regresji  $A_1$  względem wysokości miesięcznych MSDO jest niewielka. Wynosi ona bowiem jedynie 1,2 km. Maksimum

funkcji sięgające około 13,5 km występuje zarówno przy najniższych sumach opadów (do 2 mm), jak i po przekroczeniu 30 mm. Najmniejszy zasięg pierwsza składowa ma przy opadach dobowych rzędu 15 mm.

- Część obserwowanej zmienności parametrów struktury przestrzennej miesięcznych MSDO w odniesieniu do bezwzględnej wysokości opadu jest efektem zróżnicowania sezonowego. Jego rozmiar i charakterystykę jakościową zbadano używając funkcji regresji krzywoliniowych do podzbiorów MSDO dla poszczególnych miesięcy.

- Relacja wariancji nuggetowej do wysokości opadów jest silnie zróżnicowana sezonowo. Dla przedziału MSDO 5-20 mm i miesięcy zimowych typowy jest wysoki gradient przyrostu  $C_0$  wraz ze zwiększaniem się opadów. W kolejnych miesiącach wiosennych i letnich do lipca widoczna jest generalna tendencja do zmniejszania się wariancji nuggetowej i mniejszego gradientu jej przyrostu w relacji do sum opadów. Od sierpnia do listopada tendencja jest odwrotna.

- Dla większości miesięcy charakterystyczny jest monotoniczny wzrost wariancji nuggetowej wraz ze wzrostem sumy opadów. Są jednak wyjątki od tej reguły, kiedy po przekroczeniu pewnej granicy następuje spadek wielkości  $C_0$ . Dla miesięcy zimowych może być to efektem zmiennej proporcji opadów stałych i ciekłych w MSDO w zależności od ich wysokości. W półroczu letnim bardziej prawdopodobną przyczyną może być prędkość wiatru, wpływająca na błąd pomiaru, statystycznie różna przy wysokich opadach o odmiennej genezie.

- Zmienność sezonowa wariancji pierwszej składowej jest jeszcze silniej zarysowana. Wyraźnie odróżniają się trzy grupy miesięcy: styczeń – marzec, październik – grudzień i maj – kwiecień oraz czerwiec – wrzesień.

- Przy niskich MSDO (do 10-15 mm) przebieg funkcji  $C_1$  w pierwszej i drugiej grupie jest podobny: silny spadek wartości. Dalej zaznaczają się już różnice. W miesiącach zimowych minimum  $C_1$  występuje przy sumach dobowych rzędu 8-12 mm. Przy wyższych opadach wariancja pierwszej składowej bardzo szybko rośnie. Jesienią i wiosną minimum funkcji jest „głębsze” i przesunięte w stronę wyższych opadów, a gradient występującego dalej wzrostu  $C_1$  jest niższy. Trzecia wymieniona cecha nie dotyczy listopada i grudnia, które pod tym względem są podobne do miesięcy zimowych.

- Latem zróżnicowanie wielkości wariancji pierwszej składowej jest w porównaniu z pozostałą częścią roku niewielkie – i prawdopodobnie w dużym stopniu nieistotne statystycznie. Wyraźnie większy jest jednak udział  $C_1$  latem przy niskich opadach (do 15 mm) i niższy przy sumach przekraczających 30-40 mm. Jest to prawdopodobnie skutkiem zmiennej sezonowo proporcji w miesięcznych MSDO krótkodystansowych opadów konwekcyjnych.

- Wariancja drugiej składowej jest zróżnicowana sezonowo przede wszystkim przy średnich i wysokich MSDO – powyżej 10 mm na dobę.

Przebieg funkcji w 9 przypadkach na 12 jest podobny: paraboliczny wzrost do maksimum, a później liniowy lub wykładniczy spadek. Różnią je gradienty zmian i położenie maksimum. Trzy wyjątki charakteryzujące się konsekwentnym spadkiem  $C_2$  wraz z wysokością opadu to wrzesień, październik i listopad.

- Najniższe maksimum wariacji drugiej składowej występujące przy relatywnie niskich opadach (3-10 mm) jest charakterystyczne dla miesięcy zimowych (grudzień – luty). Wysoki jest w tym okresie roku także gradient spadku przy wyższych opadach. W kolejnych miesiącach – marcu, kwietniu i maju – maksima wariacji składowej drugiej rosną i przesuwają się w kierunku wyższych opadów – sięgających około 15 mm. Miesiące letnie odróżniają się od nich przede wszystkim znacznie mniejszym gradientem spadku  $C_2$  przy wysokich MSDO.

- Modele regresyjne zmienności w poszczególnych miesiącach zasięgu pierwszej składowej ( $A_1$ ), w relacji do bezwzględnej wysokości opadów są dla półrocza letniego nieistotne statystycznie. Widoczne są dwa główne typy relacji. Pierwszy z nich to układ paraboliczny z minimum wynoszącym około 11,5 km przy MSDO rzędu 15-30 mm i dłuższymi zasięgami  $A_1$  przy niższych i wyższych opadach. Do grupy tej należą miesiące od listopada do kwietnia. Drugi typ obejmuje okres od maja do sierpnia. Przeciętny zasięg pierwszej składowej jest bardzo mało zróżnicowany i wynosi 13-14 km. Wrzesień i październik mają charakter przejściowy.

- Do analizy relacji parametrów modeli struktury przestrzennej rocznych MSDO w odniesieniu do bezwzględnej wysokości opadów użyto identycznych metod jak w wypadku danych miesięcznych. Generalnie uzyskane modele regresji są lepiej dopasowane. Ogólne prawidłowości się powtarzają, choć odmienne są nieco proporcje.

- Minimalny poziom semiwariacji nuggetowej rocznych MSDO jest wyraźnie niższy niż w wypadku danych miesięcznych. Przyczyną jest najprawdopodobniej minimalny udział w analizowanych danych opadów półrocza zimowego. Charakterystyczny układ danych na wykresie daje podstawy do sugestii, że wysokie opady o różnej genezie odróżniają się mocno, ze względu na prędkość wiatru, wielkością błędu pomiarowego.

- Test hipotezy o istnieniu wieloletniego liniowego trendu parametrów struktury przestrzennej kodowanych danych MSDO dał wynik pozytywny jedynie dla wariacji składowej pierwszej ( $C_1$ ,  $p = 0,03$ , trend malejący). Pozostałe parametry nie wykazują w analizowanym 25-leciu kierunkowych tendencji o charakterze liniowym. Jest to zatem wynik nieco odmienny od uzyskanego dla danych znormalizowanych. Być może, dodatkowa zmienność wprowadzona poprzez podział na klasy wysokości opadu skutecznie „zamaskowała” niektóre relacje.

# VIII

---

## Sezonowość, wydajność i zmienność regionalna rocznych MSDO

### 1. Zmienność czasowa występowania i wydajności rocznych MSDO

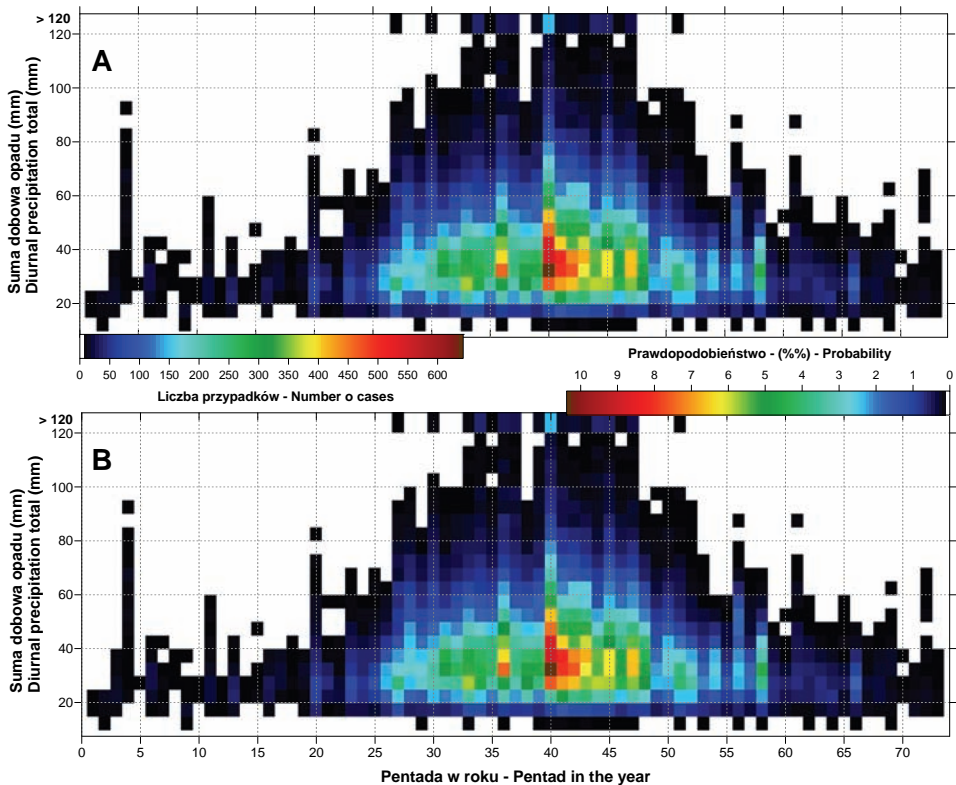
Zgromadzone w bazie danych daty wystąpienia rocznych MSDO w wieloletnim dla poszczególnych stanowisk pomiarowych stwarzają możliwość analizy nie tylko zmienności przestrzennej, ale także czasowej. Na tej podstawie podjęto w pracy problem: 1) określenia i charakterystyki zmiennego sezonowo prawdopodobieństwa rocznych MSDO o określonej wysokości dla obszaru całej Polski, 2) hipotezy dotyczącej zróżnicowania regionalnego terminów wystąpień rocznych MSDO.

Omawiając najważniejsze cechy tej części bazy danych wykorzystywanej w niniejszym opracowaniu (s. 105, ryc. 49), zwracano uwagę z jednej strony na silnie zarysowaną, choć „nieregularną” sezonowość pojawiania się na obszarze Polski rocznych MSDO, a z drugiej – na wyraźnie widoczny, nawet w 25-letniej serii obserwacyjnej, ich losowy charakter. Takie cechy analizowanych danych zadecydowały o nieparametrycznym podejściu do problemu modelowania. Wszystkie 61 940 wartości dat/sum rocznych MSDO zestawiono w jednym pliku, konwertując daty na numer dnia w roku (juliańskim). Prosta generalizacja w postaci dwuwymiarowego histogramu<sup>45</sup> dała efekty niezadowalające (ryc. 112). Z jednej strony

---

<sup>45</sup> Tabela, na podstawie której wykonano histogram została zapisana na dołączonym dysku DVD.

podział zbioru danych na przedziały pentadowe i 5 mm klasy wysokości opadów spowodował utratę części szczegółowych informacji, z drugiej nie dał ten zabieg zamierzonych efektów, tzn. całkowitej redukcji znaczenia pojedynczych ekstremalnych przypadków. Zdecydowano zatem o zastosowaniu odmiennej metodyki. Wykorzystano powszechnie stosowaną w wielu dziedzinach, np. biogeografii czy geografii społeczno-ekonomicznej, metodę oceny zagęszczenia danych punktowych tzw. *kernel density estimation* (Barbesi, Marcheselli 2002, Brunson 1995, CrimeStat III, 2004, Harada, Shimada 2006, Kharoufeh, Goulias 2002, Riva i in. 2004). Przypadki wystąpienia rocznej MSDO opadów potraktowano jako lokalizację (punkt) w dwuwymiarowym układzie współrzędnych: czas – suma opadu.

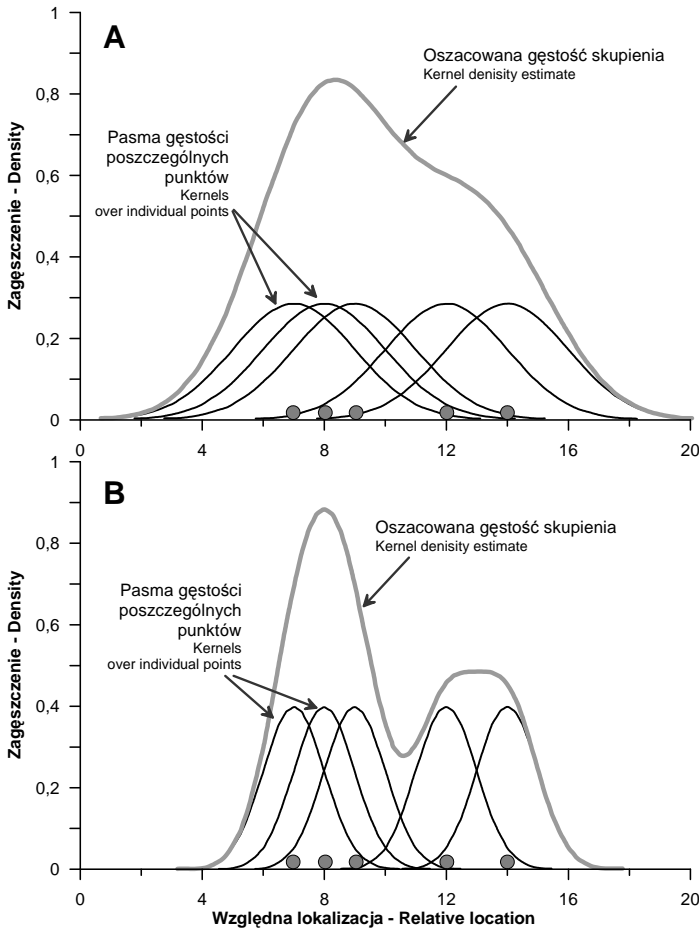


Ryc. 112. Daty wystąpienia i sumy maksymalnych opadów dobowych w Polsce w wieloleciu 1956-1980 zgeneralizowane w pentadach i przedziałach co 5 mm wysokości opadu (por. z ryc. 49). A – liczba przypadków; B – prawdopodobieństwo



## 2. Estymacja gęstości skupień

Wiąże się ona z utworzeniem symetrycznych powierzchni ponad każdym punktem, których odległość od tych punktów jest ustalona na podstawie wybranej funkcji matematycznej, a następnie zsumowanie wartości (wysokości) tych powierzchni w każdym wybranym miejscu analizowanego profilu/obszaru/objętości (ryc. 113). Technika ta została opracowana pod koniec lat pięćdziesiątych XX wieku jako alternatywna metoda szacowania gęstości histogramu (CrimeStat III, 2004).



**Ryc. 113.** Jednowymiarowy przykład interpolacji gęstości skupień (ang. *kernel density interpolation*) dla 5 punktów i zgodnym z rozkładem normalnym kształtu pasma gęstości (ang. *bandwidth*). Na rycinach A i B przedstawiono efekt interpolacji przy różnej szerokości pasma. Wykonano je według wzoru podanego w CrimeStat III (2004)

Podstawową formułę estymacji gęstości skupień w przypadku jednowymiarowym można zapisać w postaci (Kemmnade van i in. 1999):

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) = \frac{1}{n} \sum_{i=1}^n K_h(x-x_i) \quad [28]$$

gdzie  $K_h(\cdot)(t) = K(\cdot(t/h)/h)$ . W powyższym wzorze  $x$  oznacza punkt (lokalizację), dla którego określane jest zagęszczenie,  $n$  – liczbę punktów danych,  $h$  – rozmiar okna obliczeń,  $x_i$  – konkretne, poszczególne punkty danych,  $K_h(\cdot)(t)$  – rozkład (pasmo) funkcji estymacji gęstości. Rozkład ten decyduje, w jaki sposób punkty danych odległe o  $t$  od lokalizacji  $x$  wpływają na wartość gęstości w tym miejscu.

Jak to wcześniej przedstawiono na rycinie 113, „nad” każdą lokalizacją pomiaru zostaje „umieszczona” symetryczna funkcja gęstości. Określenie „symetryczna” oznacza, że jej spadek wraz ze wrastającą odległością od punktu jest taki sam w dowolnym kierunku. W przedstawionym przykładzie zastosowano funkcję gęstości rozkładu normalnego, ale wykorzystywane są także inne typy symetrycznych rozkładów. Pożądaną, syntetyczną, rozkład jest szacowany poprzez zsumowanie indywidualnych funkcji gęstości dla wszystkich lokalizacji, dając w efekcie gładką skumulowaną funkcję zagęszczenia danych. Należy zwrócić uwagę, że funkcje gęstości poszczególnych danych są sumowane w każdej wskazanej lokalizacji, a nie tylko w punktach pomiarów. Dzięki temu każdy punkt w takim samym stopniu wpływa na obliczaną powierzchnię gęstości, a uzyskana jej funkcja jest ciągła w każdym miejscu analizowanego obszaru.

Opisywana metoda wymaga jednak podjęcia, metodą prób i błędów, dwóch subiektywnych decyzji. Pierwszą z nich jest wybór funkcji gęstości (CrimeStat III 2004). Oprócz stosowanego najczęściej rozkładu normalnego, wykorzystywane są także funkcje ćwiartkowe (ang. *quartic*), stożkowe zwane trójkątnymi (ang. *conical* lub *triangular*), ujemne wykładnicze (ang. *negative exponential*) lub jednostajne (ang. *uniform*). Różnice wyników uzyskanych przy ich zastosowaniu nie są duże. Zaletą funkcji rozkładu normalnego jest możliwość ekstrapolacji gęstości. Zasięg pozostałych ograniczony jest poza lokalizację skrajnego punktu do odległości określonego przez operatora promienia pasma. Z drugiej strony, zastosowanie funkcji rozkładu normalnego czasami powoduje powstawanie artefaktów przy krawędziach (izolowane piki), zwłaszcza tam, gdzie wzdłuż krawędzi istnieją skupienia danych. Porównując wyniki z uzyskanymi przy zastosowaniu funkcji rozkładu normalnego, funkcje ćwiartkowe i jednostajne dają estymowane powierzchnie gęstości bardziej wygładzone, podczas gdy stożkowe i wykładnicze ujemne bardziej uwypuklają istniejące w obrębie niej „dolino” i „szczyty”.

Druga decyzja dotyczy szerokości (zasięgu) indywidualnej funkcji gęstości wokół każdego punktu danych. W literaturze przedmiotu parametr ten jest

określany jako szerokość pasma (ang. *bandwith*). Na rycinie 113A i B pokazano wpływ szerokości pasma na kształt uzyskanej krzywej zagęszczenia. Wyraźnie zaznacza się zależność stopnia jej wygładzenia od przyjętej szerokości pasma. Parametr szerokości pasma ma nieco inną definicję w odniesieniu do wymienionych wyżej funkcji zagęszczenia. W przypadku rozkładu normalnego jest to wartość odchylenia standardowego. Przy pozostałych funkcjach oznacza promień obszaru szukania, w zasięgu którego będzie wykonywana interpolacja. Często zalecaną opcją estymacji metodą gęstości skupień jest stosowanie zmiennej szerokości pasma, tzw. *adaptive bandwidth* (Brunsden 1995, CrimeStat III 2004). Jest ono wówczas dobierane automatycznie przez algorytm w zależności od podanej przez operatora minimalnej liczby punktów danych, które muszą się znaleźć w jego zasięgu. W efekcie pasmo jest węższe w momencie obliczeń w miejscach, gdzie zagęszczenie danych było wysokie, a szersze tam, gdzie pomiarów/obserwacji było mniej. Założeniem tego wariantu metody jest zachowanie stałej precyzji estymacji niezależnie do lokalnej ilości danych.

Przy estymacji gęstości skupień możliwe jest zastosowanie wagi różnicującej znaczenie punktów pomiarowych lub wartości uzupełniającej zmiennej pomiarowej o podobnym znaczeniu<sup>46</sup>.

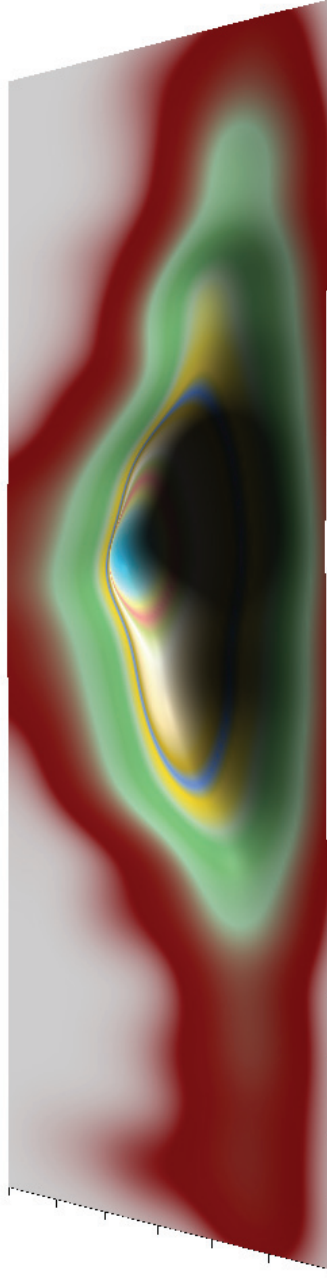
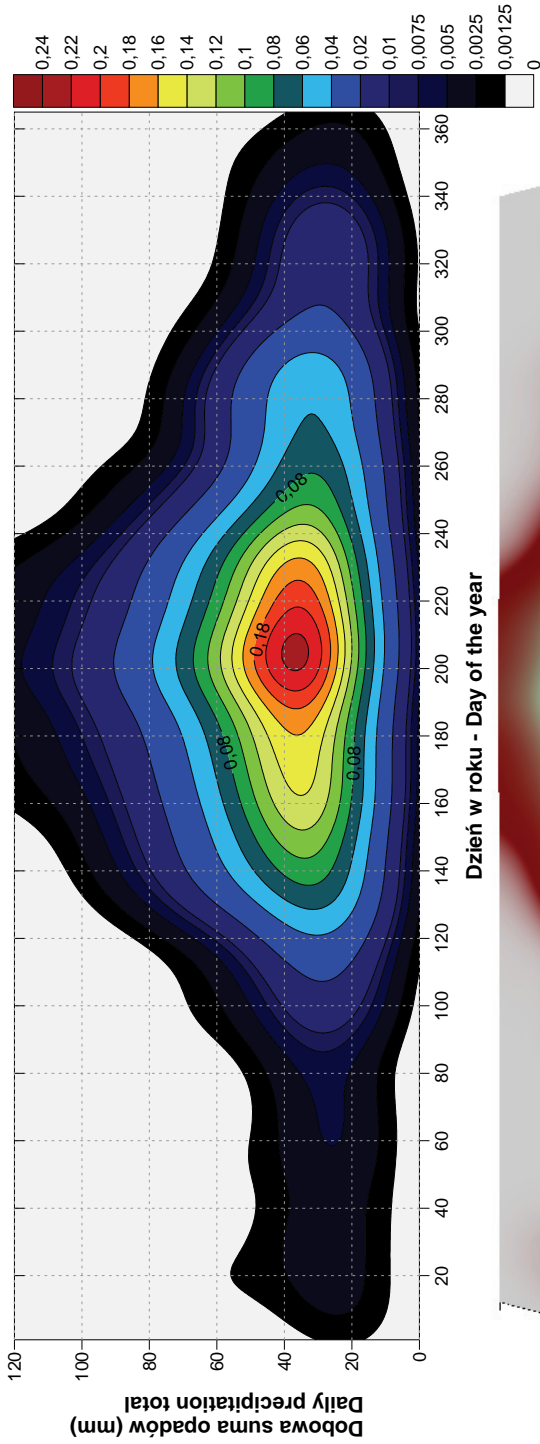
Obliczenia wykonano za pomocą programu CrimeStat III (2004), wybierając eksperymentalnie normalną funkcję gęstości z ustalonym stałym pasmem (odchyleniem standardowym) wynoszącym 10 dni/mm. Zastosowana funkcja miała postać:


$$g(x_j) = \sum \left\{ [W_i \cdot I_i] \cdot \frac{1}{h^2 \cdot 2\pi} \cdot e^{-\left[ \frac{d_{ij}^2}{2 \cdot h^2} \right]} \right\} \quad [29]$$

gdzie  $d_{ij}$  jest odległością od punktu danych do dowolnego punktu w obrębie analizowanego obszaru,  $h$  - wartością odchylenia standardowego rozkładu normalnego (szerokość pasma),  $W_i$  - wagą, natomiast  $I_i$  - natężeniem pomiaru (dodatkowe parametry, które nie były w niniejszym opracowaniu wykorzystywane). Przyjęta szerokość pasma to najniższa wartość, przy której na wynikowej powierzchni gęstości nie były widoczne ślady pojedynczych, bardzo licznych przypadków MSDO. Funkcję rozkładu normalnego zastosowano nie tylko dla tego, że dawała najbardziej „korzystny” wizualnie efekt. Jej wartość dąży bowiem do nieskończoności we wszystkich kierunkach i dlatego miała wartość niezerową w każdym punkcie analizowanego obszaru. Dzięki temu można było szacować prawdopodobieństwo wystąpienia opadów o takich charakterystykach terminu i wydajności, które w analizowanym wieloleciu nie zanotowano, a możliwych w dłuższych okresach.

---

<sup>46</sup> Na przykład, jeśli botanik analizuje zagęszczenie stanowisk jakiejś rośliny, może ich znaczenie zmodyfikować za pomocą zmiennej wieku lub wysokości poszczególnych okazów.



**Ryc. 114.** Estymowane metodą interpolacji gęstości skupień prawdopodobieństwo wystąpienia danego dnia na obszarze Polski rocznych MSDO o określonej wysokości – mapa izarytmiczna (u góry) i blokdiagram (na dole). Uwaga! Na mapie w celu dokładniejszej wizualizacji niskich prawdopodobieństw zastosowano zróżnicowaną szerokość przedziałów klasowych 

### 3. Sezonowa zmienność wydajności rocznych MSDO na terenie Polski

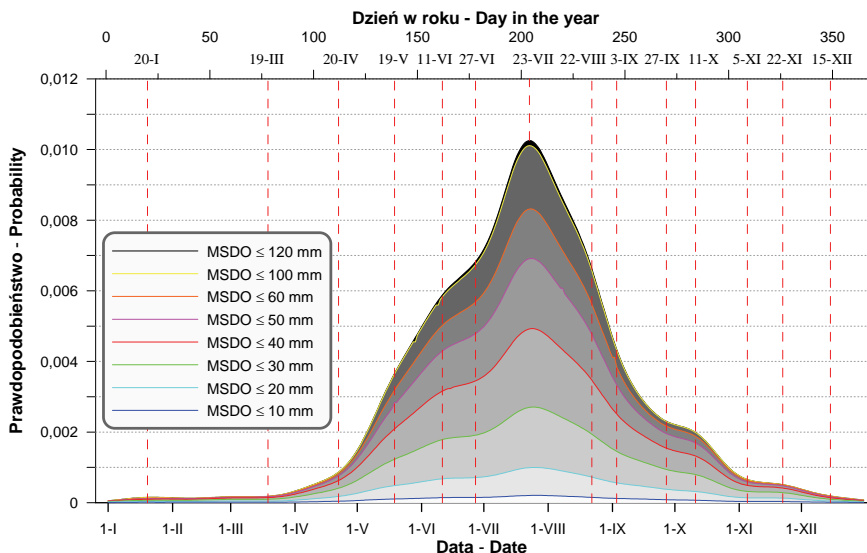
Efektom obliczeń metodą estymacji gęstości skupień są dane zestawione w tabeli 25. Zawiera ona szacowane dla każdego dnia roku prawdopodobieństwa wystąpienia na terytorium Polski rocznych MSDO o wydajności 0-120 mm, w interwale co 1 mm. Tabela ta została zamieszczona na dołączonym dysku DVD, a jej wizualizację stanowią rycyny 114-117.

Na podstawie wykresu skumulowanych prawdopodobieństw rocznych MSDO wyróżniono 14 okresów odmiennych ze względu na jego poziom i gradient zmian (ryc. 115, tab. 25). Zwrócono również uwagę na roczny przebieg prawdopodobieństwa wystąpienia MSDO o określonej wysokości (ryc. 116 i 117).

**Tabela 25.** Charakterystyka sezonów prawdopodobieństwa wystąpienia rocznych MSDO

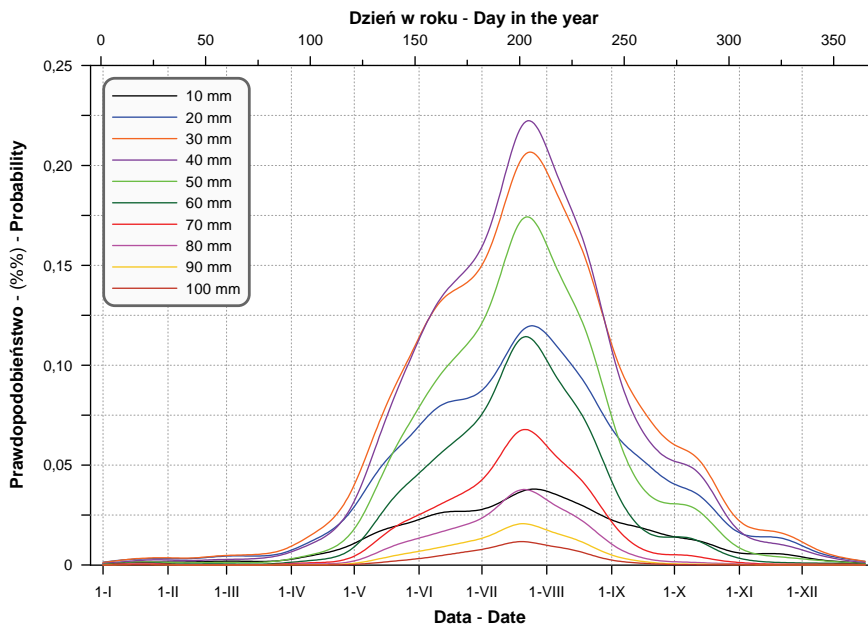
Okres Season	Początek Start	Koniec End	Dni - Days	% roku % of the year	$\Sigma P$ (%)	minP <sub>d</sub> (‰)	maxP <sub>d</sub> (‰)	średP <sub>d</sub> (‰)	gradP <sub>d</sub> (‰/d)
1	16-XII	20-I	36	9,9	0,419	0,053	0,181	0,116	-0,0005
2	21-I	19-III	58	15,9	0,876	0,127	0,182	0,151	0,0005
3	20-III	20-IV	32	8,8	1,355	0,186	0,785	0,424	0,0189
4	21-IV	19-V	29	7,9	5,935	0,824	3,661	2,047	0,0992
5	20-V	11-VI	23	6,3	11,190	3,770	5,880	4,865	0,0965
6	12-VI	27-VI	16	4,4	10,216	5,951	6,825	6,385	0,0591
7	28-VI	23-VII	26	7,1	22,536	6,902	10,256	8,668	0,1319
8	24-VII	22-VIII	30	8,2	25,926	6,630	10,244	8,642	-0,1208
9	23-VIII	3-IX	12	3,3	6,447	4,347	6,444	5,373	-0,1902
10	4-IX	27-IX	24	6,6	7,268	2,303	4,188	3,028	-0,0852
11	28-IX	11-X	14	3,8	3,004	1,991	2,275	2,146	-0,0223
12	12-X	5-XI	25	6,8	3,083	0,657	1,950	1,233	-0,0534
13	6-XI	22-XI	17	4,7	0,966	0,511	0,639	0,568	-0,0086
14	23-XI	15-XII	23	6,3	0,777	0,190	0,500	0,338	-0,0140
<b>Suma Total</b>			365	100,0	100,0				

$\Sigma P$  (%) - suma prawdopodobieństwa w sezonie w procentach, minP<sub>d</sub> (‰) - minimalne dobowe prawdopodobieństwo w sezonie w promilach, maxP<sub>d</sub> (‰) - maksymalne dobowe prawdopodobieństwo w sezonie w promilach, średP<sub>d</sub> (‰) - średnie dobowe prawdopodobieństwo w sezonie w promilach, gradP<sub>d</sub> (‰/d) - średni gradient zmian prawdopodobieństwa w sezonie w promilach na dobę.



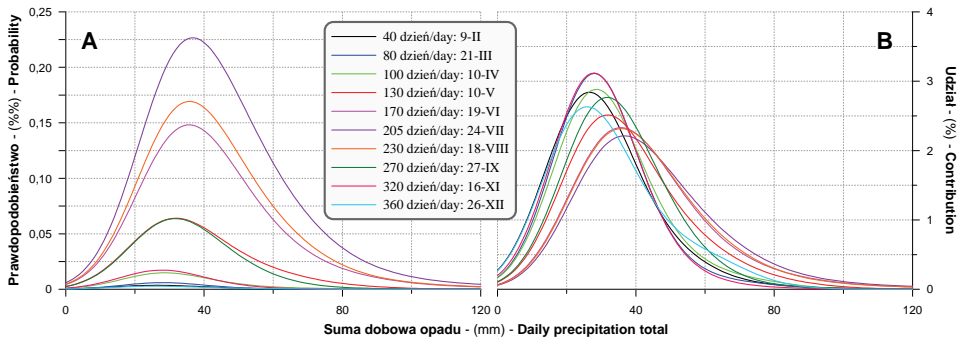
Ryc. 115. Prawdopodobieństwo wystąpienia danego dnia na obszarze Polski rocznych MSDO o określonej wysokości wraz z niższymi oraz podział roku na sezony [4]

Najbardziej znaczący jest podział roku na sezony niskiego oraz wysokiego prawdopodobieństwa wystąpienia MSDO, rozdzielone dwoma krótkimi okresami przejściowymi (ryc. 115). Od 16 grudnia do 19 marca, a więc w przez 25,8% czasu trwania roku kalendarzowego prawdopodobieństwo wystąpienia MSDO wynosi nieco poniżej 1,3%. Z drugiej strony, szansa, że największa roczna suma dobowa opadu wystąpi pomiędzy 21 kwietnia a 5 listopada (54,5% roku) wynosi aż 95,6%. Wiosenny okres przejściowy trwa od 20 marca do 20 kwietnia (8,8% roku), a prawdopodobieństwo, że w tym czasie zostanie zanotowana MSDO równa się 1,36%. Jesienią warunki przejściowe trwają nieco dłużej, bo od 6 listopada do 15 grudnia (11,0% roku). W tym okresie szansa na wystąpienie MSDO wynosi 1,74%. Inny generalny podział może obejmować sezon wzrostu prawdopodobieństwa wystąpienia MSDO trwający od 2 stycznia do 23 lipca (55,6% roku) i spadku w pozostałej części roku. Od ogólnej tendencji odbiegają dwa krótkie okresy depresji krzywej prawdopodobieństwa na przełomie stycznia i lutego (17 dni) oraz w pierwszej dekadzie marca (5 dni). Okres, kiedy szansa na wystąpienie rocznej MSDO jest największa trwa od 28 czerwca do 22 sierpnia, czyli 56 dni (15,3% roku). Możliwość, że w tym czasie zanotowany zostanie taki opad wynosi 48,5%. Krzywa prawdopodobieństwa swoje maksimum roczne osiąga 23 lipca. Wynosi ono tego dnia 1,03%. Najniższe wartości dobowe występują na przełomie roku i wynoszą 0,005-0,008%. Zróżnicowanie przekracza zatem trzy rzędy wielkości.



Ryc. 116. Profile czasowe prawdopodobieństwa wystąpienia rocznej MSDO o określonej wysokości

Ponadto, istotne prawidłowości dotyczące dynamiki prawdopodobieństwa wystąpienia MSDO w Polsce związane są ze zmiennością sezonową opadów o różnej wydajności. Na wykresie na rycinie 116 uwidacznia się to w postaci braku współkształtności, a nawet przecinanie się poszczególnych krzywych, natomiast na wykresie zamieszczonym na rycinie 117 poprzez „przesuwanie” się ich szczytów. Najbardziej znaczące są zmiany względnego udziału najniższych rocznych MSDO wynoszących poniżej 10 i 20 mm. Opady o takich sumach dominują łącznie przez trzy i pół miesiąca: od połowy grudnia do końca marca. W drugiej połowie kwietnia ich znaczenie radykalnie spada, a powrót do relacji z zimnej części roku zaczyna się od początku września. Jednakże, przez większą część roku (nawet w półroczu zimowym) dominują, jako pojedyncza klasa, opady o maksymalnych sumach dobowych rzędu 30 mm. Tylko przez nieco ponad dwa miesiące – od połowy czerwca do ostatniej dekady sierpnia – większą frekwencję występowania mają roczne MSDO w wysokości około 40 mm. Opady z tej klasy wykazują najszybszy gradient przyrostu prawdopodobieństwa szczególnie w pierwszej połowie maja i w pierwszej połowie lipca. Na uwagę zasługują także stosunkowo niewielkie, ale konsekwentne zmiany terminu, kiedy krzywe prawdopodobieństwa poszczególnych klas wysokości rocznych MSDO osiągają maksimum. Najwcześniej, bo 20 lipca szczyt ma krzywa najwyższych sum – 100 mm, a najpóźniej, 26 lipca, najniższych, rzędu 10 mm.



Ryc. 117. Profile prawdopodobieństwa wystąpienia (A) i udziału rocznej MSDO o określonej wysokości (B) dla wybranych dni w roku

Rozkład prawdopodobieństwa wystąpienia rocznych MSDO o różnej wysokości w wybranych terminach (ryc. 117) pokazuje z nieco innej strony omawiane powyżej zróżnicowanie sezonowe. Najbardziej charakterystyczne różnice dotyczą kształtu rozkładu i zmian wartości jego maksimum. Dla zimnej pory roku (grudzień – luty) charakterystyczny jest bardziej zwarty przebieg krzywej, związany głównie z niskimi prawdopodobieństwami wysokich sum dobowych. W tych miesiącach krzywa osiąga maksimum przy 27 mm opadu na dobę. Najbardziej wysmukły kształt mają rozkłady prawdopodobieństwa z 16 listopada i 21 marca, a ich szczyt odpowiada wydajności 30 mm na dobę. Od kwietnia krzywe stają się bardziej „przysadziste” (platykurtyczne), a ich maksimum przesuwają się w kierunku coraz wyższych wartości, osiągając 24 lipca 39 mm dla najbardziej prawdopodobnej sumy dobowej maksymalnego opadu rocznego.

#### 4. Zmienność przestrzenna terminu występowania rocznych MSDO

Drugim problemem, jaki zamierzano rozwiązać analizując dane terminów wystąpienia rocznych MSDO była weryfikacja hipotezy o ich potencjalnym na terenie Polski zróżnicowaniu regionalnym. Posłużono się do tego celu obliczeniami statystyk lokalnych w ruchomym oknie juliańskiego numeru dnia wystąpienia MSDO. Obliczano je dla każdego węzła stosowanej w niniejszej pracy siatki interpolacyjnej o rozdzielczości 1×1 km, wykorzystując 200 najbliższych punktów danych<sup>47</sup>. Tak dobrana wielkość próby powinna z jednej strony wyeliminować wpływ pojedynczych ekstremalnych

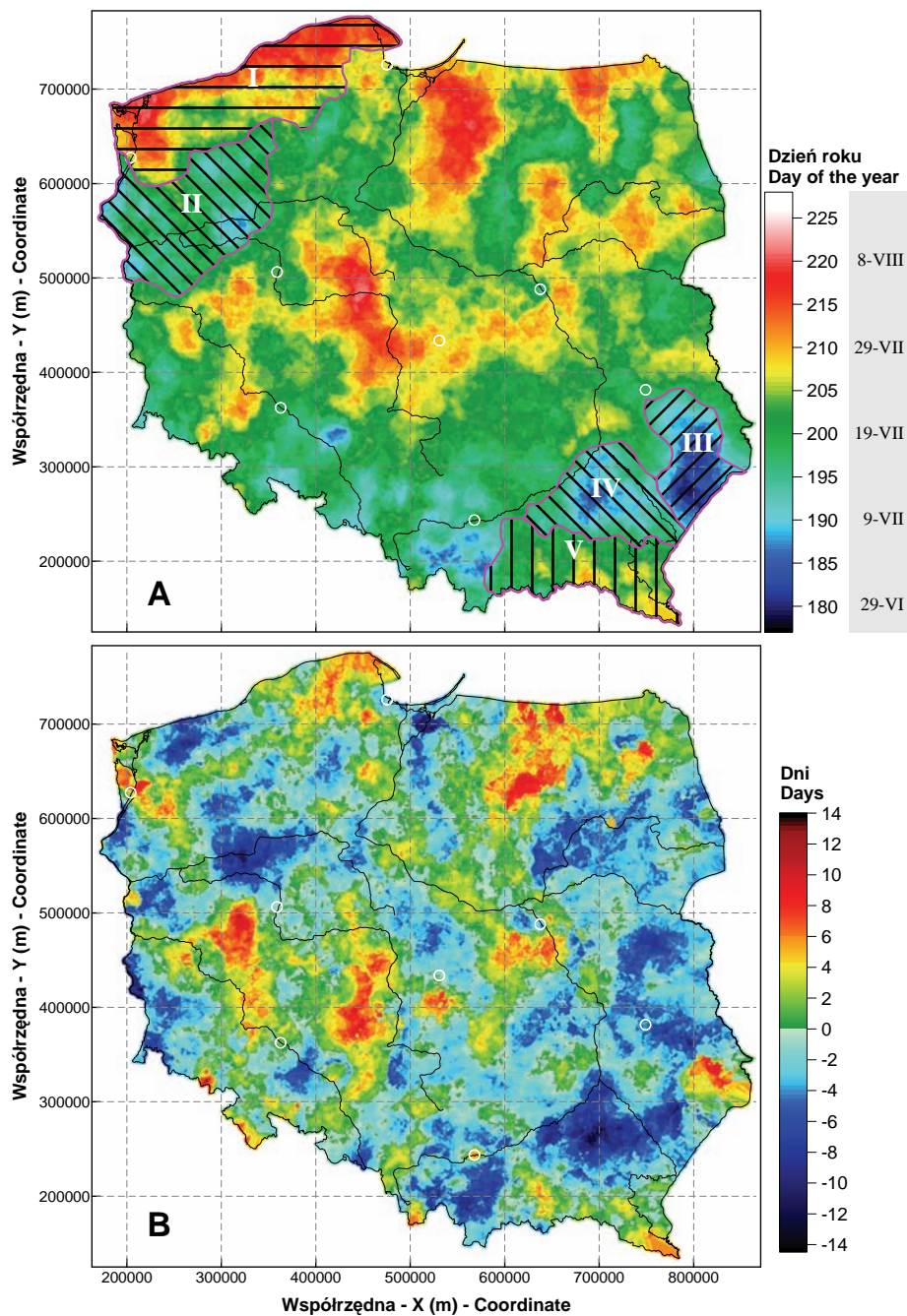
<sup>47</sup> Patrz dodatek X.6.



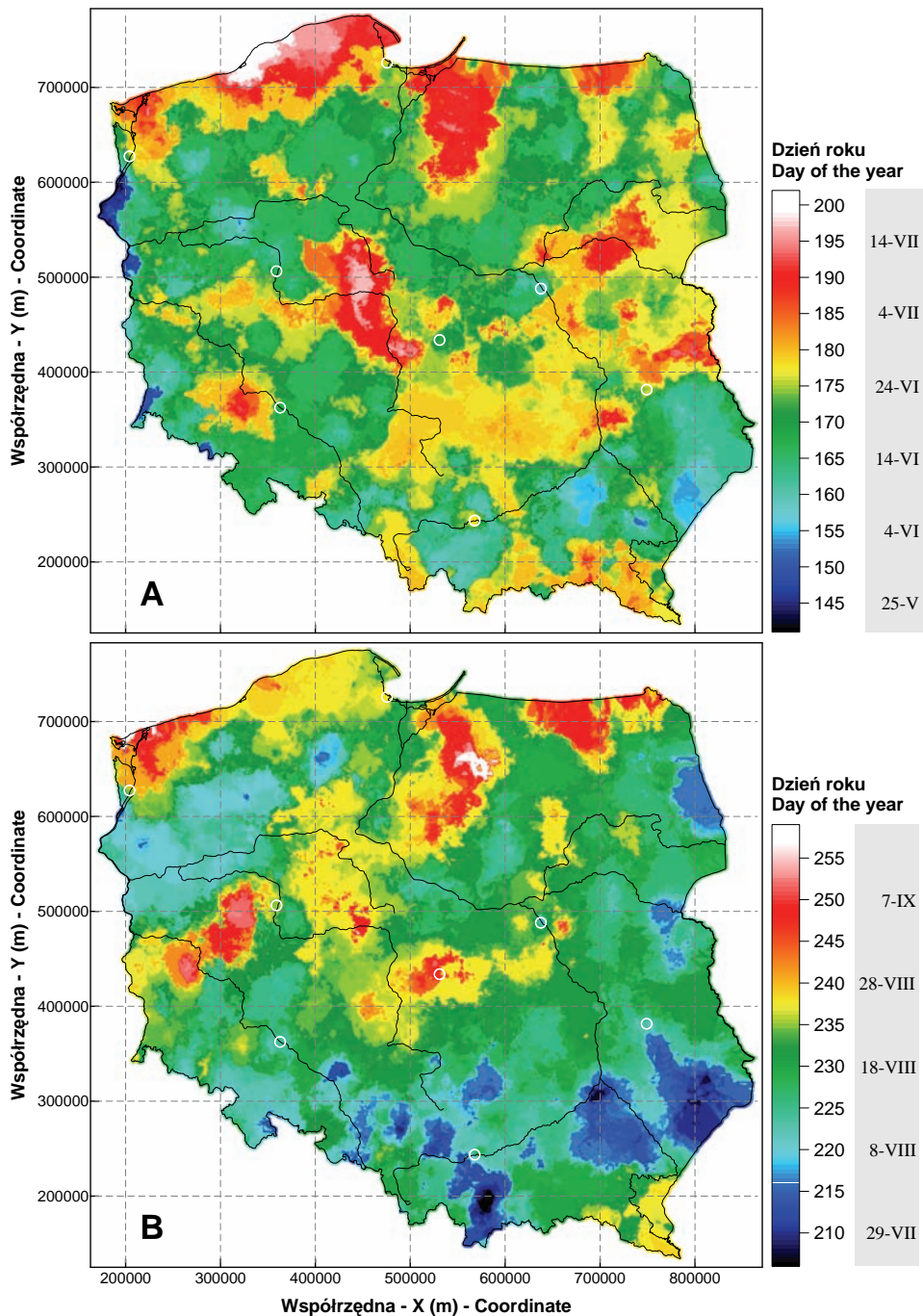
przypadków, a z drugiej zapewnić dostatecznie dobre odzwierciedlenie zmienności przestrzennej nawet w skali kilkunastu – kilkudziesięciu kilometrów. Zakładając bowiem, że znajdujące się w sąsiedztwie węzła siatki stanowiska pluwiometryczne istniały przez całe 25 lat i nie było żadnych przerw w pomiarach, wtedy ich minimalna możliwa wykorzystana w obliczeniach liczba to 8 ( $200/25 = 8$ ). W rzeczywistości, ze względu na zmiany sieci pomiarowej i luki w seriach danych poszczególnych stanowisk, do obliczeń wartości każdego węzła wykorzystywane były dane z kilkunastu do około trzydziestu punktów. Biorąc pod uwagę zróżnicowanie zagęszczenia stacji pomiarowych na terenie Polski w analizowanym wieloleciu (por. rozdz. V.3), wyznaczona liczba 200 punktów terminów rocznych MSDO znajdowała się w większości przypadków (75%) w promieniu 21-48 km od węzła (średnio 41,5 km), a przy żadnym węźle nie przekroczyła 50 km.

Przy zastosowaniu omówionych wyżej parametrów obliczono dla każdego węzła siatki interpolacyjnej następujące charakterystyki czasowe występowania rocznych MSDO: średnią i różnicę między średnią a medianą (ryc. 118), pierwszy i trzeci kwartył daty (ryc. 119) oraz jej odchylenie standardowe (ryc. 120). Stosowanie klasycznych miar statystycznych do zbioru dat pociąga za sobą ryzyko uzyskania wyników mało wiarygodnych, a w skrajnym przypadku nawet absurdalnych. Są to bowiem tak zwane dane cykliczne (ang. *circular data*), tak samo jak zbiory pomiarów kierunkowych, pojawiają się czy natężenia jakiś zjawisk w cyklach dobowych, czy tygodniowych. Obliczenie, że średnią datą dla dwóch zdarzeń jednego z 15 stycznia (15 dzień) i drugiego z 31 grudnia (365 dzień) jest 9 lipca (190 dzień) jest oczywiście skrajnie niepoprawne (prawidłowa średnia to 8 stycznia). Do danych cyklicznych stosuje się specjalnie przystosowane warianty miar statystycznych (Mardia, Jupp 2000). Niestety, program komputerowy wykorzystywany do obliczeń statystyk w ruchomym oknie takich możliwości nie posiada. W tym konkretnym przypadku nie było to wszakże konieczne potrzebne. Jeśli wyniki pomiarów skupione są bowiem w jednym sektorze cyklicznego spektrum (średnio 70% w ciągu 3 miesięcy od czerwca do sierpnia), a tak było w analizowanych danych, to rozbieżności między statystykami klasycznymi a cyklicznymi są niewielkie.

Obliczone w ten sposób średnie daty występowania MSDO (ryc. 118A) wahają się w Polsce od 27 czerwca (178 dzień roku) do 11 sierpnia (223 dzień). Przeciętny termin dla całego kraju to 22 lipca (203 dzień). Wynik ten różni się jedynie o jeden dzień w stosunku do odczytanego z krzywej prawdopodobieństwa. Węzeł o najwcześniejszej średniej dacie ma współrzędne  $X = 786\ 000$  m i  $Y = 277\ 000$  m, i położony jest w północnej części Płaskowyżu Tarnogrodzkiego (wschodnia część Kotliny Sandomierskiej). Najpóźniej roczna MSDO została oszacowana w punkcie o współrzędnych  $X = 218\ 000$  m i  $Y = 693\ 000$  m, usytuowanym w północno-wschodniej części wyspy Wolin.



Ryc. 118. Średnia (A) i różnica pomiędzy średnią a medianą (B) daty wystąpienia rocznej MSDO w wieloleciu 1956-1980. Szrafem zaznaczono zasięg prowizorycznych regionów, których dane wykorzystano do testowania istotności różnic



Ryc. 119. Pierwszy (A) i trzeci (B) kwartył daty wystąpienia rocznej MSDO w wieloleciu 1956-1980

Największy względnie zwarty obszar wczesnego występowania rocznych MSDO znajduje się w południowo-wschodniej Polsce, obejmując Kotlinę Sandomierską i sąsiadujące z nią regiony Pogórza Karpackiego, południową część Wyżyny Lubelskiej i Rostocze, a także Pogórza i część Karpat na południe i zachód od Krakowa. Mniej „wyraźnie” zaznaczają się podobne strefy dalej na zachód, w obrębie Niziny Śląskiej i Pogórza Sudeckiego oraz w północnej Polsce nad dolną Odrą, Wartą i Notecią. Przeciętne występowanie rocznych MSDO pod koniec lipca i na początku sierpnia notowane jest w kilku obszarach usytuowanych w północnej i centralnej Polsce. Najbardziej zwarty jest ten, który obejmuje Pojezierze Iławskie i sąsiadujące z nim regiony położone na wschód od dolnej Wisły, a zwłaszcza Wysoczyznę Elbląską, Równinę Warmińską i wschodnią część Żuław Wiślanych. Drugi z takich obszarów ciągnie się wzdłuż wybrzeża od Szczecina po Puck i Hel. Pas ten jest szeroki w zachodniej i północnej części Pobrzeża Szczecińskiego, wąski wzdłuż Wybrzeża Trzebiatowskiego i Równiny Białogardzkiej oraz ponownie szeroki na odcinku Wybrzeża Słowińskiego. Trzeci obszar, gdzie roczne MSDO występują względnie późno to tereny na północ i południe od środkowej Warty, w okolicach Konina. Pozostała część kraju stanowi mało zróżnicowaną „powierzchnię”, gdzie maksymalne roczne sumy dobowe notowane są przeciętnie w ciągu 2 tygodni, między 16 a 30 lipca.

Różnicę między średnią a medianą wykorzystano jako miarę rodzaju i wielkości skośności rozkładu terminów pojawiania się rocznych MSDO (ryc. 118B). Wynosi ona maksymalnie dokładnie dwa tygodnie *in plus* i *in minus*, a rozkład wartości jest prawie idealnie symetryczny. Pięćdziesiąt procent zmienności tej różnicy (przedział międzykwartylowy) mieści się jednakże w 4,5 dnia. Jakkolwiek obszary różnic dodatnich (rozkłady prawoskośne) i ujemnych (lewośkośne) są czasami dość duże i zwarte, to występowanie także mozaiki małych przeplatających się powierzchni skłania do ostrożności przy interpretacji. Część z nich bowiem to z pewnością efekt pojedynczych losowych przypadków bardzo wczesnych albo bardzo późnych rocznych MSDO. Generalnie rozkłady prawoskośne dominują na Pojezierzu Mazurskim, w północno-wschodniej części Pojezierza Pomorskiego oraz południowej i zachodniej Wielkopolsce. Rozkłady lewośkośne, świadczące o większym udziale wysokich opadów w pierwszej połowie roku, dominują we wschodniej i południowo-wschodniej Polsce, ale także w pasie wzdłuż granicy z Niemcami oraz w północnej Wielkopolsce i południowej części Pojezierza Pomorskiego.

Rozkład przestrzenny pierwszego i trzeciego kwartyla zbioru terminów pojawiania się rocznych MSDO wykazuje oczywiście dużą zbieżność z mapą wartości średnich (ryc. 119). Istnieją jednakże znaczące, warte zauważenia różnice. W obu przypadkach większy jest odstęp między skrajnymi datami. Średnie wartości dla całego kraju różniły się między sobą maksymalnie o 45 dni,

natomiast terminy, przy których przekroczone zostaje 25% całego zbioru zawierają się w przedziale 61 dni (od 21 maja do 21 lipca). Te same charakterystyki dla trzeciego kwartyła wynoszą odpowiednio 53 dni i przedział od 25 lipca do 16 września. Najwcześniejszy termin pierwszego kwartyła notowany jest nad Dolną Odrą, na południe od Szczecina, najpóźniej – na Pobrzeżu Słowińskim, w okolicach jeziora Łebsko. Trzy czwarte przypadków wystąpienia rocznych MSDO najwcześniejszy następuje na pograniczu Beskidu Żywieckiego i Gorców, najpóźniej, jak w wypadku średniej, na północno-wschodnim krańcu wyspy Wolin, w okolicach Dziwnowa. Porównanie map rozkładu pierwszego kwartyła i średniej terminów<sup>48</sup> występowania rocznych MSDO ujawnia, że generalnie różnice są odmienne po obu stronach linii dzielącej Polskę z północnego wschodu na południowy zachód. Biegnie ona od Bielska Podlaskiego, wzdłuż południowej krawędzi Wysoczyzny Rawskiej, na południe od Warszawy, do Wzgórz Ostrzeszowskich i dalej do Gór Sowich. Na południe od tej linii MSDO występują względnie później, a na północ – względnie wcześniej niż przeciętnie w całym zbiorze danych. Od reguły tej odbiega szereg izolowanych obszarów zarówno na północy, jak i na południu, a wśród nich przede wszystkim: północno-wschodnia część Pojezierza i Pobrzeża Pomorskiego, Pojezierze Elckie i obszar leżący na południe od Dolnego Bugu. Podobne porównanie dokonane między rozkładem przestrzennym średniego terminu i jego trzeciego kwartyła wskazuje na brak generalnych różnic obejmujących całe terytorium Polski. Dotyczą one tylko wybranych regionów. Najbardziej odróżnia się stosunkowo wąska strefa biegnąca od zachodniej granicy kraju, na odcinku Nysy Łużyckiej, przez północną Wielkopolskę, po dolną Wisłę w okolicach Bydgoszczy i Torunia, a nawet nieco dalej na wschód, aż do granicy z Obwodem Kalinińskim. W jej obrębie trzeci kwartył terminów rocznych MSDO wystaje relatywnie później. Podobne cechy wykazuje znacznie mniejszy obszar na południowym wschodzie kraju, obejmujący część Polesia, Wyżyny Lubelskiej i Roztocza. Natomiast *in minus* od średniej odbiegają najbardziej Nizina Podlaska i obszar na wschód od Zatoki Puckiej.

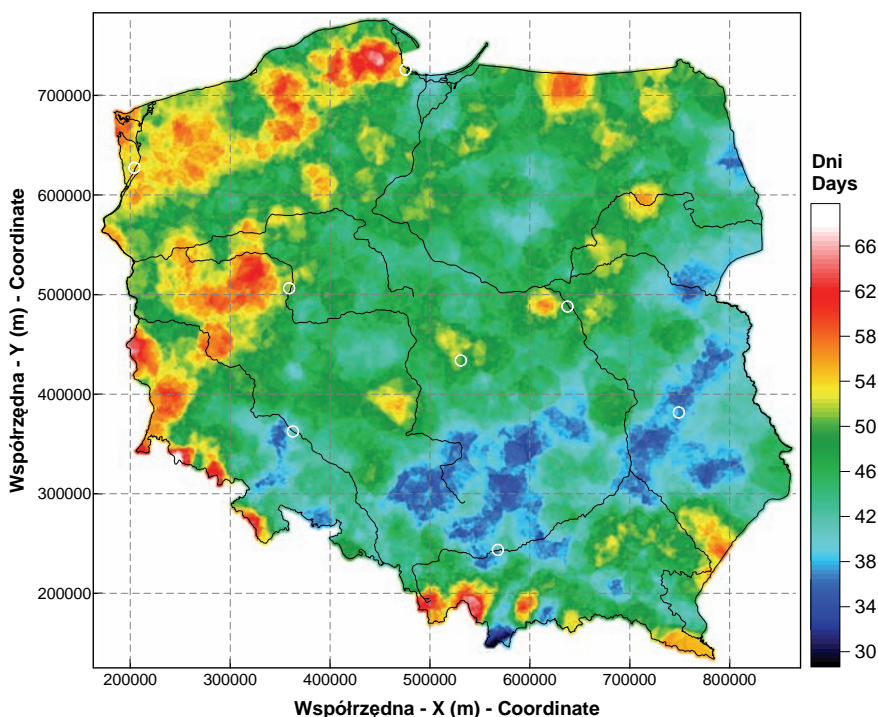
Odchylenie standardowe terminów pojawiania się rocznych MSDO dla poszczególnych węzłów siatki i przyjętej wielkości okna lokalnych statystyk wahało się od 29 do 70 dni, wynosząc przeciętnie dni 46 (ryc. 120). W jego układzie przestrzennym najbardziej znacząca jest obecność trzech głównych stref – w układzie stopniowym. Pierwsza obejmuje północno-zachodnią Polskę, a jej w miarę jednolity zasięg wyznacza linia biegnąca od Gdańska do Worka Turowskiego. Dominują na tym obszarze wysokie wartości odchylenia standardowego – generalnie powyżej 50 dni. Druga strefa rozciąga

---

<sup>48</sup> Porównania te wykonano poprzez przeskalowanie map pierwszego i trzeciego kwartyła do zakresu średniej. Map tych w publikacji nie zamieszczono.



się nieco na północ od linii: Białystok – Warszawa i Łódź w kierunku Wrocławia. Na większości obszaru znajdującego się między wymienionymi wyżej granicami odchylenia standardowe mieszczą się w granicach 40-50 dni. Trzecia, nieco porozrywana, strefa biegnąca łukiem od północno-wschodniej granicy kraju do okolic Wrocławia i opierająca się na południu o przedgórze Karpat, Karpaty i Sudety Zachodnie charakteryzuje się najniższymi wartościami – generalnie poniżej 40 dni. W górach sąsiadują ze sobą, nawet na niewielkich dystansach, obszary o skrajnie odmiennych charakterystykach zmienności terminów występowania rocznych MSDO. Najbardziej widoczne występują w Karpatach Zachodnich. Tatry, gdzie zmienność ta jest najmniejsza, i wynosi 29-37 dni, „sąsiadują” od północnego zachodu z Beskidem Żywieckim i dalej Śląskiem, a od północnego wschodu – z Pieninami i Gorcami. W regionach tych normą są z kolei wartości odchylenia standardowego przekraczające 55, a nawet 60 dni. Analizowana cecha ukazuje obecność na terenie naszego kraju generalnego gradientu zmian, ciągnącego się z północnego zachodu na południowy wschód, uwzględniającego specyfikę poszczególnych łańcuchów górskich.



Ryc. 120. Odchylenie standardowe daty wystąpienia rocznej MSDO w wieloleciu 1956-1980

## 5. Testowanie istotności różnic regionalnych terminu występowania rocznych MSDO

Obliczenia i charakterystykę zmienności terminów występowania rocznych MSDO przeprowadzono również w celu oceny potencjalnego zróżnicowania regionalnego tej cechy. Biorąc pod uwagę stałą liczebność próbki, z której obliczano statystyki lokalne (200 dat), błąd standardowy średniej określonej dla każdego węzła siatki waha się 2-5 dni. Z tego względu różnice przekraczające 5 dni można uznać za bardzo istotne statystycznie. Średnie terminy rocznych MSDO w sąsiadujących węzłach siatki różnią się przeciętnie o około 0,5 dnia (od 0 do 7), czyli poziom istotnych różnic może być osiągnięty na dystansie od jednego do kilkudziesięciu kilometrów.

Omówione powyżej wyniki wskazują na wysokie prawdopodobieństwo istotnych różnic przestrzennych w terminach pojawiania się na terenie Polski rocznych MSDO. Dotyczy to wszakże tylko pojedynczych lokalizacji – węzłów siatki interpolacyjnej. Aby stwierdzić, czy takie relacje obejmują też większe, z natury heterogeniczne obszary, posługiwano się tym przypadkiem odpowiednimi narzędziami statystyki danych cyklicznych i przeprowadzono następującą procedurę. Wydzielono na podkładzie mapy średnich dat występowania rocznych MSDO 5 prowizorycznych regionów<sup>49</sup>, zarówno sąsiadujących ze sobą, jak i „rozłącznych” geograficznie (ryc. 118). Regiony te oznaczone cyframi rzymskimi od I do V charakteryzowały się całym spektrum zmienności analizowanej cechy. Region „I” to Pobrzeża Południowo-bałtyckie od Szczecina do Gdańska i znaczna część sąsiadujących z nimi od południa pojezierzy, gdzie roczne MSDO występują późno – na większości obszaru średnio w sierpniu. Nad dolną Odrą, Wartą i Notecią wyznaczono sąsiedni region „II”, gdzie średnia data rocznego MSDO to połowa lipca. Kolejny region – „III” – zlokalizowano w obrębie Wyżyny Lubelsko-Lwowskiej i Podkarpacia, nad górnym Bugiem. Tam roczne MSDO występują najwcześniej w całej Polsce, bo pod koniec czerwca i na początku lipca. Region „IV” obejmuje przede wszystkim Kotlinę Sandomierską i część Pogórza Środkowobeskidzkiego. Tu też roczne maksymalne sumy dobowe opadów zdarzają się stosunkowo wcześnie. Ostatni wydzielony obszar „V” obejmuje Karpaty, na wschód od Krakowa, gdzie roczne MSDO najczęściej notowane są w drugiej połowie lipca. Kolejnym krokiem było wyselekcjonowanie wszystkich przypadków rocznych MSDO zarejestrowanych w analizowanym 25-leciu w obrębie wyróżnionych regionów. Do charakterystyki i porównania tych zbiorów dat posłużono się miarami i testami statystyki

---

<sup>49</sup> Granice regionów zostały zdigitalizowane arbitralnie, bez zastosowania jakichkolwiek obiektywnych kryteriów podziału.

danych cyklicznych (Fisher, 1993, Mardia, Jupp 2000). Wyniki zostały zestawione w tabelach 26 i 27 oraz na rycinie 121.

Histogramy kołowe (ryc. 121) potwierdzają, że jak ma to miejsce w odniesieniu do całego kraju, roczne MSDO mogą w wyróżnionych regionach wystąpić w każdej porze roku. Mimo to długość średniego wektora charakteryzującego stopień koncentracji danych jest stosunkowo wysoka – od 0,722 do 0,772<sup>50</sup>. Potwierdzają to także testy zgodności obserwowanych dat z rozkładem równomiernym (ang. *uniform distribution*) i rozkładem von Misesa, będącym odpowiednikiem rozkładu normalnego dla danych cyklicznych. Na omawianych wykresach (ryc. 121) pojedyncze przypadki rozległych i wydajnych sum dobowych zaznaczają się silnie w 25-letniej serii obserwacyjnej nawet po agregacji w pentady. Cecha ta jest najsłabiej widoczna w wypadku regionu „I” – nadmorskiego. Średnie daty mieszczą się przedziale od 5 (region III) do 26 lipca (region I), a standardowy błąd ich określenia wynosi 1-2 dni. Różnice między średnią „cykliczną” a tradycyjną średnią arytmetyczną tych danych wynoszą od -4,23 dnia (region II) do +0,95 dnia (region III). Generalnie jednak średnia arytmetyczna wskazywała na daty późniejsze. Również „cykliczne” odchylenie standardowe różni się konsekwentnie od jego tradycyjnego odpowiednika, dając wartości od 1,4 (region II) do 4,2 dnia (region V) niższe.

Zmienność czasów występowania MSDO jest w każdym z rozpatrywanych regionów jednak na tyle duża, że nie sposób wizualnie ocenić, czy obserwowane różnice są istotne. W celu zobjektywizowania takiego porównania posłużono się dwoma testami statystycznymi, przeznaczonymi specjalnie do danych cyklicznych (tab. 27). Jeden z nich (test F Watsona-Williamsa) służy do testowania hipotezy o zgodności kierunków średnich wektorów (w tym wypadku średnich dat), a drugi (test U<sup>2</sup> Watsona) – hipotezy o identyczności rozkładów analizowanych danych. Dały one wyniki różniące się jedynie w szczegółach. Rozkładem wartości dat wystąpienia rocznych MSDO różnią się między sobą istotnie (na poziomie  $p < 0,001$ ) wszystkie porównywane regiony oprócz „III” i „IV”. Natomiast, różnice średnich nie są istotne na poziomie  $p < 0,05$  tylko w dwóch z dziesięciu porównywanych par: „II” z „III” i „II” z „IV”. Jak można się było spodziewać, najbardziej odróżniają się regiony „I” i „V”. Wyniki przeprowadzonej analizy potwierdzają ponownie, że w obrębie granic Polski istnieje istotne zróżnicowanie regionalne czasów występowania rocznych MSDO. Przeprowadzenie formalnego podziału regionalnego w sytuacji posiadania praktycznie ciągłego obrazu zmienności analizowanej cechy uznano jednak za mało sensowne. Ewentualna modyfikacja

---

<sup>50</sup> Długość średniego wektora może zmieniać się w przedziale od 0 do 1. Wartość dolnej granicy oznacza idealnie równomierne rozproszenie danych, natomiast górnej – maksymalną koncentrację wokół średniej (w tym wypadku wszystkie przypadki rocznych MSDO tego samego dnia roku).

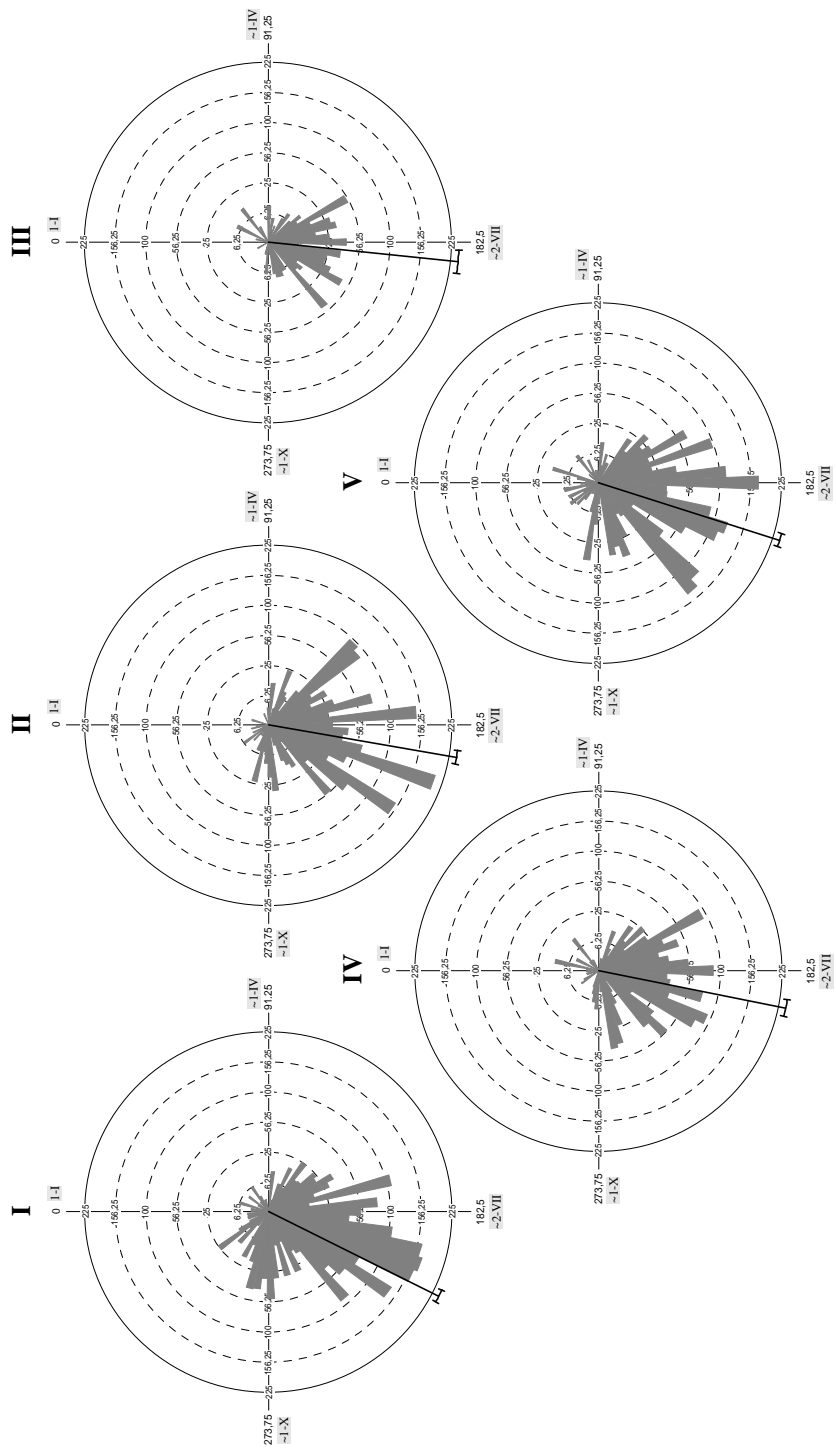


sezonowego rozkładu prawdopodobieństwa występowania rocznych MSDO o określonej wysokości (ryc. 115), uwzględniająca ich zmienność przestrzenną musi być dokonana na innej zasadzie niż tylko poprzez wydzielenie z całego zbioru dat ich regionalnych podzbiorów. Przy podkreślanym kilkakrotnie dużym wpływie pojedynczych opadów, notowanych w dziesiątkach, a nawet setkach lokalizacji, uzyskane statystyki mogłyby być silnie obciążone. Wydaje się zatem, że zamiast danych empirycznych można by wówczas wykorzystać wyniki kosymulacji czasoprzestrzennych, tym bardziej uzasadnionych, jeśli terminy występowania rocznych MSDO wykazują również autokorelację przestrzenną. Aspekt ten zostanie poruszony w następnym podrozdziale.

**Tabela 26.** Statystyki dat występowania rocznych MSDO w pięciu wyróżnionych regionach

Parametr	Parameter	Region				
		I	II	III	IV	V
Ilość danych	Number of observations	1943	1479	577	1133	1710
Średni wektor ( $\mu$ )	Mean vector ( $\mu$ )	206,818	189,941	186,01	191,488	197,868
Data	Date	26-VII	9-VII	5-VII	10-VII	17-VII
Długość średniego wektora ( $r$ )	Length of mean vector ( $r$ )	0,722	0,772	0,744	0,746	0,742
Mediana	Median	203,671	196,767	186,904	193,808	199,726
Data	Date	23-VII	16-VII	6-VII	13-VII	19-VII
Odchylenie standardowe	Circular standard deviation	46,219	41,184	44,097	43,866	44,253
Błąd standardowy średniej	Standard error of mean	1,043	1,059	1,819	1,291	1,061
95% przedział ufności $\mu$ (-/+)	95% Confidence interval (-/+ for $\mu$ )	204,774	187,865	182,444	188,957	195,788
		208,862	192,017	189,577	194,02	199,947
99% przedział ufności $\mu$ (-/+)	99% Confidence interval (-/+ for $\mu$ )	204,132	187,213	181,323	188,162	195,135
		209,504	192,67	190,698	194,815	200,601
Test $U^2$ Watsona (von Mises, $U^2$ )	Watson's $U^2$ test (von Mises, $U^2$ )	1,176	0,422	0,225	3,67	1,394
Test $U^2$ Watsona ( $p$ )	Watson's $U^2$ test ( $p$ )	< 0,005	< 0,005	< 0,005	< 0,005	< 0,005
Test $U^2$ Watsona (równom., $U^2$ )	Watson's $U^2$ test (uniform, $U^2$ )	56,776	48,471	17,651	34,591	52,162
Test $U^2$ Watsona ( $p$ )	Watson's $U^2$ test ( $p$ )	< 0,005	< 0,005	< 0,005	< 0,005	< 0,005

Zamieszczono również wyniki testów  $U^2$  Watsona zgodności obserwowanych rozkładów z rozkładem von Misesa (odpowiednik rozkładu normalnego dla danych cyklicznych) i rozkładem równomiernym.



**Ryc. 121.** Rozkład dat wystąpienia rocznej MSDO w pięciu wyróżnionych prowizorycznie regionach (patrz ryc. 118). Na diagramach zaznaczone są także średnie daty i ich 95% przedziały ufności. Skala częstotliwości jest na wszystkich wykresach jednokowa

**Tabela 27.** Górna część tabeli: macierz porównań średnich dat wystąpienia rocznych MSDO w regionach za pomocą testu F Watsona-Williamsa. Dolna połowa macierzy zawiera wartości statystyki F, a górna szacowane prawdopodobieństwo hipotezy zerowej o równości średnich. Wartości  $p$  podane jako 0 oznaczają, że były niższe niż dokładność numeryczna komputera ( $1 \times 10^{-16}$ ).

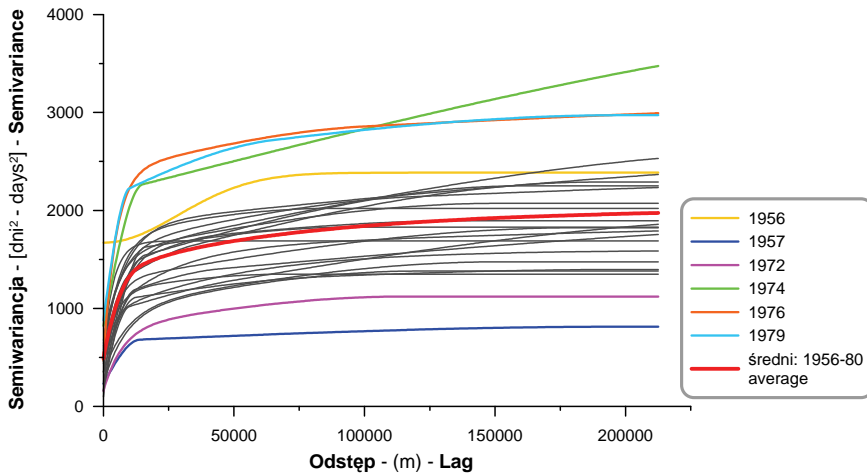
Dolna część tabeli: macierz porównań rozkładów wystąpienia rocznych MSDO w regionach za pomocą testu  $U^2$  Watsona. Dolna połowa macierzy zawiera wartości statystyki  $U$ , a górna szacowane prawdopodobieństwo hipotezy zerowej o identyczności rozkładów

		Regiony: F Watson-Williams				
		I	II	III	IV	V
Regiony	I	-----	0,000	0,000	0,000	2,99 <sup>-9</sup>
	II	123,660	-----	0,059	0,356	1,93 <sup>-7</sup>
	III	92,496	3,580	-----	0,015	2,98 <sup>-8</sup>
	IV	81,870	0,853	5,921	-----	1,64 <sup>-4</sup>
	V	35,368	27,226	30,936	14,242	-----
		Regiony: $U^2$ Watson				
		I	II	III	IV	V
Regiony	I	-----	< 0,001	< 0,001	< 0,001	< 0,001
	II	1,789	-----	< 0,001	< 0,001	< 0,001
	III	1,707	0,65	-----	0,1 > p > 0,05	< 0,001
	IV	1,913	0,754	0,167	-----	< 0,001
	V	1,813	1,332	0,545	0,495	-----

## 6. Autokorelacja przestrzenna terminów rocznych MSDO

Zastosowana w tym kontekście metodyka nie różniła się od wykorzystanej do badania struktury przestrzennej sumy opadów. Daty wystąpienia poszczególnych rocznych MSDO zamienione na numery dnia w roku posłużyły do obliczenia izotropowych semiwariancji empirycznych, do których następnie dopasowano dopuszczalne funkcje (ryc. 122)<sup>51</sup>. Tak jak w przypadku sum opadów, konieczne było przeprowadzenie przy każdym zbiorze terminów rocznych MSDO interaktywnego maskowania danych odstających – pojedynczych anomalnych przypadków (por. dodatek X.5). Ich liczba wahała się od 9 do 37 (średnio 19), a udział od 0,4 do 1,6% (średnio 0,8%), natomiast rozkład przestrzenny na terenie kraju był znacznie bardziej równomierny niż przy sumach (por. dodatek X.5), choć również dominowały obszary górskie.

<sup>51</sup> Tabelę z zestawieniem parametrów wszystkich modeli zamieszczono na załączonym DVD.



Ryc. 122. Bez kierunkowe (izotropowe) modele semiwariogramów terminów występowania rocznych MSDO. Grubą czerwoną linią zaznaczono średni semiwariogram dla 25-lecia 1956-1980. Kolorem wyróżniono kilka przypadków ekstremalnych [4]

Spodziewano się, że uzyskane modele mogą być pod względem typu, liczby i zasięgu składowych zbliżone do tych, które opracowano na podstawie danych sumy opadów (por. ryc. 53). Charakterystyka geometryczna powierzchni objętych poszczególnymi, występującymi w różnych terminach, epizodami najwyższych rocznych opadów miała przecież znaczący wpływ na globalny model struktury przestrzennej. Okazało się jednak, że różnice są bardzo znaczące.

Na rycinie 122 wyraźnie zaznacza się zróżnicowanie prawidłowości autokorelacji przestrzennej terminów występowania rocznych MSDO. Przede wszystkim, uwagę zwracają pojedyncze krzywe wyraźnie odstające od głównej wiązki. Z jednej strony są to lata 1957 i 1972 o bardzo małym zróżnicowaniu czasowym rocznych MSDO (odchylenie standardowe odpowiednio 31,3 oraz 35,0 dnia<sup>52</sup>), a z drugiej lata 1974, 1976 i 1979 o dużej zmienności terminów (SD odpowiednio 63,0, 56,5 oraz 55,6 dnia). Rok 1956 odróżniał się nie tyle ze względu na poziom zróżnicowania terminów występowania rocznych MSDO, ale przede wszystkim na całkowicie odmienny model ich autokorelacji. O ile bowiem we wszystkich pozostałych przypadkach używano kombinacji funkcji nuggetowej, wykładniczych i sferycznych, to dla danych z tego roku konieczne było użycie modelu gaussowskiego. Przypuszczalnie jest to tylko artefakt związany z szybkimi w tym czasie zmianami sieci pomiarowej, a nie odbicie rzeczywistej anomalii przyrodni-

<sup>52</sup> Statystyka klasyczna.

czej. Poza tym, główna wiązka krzywych modeli wykazuje wyraźną dwudzielność, widoczną szczególnie na dystansach rzędu 10-70 km. Do „dolnej” grupy należą lata 1958, 1959, 1960, 1965, 1969, 1975, 1977, 1978 i 1980, zaś do „górnej”: 1961, 1962, 1963, 1964, 1966, 1967, 1968, 1970, 1971, 1973. Przy odległościach większych niż 70 km trzy krzywe z wiązki górnej (lata 1961, 1964 i 1973) przebiegają bardziej „płasko” niż pozostałe, krzyżując się z dolnymi. Przeprowadzona powyżej ogólna charakterystyka zbioru modeli struktury przestrzennej terminów rocznych MSDO pokazuje jeden z aspektów ich odmienności względem tych, które opisywały prawidłowości rozkładu przestrzennego wysokości opadów (por. rozdz. VI.3 i ryc. 53). Lata anomalne w jednym kontekście są jak najbardziej typowe w drugim i na odwrót. Poza tym, modele semiwariancji sum opadów wykazywały daleko większą jednorodność – odchylenia od przeciętnej były znacznie mniejsze.

Biorąc również pod uwagę liczbę i charakter składowych budujących modele, zauważalne są istotne różnice. Analizując sumy opadów, stwierdzono występowanie trzech typów, to jest 2, 4 i 5 (por. podrozdz. VI.4), z których ten pierwszy dominował (odpowiednio 11, 5 i 9 przypadków). Tym razem wystąpiły także typy 1 i 6, czyli „nugget + 1 składowa + trend” oraz „nugget + 1 składowa” (odpowiednio 2 i 3 przypadki), a najliczniej, bo w 11 przypadkach stwierdzono typ 4, czyli „nugget + 2 składowe”. Prawdopodobnie różnice te są efektem zredukowanej znacznie roli składowej długodystansowej w zróżnicowaniu przestrzennym terminów rocznych MSDO (typ 2 i 4 różnią się jej obecnością lub brakiem). Można by na podstawie tego sądzić, że jej wysoki udział w sumach MSDO był do pewnego stopnia pozorny – „łączone” były bowiem zbliżone wysokością opady pochodzące z różnych epizodów i o różnej genezie.

Zasięg pierwszej składowej waha się od 9 do 28 km i średnio wynosi 15,8 km. Wartości te są bardzo zbliżone do uzyskanych dla pierwszej składowej modeli sum opadów (9-30 km, 17,5 km, por. rozdz. VI). Podobieństwo dotyczy jednakże jedynie tych ogólnych statystyk, a nie wartości w kolejnych latach. Współczynnik korelacji liniowej pomiędzy nimi wynosi bowiem zaledwie 0,16. To samo można powiedzieć o zasięgach drugiej i trzeciej składowej, odpowiednio: 102,8 oraz 105,3 km oraz 180,0 i 177,5 km. W wypadku tych parametrów oraz składowej trendu, dodatkowa różnica wynika z faktu, że nie były one notowane każdego roku i często występuje ich niesynchroniczność. Oznacza to, że na przykład składowa występująca w modelu rocznych MSDO nie pojawia się na modelu terminów ich występowania lub odwrotnie.

Wariancja nuggetowa ( $C_0$ ) modeli terminów występowania rocznych MSDO wynosiła średnio 21 dni<sup>53</sup> (9,8-40,9) i stanowiła 23,6% wariancji pró-

---

<sup>53</sup> Dla uproszczenia wariancje składowych podawane są w tym przypadku w oryginalnej skali pomiarowej, czyli po operacji pierwiastkowania.

by. Te same wskaźniki dla kolejnych składowych ( $C_1$ ,  $C_2$  i trend) wynosiły 28,8, 16,0 i 11,5 dni oraz 45,1, 17,3 i 14,1%. Całkowita zmienność przestrzenna jest zatem w wypadku tego parametru głównie kształtowana występowaniem opadów o krótkim zasięgu (z pojedynczych komórek konwekcyjnych) oraz czynnikami losowymi wpływającymi na ich przemieszczanie się i wewnętrzną strukturę.

Stwierdzone i opisane powyżej charakterystyki autokorelacji przestrzennej terminów występowania rocznych MSDO potwierdzają sugerowaną w poprzednim podrozdziale możliwość wykonywania czasoprzestrzennych kosymulacji i tworzenia wiarygodnych, praktycznie użytecznych, tablic prawdopodobieństwa czasu i wysokości maksymalnych rocznych sum dobowych opadów nawet w skali pojedynczych zlewni rzecznych, mezoregionów, czy obszarów zurbanizowanych.

## 7. Podsumowanie rozdziału

- Do oceny prawdopodobieństwa, na terenie Polski, terminu i wysokości rocznych MSDO użyto metody interpolacji gęstości skupień (ang. *kernel density estimation*). Użyty wariant obliczeń umożliwił szacowanie prawdopodobieństwa sumy opadu w przedziale 0-120 mm co 1 mm, w każdym dniu roku, a więc także takich przypadków, które nie zanotowano w analizowanym wieloleciu (interpolacja i ekstrapolacja).

- Na podstawie uzyskanych wyników wydzielono w trakcie roku 14 okresów o zmiennym poziomie i gradiencie prawdopodobieństwa wystąpienia rocznych MSDO w różnych klasach wysokości. Okresy te trwają 12-58 dni.

- Jakkolwiek na obszarze naszego kraju roczna MSDO może potencjalnie wystąpić każdego dnia, to prawdopodobieństwo jej wystąpienia między 16 grudnia a 19 marca jest skrajnie niskie (tylko 1,3%), natomiast że taki fakt nastąpi między 21 kwietnia a 5 listopada wynosi aż 95,6%. Okres, kiedy szansa na wystąpienie rocznej MSDO jest największa trwa od 28 czerwca do 22 sierpnia. Skumulowane prawdopodobieństwo tego zdarzenia sięga 48,5%.

- Krzywa prawdopodobieństwa swoje maksimum roczne osiąga 23 lipca. Wynosi ono tego dnia 1,03%. Najniższe wartości dobowe występują na przełomie roku i wynoszą 0,005-0,008%. Zróżnicowanie przekracza zatem trzy rzędy wielkości.

- Od połowy grudnia do końca marca dominują roczne MSDO mniejsze od 20 mm. Od połowy czerwca do ostatniej dekady sierpnia najczęściej występują roczne MSDO o wysokości około 40 mm i więcej. W pozostałej części roku klasą dominującą są opady w przedziale 20-40 mm. Maksimum praw-

dopodobieństwa wystąpienia rocznych MSDO z poszczególnych klas wysokości zawiera się w krótkim okresie między 20 a 26 lipca.

- Najbardziej prawdopodobna wysokość rocznego MSDO w miesiącach zimowych wynosi 27 mm; 24 lipca wartość ta jest najwyższa i osiąga 39 mm w ciągu doby.

- Do oceny zmienności przestrzennej terminów występowania rocznych MSDO użyto obliczeń statystyk lokalnych ich daty (numeru dnia w roku juliańskim).

- Średnie daty występowania MSDO wahają się w Polsce od 27 czerwca (178 dzień roku) do 11 sierpnia (223 dzień), z przeciętnym terminem dla całego kraju 22 lipca (203 dzień). Wynik ten różni się jedynie o jeden dzień względem odczytanego z krzywej prawdopodobieństwa.

- Największy względnie zwarty obszar wczesnego występowania rocznych maksymalnych sum dobowych znajduje się w południowo-wschodniej Polsce, nie obejmując Karpat Wschodnich, oraz nad dolną Wartą, Notecią i Odrą. Najpóźniej, bo pod koniec lipca i w pierwszej połowie sierpnia, największe roczne sumy dobowe notowane są najczęściej na wybrzeżu oraz nad środkową Wartą i Notecią.

- Odchylenie standardowe terminów pojawiania się rocznych MSDO wahało się od 29 do 70 dni, wynosząc przeciętnie dni 46. Cecha ta wykazuje generalną tendencję, modyfikowaną regionalnie przez rzeźbę terenu, do spadku z północnego zachodu na południowy wschód (gradient oceaniczny?).

- Używając testów przeznaczonych do danych cyklicznych, wykazano wysoką istotność statystyczną różnic regionalnych w terminach występowania rocznych MSDO. Skupienie terminów w jednym sektorze spektrum spowodowało, że odchylenia statystyk klasycznych od cyklicznych były wprawdzie niewielkie, lecz miały charakter systematyczny.

- Stwierdzono, że dane terminów występowania rocznych MSDO wykazują również wyraźną autokorelację przestrzenną, która może być w sposób standardowy modelowana.

- Przeciętne zasięgi składowych modeli struktury przestrzennej terminów MSDO są zbliżone do stwierdzonych dla ich wysokości (15,8, 102,8 i 180 km). W przypadkach jednak poszczególnych zbiorów danych nie ma między nimi wyraźnych zależności. Dotyczy to także liczby i typu składowych tworzących model. Prawdopodobnie różnice te są efektem zredukowanej znacznie roli składowej długodystansowej.

- Całkowita zmienność przestrzenna terminu występowania rocznych MSDO jest głównie, bo w około 69%, kształtowana przez opady o krótkim zasięgu (z pojedynczych komórek konwekcyjnych) oraz czynniki losowe wpływające na ich przemieszczanie się i wewnętrzną strukturę.

- Zaproponowano wykorzystanie kosymulacji czasoprzestrzennych do sporządzenia regionalnych tablic prawdopodobieństwa terminu i wysokości rocznych MSDO.

# IX

---

## Podsumowanie i wnioski

Zestawienie szczegółowych wyników pracy zamieszczono po każdym rozdziale analitycznym (podrozdz. V.9, VI.10, VII.9 i VIII.7). W tym miejscu uwaga zostanie skupiona na bardziej ogólnych refleksjach zarówno natury metodycznej, jak i merytorycznej.

Na podkreślenie zasługuje przede wszystkim najbardziej ogólny wniosek wypływający z przeprowadzonych prac: dane maksymalnych sum dobowych opadów, zarówno miesięcznych jak i rocznych, wykazują na terenie Polski klarowne prawidłowości zróżnicowania przestrzennego. Mogą one być matematycznie modelowane i służyć zarówno do celów lepszego zrozumienia mechanizmów kształtujących pole opadów, jak i do tworzenia prognoz dla potrzeb praktyki. Kiedy zainteresowano się tym tematem, efekt taki wcale nie był wówczas oczywisty. Wręcz przeciwnie, nieciągłość przestrzenna opadów w skali jednej doby oraz asynchroniczność danych MSDO rodziły obawy, że relacje te będą na tyle niewyraźne, iż nie da się zidentyfikować żadnych istotnych prawidłowości. W dotychczas opublikowanych pracach odnośnie struktury przestrzennej opadów tego typu danych jeszcze nie analizowano.

Pole MSDO na obszarze Polski jest najczęściej efektem sumowania się skutków działania trzech typów procesów operujących w różnych skalach przestrzennych: lokalnej (< 10-20 km), regionalnej (50-150 km) i ponadregionalnej (> 200 km). Wyróżnione skale przestrzenne wiążą się prawdopodobnie z konwekcyjną/orograficzną, synoptyczną (frontalną) i „klimatologiczną” genezą wysokich opadów (fale planetarne?). Ich udział jest bardzo zmienny. Generalnie dominują jednak wysokie sumy dobowe opadów o rozciągłości przestrzennej od 50 do 150 km związanych ze zjawiskami mezoskalowymi i migracją frontów atmosferycznych (35-38%). Obszary wysokich opadów w skali lokalnej mają częściej zasięgi zmieniające się w przestrzeni losowo, szczególnie w cieplej porze roku. W skalach ponadlokalnych dominują struktury o rozmiarach powtarzalnych.

Również wykazanie, że mimo dużej, typowej dla klimatu Polski zmienności, parametry charakteryzujące strukturę przestrzenną MSDO wykazują



wyraźną cykliczność sezonową oraz zróżnicowanie zależne od wysokości opadów, może być uznane za ważne osiągnięcie. Potwierdzono bowiem nie tyle dość oczywiste hipotezy, ale przede wszystkim dostarczono informacji ilościowych umożliwiających bliższą charakterystykę typu i skali owych prawidłowości. Zmienność sezonowa udziału składnika losowego (wariancji nuggetowej) jest na przykład bardzo zbliżona do średniego wieloletniego przebiegu współczynników korygujących błędy standardowych pomiarów opadów. Ich źródłem jest przede wszystkim turbulencja spowodowana wiatrem, występowanie opadów w postaci stałej, ale także straty związane z parowaniem i zwilżaniem. Stwierdzono również, że bezwzględna wielkość  $C_0$  jest uzależniona także od zmienności pola MSDO, zarówno globalnej, jak i lokalnej. Również względny udział wariancji pierwszej składowej wykazuje wysoce istotną zmienność sezonową z maksimum w czerwcu i maju oraz minimum w październiku. Taki cykl roczny odzwierciedla prawdopodobnie zmienny udział opadów konwekcyjnych – maksymalny w miesiącach o największych kontrastach termicznych. Charakterystyki struktury przestrzennej MSDO są bardziej uzależnione od względnej, związanej z usytuowaniem w ramach dystrybuanty empirycznej, niż bezwzględnej sumy opadów.

Analizowane wielolecie było zbyt krótkie, aby móc ocenić w sposób wiarygodny długotrwale tendencje związane z ewolucją klimatu. Stwierdzone jednak zmiany elementów struktury przestrzennej związane głównie ze zjawiskami zachodzącymi w skali lokalnej wskazują, że hipoteza postawiona na początku niniejszej rozprawy (por. rozdz. I), a dotycząca potencjalnego wpływu zmian struktury przestrzennej ekstremalnych opadów na rejestrowaną punktowo ich częstość, stwarza obiecującą perspektywę badawczą i powinna być dalej weryfikowana. Spadek lokalnego zróżnicowania wysokich opadów w okresie zimowym wymaga głębszej refleksji nad jego genezą w kontekście innych źródeł danych.

Na podstawie estymacji gęstości skupień wydzielono 14 okresów o zmiennym poziomie i gradientie prawdopodobieństwa wystąpienia rocznych MSDO w różnych klasach wysokości. Okresy te trwają od 12 do 58 dni. Jakkolwiek na obszarze naszego kraju roczna MSDO może potencjalnie wystąpić każdego dnia, to prawdopodobieństwo jej wystąpienia jest między 16 grudnia a 19 marca skrajnie niskie. Z drugiej strony możliwość, że taki fakt nastąpi między 21 kwietnia a 5 listopada wynosi aż 95,6%. Okres, kiedy szansa na wystąpienie rocznej MSDO jest największa trwa od 28 czerwca do 22 sierpnia. Krzywa prawdopodobieństwa swoje maksimum roczne osiąga 23 lipca. Od połowy grudnia do końca marca dominują roczne MSDO mniejsze od 20 mm. Od połowy czerwca do ostatniej dekady sierpnia najczęściej występują roczne MSDO o wysokości około 40 mm i więcej. W pozostałej części roku klasą dominującą są opady w przedziale 20-40 mm. Maksimum prawdopodobieństwa wystąpienia rocznych MSDO z poszcze-

gólnych klas wysokości zawiera się w krótkim okresie między 20 a 26 lipca. Średnie daty występowania MSDO wahają się w Polsce od 27 czerwca (178 dzień roku) do 11 sierpnia (223 dzień). Zmienność terminów występowania rocznych MSDO, mimo losowego charakteru najwyższych opadów, okazała się wyraźnie zróżnicowana regionalnie. Stanowi ona, jak się wydaje, efekt czynników działających w różnych skalach, a jednym z ważniejszych jest ogólny gradient oceaniczno-kontynentalny pomiędzy Europą Zachodnią i Wschodnią, modyfikowany przez bezpośrednie oddziaływanie Bałtyku oraz topografii terenu. Stwierdzono, że dane terminów występowania rocznych MSDO wykazują również wyraźną autokorelację przestrzenną, która może być modelowana w sposób standardowy. Przeciętne zasięgi składowych modeli struktury przestrzennej terminów MSDO są zbliżone do tych, jakie stwierdzono dla ich wysokości (15,8, 102,8 i 180 km). Całkowita zmienność przestrzenna terminu występowania rocznych MSDO jest głównie, bo w około 69%, kształtowana przez opady o krótkim zasięgu (z pojedynczych komórek konwekcyjnych) i czynniki losowe wpływające na ich wewnętrzną strukturę oraz przemieszczanie.

Jakkolwiek wyniki wykonanych analiz rzuciły sporo światła na problematykę zmienności przestrzennej najwyższych sum dobowych opadów, to jednak sporo zostało jeszcze do zrobienia. Wiele stwierdzonych prawidłowości wymaga głębszej interpretacji genetycznej, trudnej, czy wręcz niemożliwej do przeprowadzenia bez dodatkowych danych i szerszego kontekstu ich analizy.

Szczególną uwagę należy poświęcić na studia poszczególnych przypadków, typowych dla wyróżnionych klas struktury przestrzennej. Do ich przeprowadzenia konieczny jest kompletny zbiór dobowych sum opadów z terenu całej Polski, obejmujących kilkudniowy okres przed, w trakcie i po analizowanym epizodzie, a przede wszystkim mapy synoptyczne ukazujące warunki tworzenia opadów.

Inny potencjalny owocny kierunek, niewymagający pracochłonnego tworzenia nowych baz danych, stanowi zbadanie relacji pomiędzy pojawianiem się wyróżnionych typów struktury przestrzennej miesięcznych i rocznych MSDO a częstością pojawiania się nad Polską różnych mas powietrza, określonych na podstawie kalendarzy typów cyrkulacji (Lityński 1969, Sępniewska-Podraźka 1991).

Uwagę należy również poświęcić potencjalnemu regionalnemu zróżnicowaniu struktury przestrzennej maksymalnych opadów dobowych. Stwierdzone w niniejszej pracy prawidłowości mają najprawdopodobniej zastosowanie dla obszaru znacznie większego niż terytorium naszego kraju, być może w dużym stopniu dla całej strefy klimatu umiarkowanego wilgotnego. Regionalne odmienności dotyczyć zatem będą nie charakterystyk jakościowych, ale przede wszystkim, mniej lub bardziej wyraźnych, statystycznych różnic w częstości pojawiania się i względnym znaczeniu

poszczególnych składowych. Badanie takiego zróżnicowania jest aktualnie jak najbardziej możliwe i planowane w ramach kolejnego etapu prac. Wydaje się, że ze względu na skokowe, jakościowe i ilościowe, zróżnicowanie reżimu opadów w Polsce pomiędzy Niżem i pasem wyżyn a górami, analizę taką należy przeprowadzić osobno dla obu obszarów, korzystając z charakterystyk struktury przestrzennej obliczanych lokalnie w ruchomym oknie (Hass 1990). Przy takim wariancie obliczeń uwzględniona mogłaby być także w pełni anizotropia (por. rozdz. III.2.E). Niestety, do tej pory nie opracowano zadowalającej automatycznej procedury tworzenia anizotropowych modeli struktury przestrzennej. Być może problem będzie można rozwiązać stosując podejście nieparametryczne (Yao, Journel 1998, Ma, Yao 2001), połączone z przestrzenną PCA (Stach, Tamulewicz 2005a).

Obiecującą perspektywę stanowi również możliwość wykonywania dla rocznych MSDO kosymulacji, zarówno ich wysokości jak i czasu wystąpienia. Jest to logiczną konsekwencją istnienia autokorelacji przestrzennej zarówno w zbiorze sum opadów, jak i terminów ich pojawiania się. Dzięki temu będzie można wykonywać wiarygodne, w pełni probabilistyczne, czasoprzestrzenne oceny ryzyka wysokich opadów w skali regionów, a nawet średniej wielkości zlewni rzecznych. Do tej pory robiono to wyłącznie na podstawie danych punktowych, których reprezentatywność przestrzenna była trudna do zweryfikowania. Nieznane były relacje skali pozwalające na „konwersję” prawdopodobieństw punktowych na obszarowe.

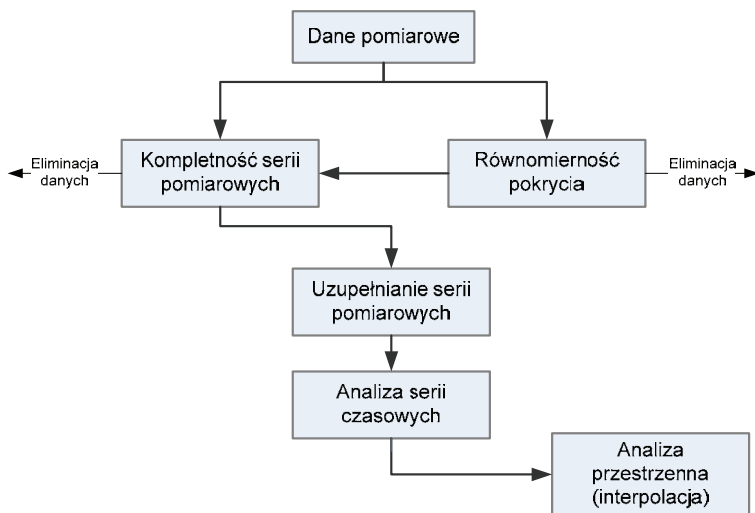
Semiwariogramy empiryczne danych MSDO mają charakter złożony – składają się z dwóch do pięciu modeli elementarnych. Jest to zazwyczaj interpretowane, jak to już wspomniano w wcześniej, jako wynik działania i sumowania się efektów zjawisk operujących w różnych skalach przestrzennych. Można to przyrównać do hałasu ulicznego powstałego poprzez „mieszanie się” dźwięków pochodzących z różnych źródeł: samochodów, tramwajów, ludzi itp. Dźwięki te poprzez swoje specyficzne cechy: amplitudę, częstotliwość, wysokość itp. mogą być jednak łatwo zidentyfikowane i oddzielone od siebie (odfiltrowane). Dokonuje się to w pewnym stopniu w mózgu każdej osoby słyszającej hałas, a bardziej precyzyjnie przez inżyniera akustyka, za pomocą odpowiednich urządzeń technicznych. Tego typu zabieg, rozdzielania czy odfiltrowania składowych „mieszaniny”, możliwy jest też w wypadku danych przestrzennych. Nosi on nazwę krigingu czynnikowego (ang. *factorial kriging*, Deutsch, Journel 1998, Goovaerts 1997). Warunkiem uzyskania sensownych wyników jest niezależność składowych oraz ich „prosta” addytywność. W odniesieniu do danych MSDO być może nie zawsze spełniane jest pierwsze kryterium – w sytuacji kiedy komórki konwekcyjne stanowią element złożonych układów frontalnych. Mimo to, metoda krigingu czynnikowego stwarza obiecującą możliwość rozdzielania, i ilościowej oceny, głównych komponentów tworzących pole MSDO.

## Dodatek: Szczegółowe zagadnienia metodyczne

### 1. Problem zmian ilości i rozkładu przestrzennego danych

Większość opracowań klimatologicznych dotyczących zmienności przestrzennej parametrów klimatu jest wykonywana według ustalonego, tradycyjnego schematu metodycznego: „najpierw agregacja, później interpolacja”. Procedura taka obejmuje zazwyczaj kilka etapów (ryc. 123):

- selekcję punktów pomiarowych ze względu na jednolitość pokrycia analizowanego obszaru. Zazwyczaj, jeśli istnieją skupienia stacji leżących blisko siebie, część z nich eliminuje się w celu uzyskania zagęszczenia typowego dla pozostałego terenu. Zwykle nie testuje się, czy i w jakim stopniu informacje z blisko leżących punktów są nadmiarowe;



Ryc. 123. Typowy przebieg analizy danych klimatologicznych. Dokładne objaśnienia w tekście

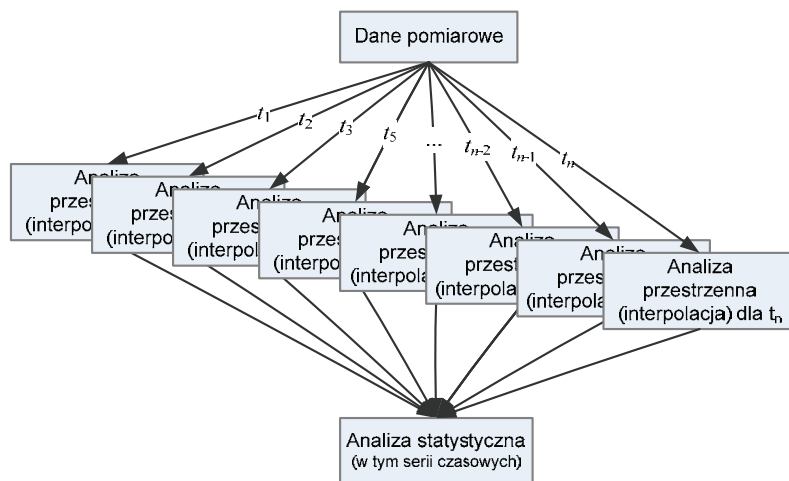
- sprawdzanie jednorodności i kompletności ciągów pomiarowych dla poszczególnych stanowisk (punktów). Stacje z danymi budzącymi wątpliwości i/lub z dużymi lukami w ciągach pomiarowych są albo eliminowane z analizowanej bazy, albo z uwagi na „ważność” stanowiska stosuje się różnorodne techniki statystyczno-matematycznego wyrównywania i uzupełniania danych;

- obliczanie różnorodnych statystyk dla zweryfikowanych punktów pomiarowych;

- interpolację uzyskanych statystyk na cały analizowany obszar.

Opisana wyżej procedura ma wiele zalet, ale także dwie poważne wady. Po pierwsze, w trakcie przygotowania materiału eliminowana jest część potencjalnie bardzo użytecznych danych z punktów o niekompletnych ciągach pomiarowych i/lub z obszarów o dużym zagęszczeniu stacji. Po drugie, agregacja danych dla punktów i interpolacja wyłącznie parametrów statystycznych, jak na przykład średnie okresowe, powoduje, że nieznaną pozostaje zmienność czasowa struktury przestrzennej analizowanego parametru. W wypadku wielu parametrów klimatycznych występują powtarzalne układy przestrzenne, związane na przykład z warunkami cyrkulacyjnymi, które zazwyczaj pozostają zamazane lub kompletnie nieczytelne, kiedy analiza ogranicza się do statystyk obliczanych dla miesięcy, sezonów, lat, czy wielolecia.

W niniejszym opracowaniu przyjęto odmienną strategię analizy danych: „najpierw interpolacja, później agregacja” (ryc. 124). Jej podstawą jest wykorzystanie wszystkich dostępnych punktowych danych w elementarnym



**Ryc. 124.** Przebieg analizy danych klimatologicznych uwzględniającej wszystkie dostępne wyniki pomiarów oraz najpierw ich przestrzenną, a później czasową syntezę. Dokładne objaśnienia w tekście

kroku czasowym dla uzyskania obrazów zmienności przestrzennej analizowanego zjawiska. Zagregowane charakterystyki pola zostają określone w drugim etapie, na podstawie nie oryginalnych punktów danych, ale albo metodami GIS-u z rastrowych obrazów etapowych, albo jako statystyki nie-przestrzenne.

Aby móc zastosować takie podejście, najpierw należy jednak mieć pewność, że zmiany rozmiarów i struktury sieci pomiarowej zachodzące pomiędzy kolejnymi terminami pomiarów nie mają istotnego wpływu na jakość uzyskiwanych wyników. Przeprowadzone przez Stacha i Tamulewicza analizy przestrzenne sum miesięcznych opadów dla tego samego wielolecia 1956-1980 (2005a, 2005b, 2005c) wykazały, że zachodzące wówczas względnie niewielkie zmiany ilości i rozmieszczenia danych pomiarowych nie zmieniały w wyraźny sposób dokładności uzyskanego rozkładu przestrzennego zjawiska. Dokładność ta uzależniona była bowiem przede wszystkim od specyfiki samego zjawiska w postaci lokalnej zmienności pola opadów, wyrażonej poprzez statystyki najbliższego sąsiada (por. rozdz. V). Zróżnicowanie globalne sum opadów, charakteryzowane na przykład przez odchylenie standardowe, miało znaczenie marginalne.

W cytowanych powyżej pracach wykonywano jednak ocenę jedynie względem sum miesięcznych. Dane MSDO będące podzbiorem kompletnych ciągów sum dobowych opadów mają wyraźnie odmienny charakter. Nie mając możliwości przeprowadzenia odpowiednich testów na zbiorach sum dobowych opadów, uznano za wystarczające uzasadnienie wyżej postawionej tezy dobrze udokumentowany przykład z literatury. Są nim wyniki analiz przeprowadzonych w trakcie eksperymentu zwanego *Spatial Interpolation Comparison 97* (SIC-97, Dubois i in. 2003). Celem tego przedsięwzięcia było przetestowanie na jednym zbiorze danych opadowych estymacji przestrzennych wykonanych za pomocą różnych metod interpolacyjnych pod kątem ich dokładności i przydatności do rutynowych ocen zagrożeń skażeniami (szybkość obliczeń, stopień komplikacji procedury, możliwość automatyzacji itp.). Jako przykładowe dane posłużyły sumy dobowe opadów z 8 maja 1986 roku, kiedy nad Szwajcarią przemieszczała się chmura radioaktywnych aerozoli pochodząca z uszkodzonego reaktora elektrowni atomowej w Czarnobylu na Ukrainie.

W ramach SIC-97 Dubois i Shibli (2003) przeprowadzili także ocenę wpływu ilości danych na jakość uzyskiwanych wyników, wyrażoną poprzez pierwiastek średniego błędu kwadratowego (RMSE). Eksperyment polegał na losowaniu z całego zbioru 467 danych podzbiorów od 25 do 300 punktów w kroku co 25. Na podstawie tych danych dokonywano estymacji metodami IDW - odwrotnej wagi odległości (ang. *inverse distance weighting*), MQ - funkcji multikwadratowych (ang. *multiquadratic function*), OK - zwykłego krigingu (ang. *ordinary kriging*), optymalizując parametry obliczeń za pomo-

czą krosvalidacji (Davis 1987, Deutsch, Journal 1998, Goovaerts 1997, Wackernagel 2003). W celu uniknięcia przypadkowych wyników, losowanie i obliczenia dla każdej wielkości podzbioru powtarzano 10 razy. Dla każdej z trzech metod tendencje RMSE były identyczne. Po przekroczeniu 100, kolejne przyrosty ilości wykorzystywanych danych nie miały już istotnego wpływu na wielkość błędu estymacji. Ponieważ optymalizacja parametrów obliczeń metodą krosvalidacji ma wadę polegającą na tym, że te same dane służą do wykonania estymacji i oceny jej jakości, przeprowadzono dodatkowy test. Użyto do niego pozostałe z losowania dane to jest 467 - N, gdzie N oznacza wielkość wylosowanej próby = 25, 50, 100, ..., 300. Tym razem wartość błędu stabilizowała się po przekroczeniu 275 danych.

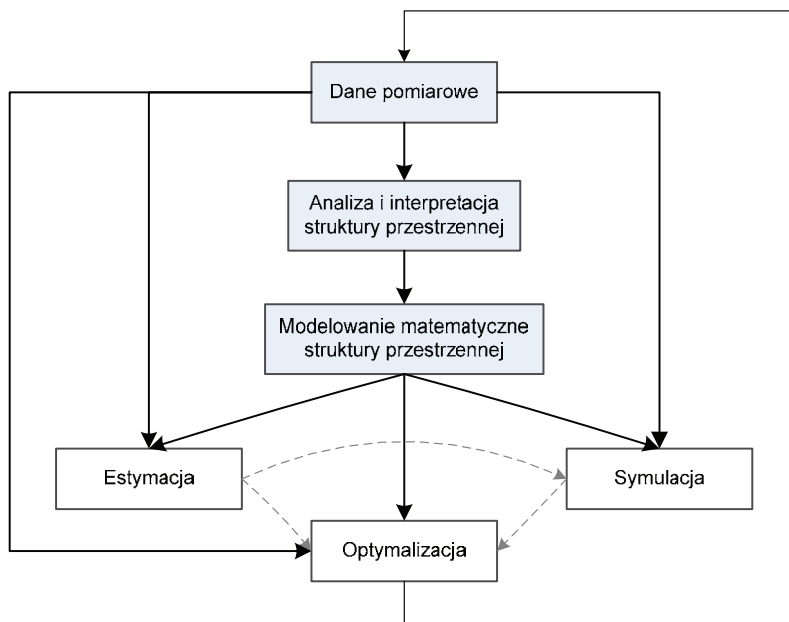
Żeby ocenić znaczenie przedstawionych powyżej wyników eksperymentu numerycznego wykonanego w ramach SIC-97, trzeba brać pod uwagę kilka czynników. Po pierwsze, eksperyment był wykonany na jednym, konkretnym zbiorze sum dobowych opadów. Relacje te mogą się zmieniać prawdopodobnie w znacznym zakresie w zależności od genezy i charakteru opadów. Po drugie, Szwajcaria jest znacznie mniejsza od Polski (tylko 41,3 tys. km<sup>2</sup>), ale w całości górzysto-wyżynna, co bardzo wpływa na zmienność przestrzenną opadów. Porównując zagęszczenie sieci pomiarowej obu krajów i przyjmując 300 stanowisk w Szwajcarii jako poziom zapewniający stabilne wyniki estymacji, można oceniać, że dla Polski oznaczałoby to około 1200-1500 punktów. Jest to, zdaniem autora, szacunek bardzo ostrożny, a rzeczywista wartość - niższa.

## 2. Estymacja, symulacja i optymalizacja

Opis i modelowanie (por. rozdz. III.2) struktury przestrzennej analizowanej cechy rzadko stanowią główny cel prac wykorzystujących metody geostatystyki. Jest to zazwyczaj tylko etap, w którym zidentyfikowane i matematycznie określone reguły zmienności przestrzennej są następnie wykorzystywane do estymacji, symulacji lub optymalizacji (ryc. 125). Takie motywy stanowiły również genezę niniejszej rozprawy, chociaż w dalszym etapie, o czym wspomniano we wprowadzeniu (por. rozdz. I), jej zakres został ograniczony. Poniżej zostaną w skrócie omówione te elementy wybranej metodyki estymacji i symulacji pola prawdopodobieństwa MSDO<sup>54</sup>, które miały największy wpływ na sposób rozwiązania problemu stanowiącego istotę niniejszej rozprawy - analizy i modelowania struktury przestrzennej maksymalnych sum dobowych opadów na terenie Polski.

---

<sup>54</sup> Wyniki tych prac zostaną opublikowane w osobnej rozprawie.



Ryc. 125. Schemat przebiegu pełnej analizy geostatystycznej. Elementy ujęte w niniejszej rozprawie zaznaczono szarym tłem. Szersze objaśnienia w tekście

Sednem geostatystyki jest problem estymacji wartości cechy  $z$  w nieopróbowanej lokalizacji  $\mathbf{u}$ . Dysponujemy w tym momencie zbiorem  $n$  pomiarów cechy  $z$  dokonanych w szeregu punktów o znanych współrzędnych  $\mathbf{u}_\alpha$  czyli  $z(\mathbf{u}_\alpha)$ . Wykazują one wszystkie pewien stopień autokorelacji przestrzennej. W większości metod interpolacyjnych ową nieznaną, szacowaną wartość oblicza się jako liniową kombinację z  $n(\mathbf{u}_\alpha)$  sąsiednich wyników pomiarów [30]:

$$\hat{z}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u}_\alpha)} \lambda_\alpha z(\mathbf{u}_\alpha) \quad [30]$$

W geostatystyce nieznaną wartość  $z(\mathbf{u})$  jest interpretowana jako realizacja funkcji losowej  $FL(\mathbf{u})$ , która jest w pełni charakteryzowana przez jej rozkład prawdopodobieństwa  $F(\mathbf{u};z) = \text{Prob}\{Z(\mathbf{u}) \leq z\}$ . Wagi  $\lambda_\alpha$  uwzględniają odległość lokalizacji danych od miejsca estymacji, ich przestrzenną konfigurację (w tym ich potencjalne skupienia), a przede wszystkim ich przestrzenną strukturę określoną poprzez model uzyskany na podstawie autokowariancji empirycznej [2] (por. rozdz. III.2.2). Uzyskuje się je z rozwiązania układu  $(n(\mathbf{u}_\alpha) + 1)$  równań liniowych, zwanych równaniami zwykłego krigingu (ang. *Ordinary Kriging* = **OK**, [31]):



$$\sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta} \gamma(\mathbf{u}_{\alpha} - \mathbf{u}_{\beta}) - \mu = \gamma(\mathbf{u}_{\alpha} - \mathbf{u}) \quad \alpha = 1, \dots, n(\mathbf{u})$$

$$\sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta} = 1$$
[31]

gdzie  $\mu$  to mnożnik Lagrange'a konieczny ze względu na ograniczenie wag (sumowanie się wag do jedności).

Kriging daje w większości wypadków znacznie lepsze efekty niż inne metody estymacji (= interpolacji, Dubois i in. 2003, Dubois, Shibli 2003, Lam 1983, Weber, Englund 1992, 1994, Zimmerman i in. 1999), ponieważ jego algorytm zawiera procedury minimalizacji błędów i zapewnienia ich losowego rozkładu (estymator jest nieobciążony, co oznacza bilansowanie się błędów ujemnych i dodatnich).

W wielu zastosowaniach, zarówno teoretycznych, jak i praktycznych, równie ważne jak ocena nieznannej, lokalnej (punktowej), wartości analizowanej cechy jest określenie niepewności tej oceny. Poza tym, zwykły kriging, jak wiele innych metod interpolacyjnych, daje wyniki będące optymalną estymacją z punktu widzenia kryterium najmniejszych kwadratów. Często jednakże, dla konkretnych potrzeb, niezbędny jest szacunek nieznannej wartości dokonany względem innych kryteriów. Dlatego, w obrębie geostatystyki opracowano kilka metod umożliwiających, zamiast określenia jedynie jednej „optymalnej” wartości analizowanej cechy w nieopróbowanej lokalizacji, budowę kompletnych lokalnych modeli jej rozkładu (dystrybuanty). W odniesieniu do danych MSDO takie podejście ma dodatkowy, metodologiczny aspekt. Ze względu na nieciągłość przestrzenną epizodów opadowych i asynchroniczność MSDO, celem estymacji nie może być określenie konkretnej wartości, ale raczej prawdopodobieństwa jej wystąpienia. Efektem końcowym nie jest wówczas typowa mapa rastrowa czy izarytmiczna, ale model pola prawdopodobieństwa, z którego możemy zależnie od potrzeb uzyskać dowolne przekroje (mapy dla określonych progów prawdopodobieństwa) lub bardziej złożone charakterystyki. Do wspomnianych wyżej metod geostatystycznych umożliwiających budowę kompletnych lokalnych modeli rozkładu estymowanej cechy należą kriging multigaussowski (ang. *multigaussian kriging*), kriging dysjunktywny (ang. *disjunctive kriging*) oraz kriging wartości kodowanych (ang. *indicator kriging - IK*), zwany czasem w polskiej literaturze tematu krigingiem wskaźnikowym bądź indyktorowym (Zawadzki 2005). Są one opisane i porównane w kilku najważniejszych podręcznikach teorii geostatystyki (Chilès, Delfiner 1999, Goovaerts 1997, Webster, Oliver 2001), a także w szeregu prac eksperymentalnych (Carr, Deng 1987, Goovaerts 2001, Lajaunie 1990, Lark, Ferguson

2004, Moyeed, Papritz 1999, Papritz, Dubois 1999, Solow 1986). W ramach projektu badawczego, którego częściowym efektem jest niniejsza rozprawa, zdecydowano się zastosować kriging danych kodowanych (IK). Decyzję tę podjęto analizując bilans wad i zalet tej metodyki.

Wady krigingu wartości kodowanych obejmują (Christakos 2000, Deutsch, Journel 1998, Goovaerts 1997, Olea 1999):

- utratę części informacji ze względu na dyskredytyzację ciągłej dystrybuanty empirycznej;

- pracochłonność – konieczność czasochłonnego budowania modelu semiwariancji dla każdej wartości progowej; często występujące trudności w określeniu modeli dla wartości bardzo niskich i bardzo wysokich zmuszają do subiektywnych decyzji, a te rodzą wątpliwości co do optymalności uzyskanych estymacji;

- wykraczanie estymowanych prawdopodobieństw poza dopuszczalny zakres (0, 1) oraz błędy w ich relacjach porządkowych;

- arbitralnie przyjmowaną metodę interpolacji/ekstrapolacji uzyskanej warunkowej dystrybuanty.

Zaletami krigingu wartości kodowanych są:

- potwierdzona w dziesiątkach zastosowań i testów metodycznych skuteczność;

- brak trudnych do weryfikacji założeń dotyczących rozkładu statystycznego populacji (metoda nieparametryczna);

- że żadna z alternatywnych metod nie jest wyraźnie lepsza;

- łatwość ich używania; alternatywne metody są bardziej skomplikowane – bardziej „podatne” na błędy metodyczne;

- łatwa możliwość uwzględnienia danych uzupełniających („twardych” i „miękkich”);

- powszechna dostępność oprogramowania (Deutsch, Journel 1998, Mao, Journel 1998, Pardo-Igúzquiza, Dowd 2005, Richmond 2002); na bazie publicznie dostępnego kodu komputerowego tradycyjnych metod estymacji geostatystycznej (SK, OK) łatwo można samemu napisać program komputerowy realizujący obliczenia IK.

Istotą metody IK jest najpierw budowa globalnego modelu niepewności opartego na dystrybuancie empirycznej wyliczonej ze wszystkich danych pomiarowych, a następnie jego lokalna modyfikacja na podstawie wektora danych kodowanych z sąsiadujących z lokalizacją punktu estymacji stanowisk pomiarowych.

Na podstawie rozkładów prawdopodobieństwa uzyskanych metodą krigingu wartości kodowanych obliczane są także dwie charakterystyki ogólne: średnia wartość szacowanej dystrybuanty, zwana średnią oczekiwaną (ang. *E-type mean* [32], ryc. 128) oraz wariancja warunkowa:

$$\begin{aligned}
 [z(\mathbf{u})]_E^* &= \int_{-\infty}^{+\infty} z dF(\mathbf{u}; z|(n)) \\
 &\approx \sum_{k=1}^{K+1} z_k [F(\mathbf{u}; z_k|(n)) - F(\mathbf{u}; z_{k-1}|(n))]
 \end{aligned}
 \tag{32}$$

gdzie  $z_k, k = 1, \dots, K$  to  $K$  przyjętych wartości progowych, a  $z_0 = z_{\min}, z_{K+1} = z_{\max}$  to minimum i maksimum zakresu cechy  $z$ , które są parametrami podawanymi przez osobę wykonującą obliczenia (operatora). Średnia warunkowa każdej klasy ( $z_{k-1}, z_k$ ) określona jako  $z_k$  jest uzyskiwana na podstawie procedury interpolacji, której rodzaj i parametry są również specyfikowane przez operatora.

Wszystkie metody estymacji oparte na regresji liniowej liczonej metodą najmniejszych kwadratów mają jednak istotną wadę (tab. 28). Jest nią zmienne lokalnie „wygładzanie” rozkładu statystycznego estymowanej cechy. W kategoriach ilościowych wada metod interpolacyjnych polega na tym, że wygenerowany za ich pomocą model (powierzchnia opisana izoliniami) ma inną charakterystykę statystyczną (rozkład zmiennej/histogram) i przestrzenną (semiwariogram) niż oryginalna próbka, na podstawie której powstał. Aby ten problem choć częściowo pokonać, na gruncie geostatystyki opracowano kilka metod symulacji przestrzennej, które honorując dane pomiarowe, umożliwiają generowanie szeregu równie prawdopodobnych obrazów (modeli), w pełni zachowujących charakterystyki rozkładu statystycznego próbki i strukturę przestrzenną zjawiska (Deutsch, Journel 1998, Goovaerts 1997, tab. 28). Daje to duże możliwości ilościowej oceny niepewności dokonanych szacunków. Geostatystyczne metody estymacji dają

**Tabela 28.** Cechy i zastosowania estymacji oraz symulacji geostatystycznych

	<b>Estymacja (kriging)</b>	<b>Symulacja warunkowa</b>
<b>Efekt obliczeń</b>	jeden model „deterministyczny”	wiele realizacji (modeli „probabilistycznych”)
<b>Właściwości</b>	honoruje dane, minimalizuje wariancję błędu	honoruje dane, histogram, wariogram, gęstość spektralną i in.
<b>Obraz</b>	łagodny (gładki), zwłaszcza wobec „chaotycznego” modelu wariogramu	bardzo zróżnicowany, zwłaszcza wobec „chaotycznego” modelu wariogramu
<b>Dane wynikowe</b>	tendencja do tworzenia powierzchni trendu z dala od danych; możliwość lokalizacji położenia punktów pomiarowych	taka sama zmienność obrazu w każdej części; niemożność odgadnięcia lokalizacji punktów pomiarowych
<b>Zastosowania</b>	tworzenie map izarytmicznych; estymacja punktowa	modelowanie niejednorodności; szacowanie niepewności; estymacja obszarowa

optymalne oceny (wraz z szacunkiem ich niepewności) wartości analizowanego parametru w punkcie, metody symulacji dla obszarów (zlewni rzecznych, regionów fizycznogeograficznych, jednostek administracyjnych itp.). Istotą symulacji geostatystycznych jest zastosowanie metodyki Monte Carlo (próbkiowania rozkładu statystycznego), ograniczonej poprzez uwzględnienie danych pomiarowych i modelu struktury przestrzennej zjawiska. Z tego względu w nazwach poszczególnych procedur stosowany jest określnik „symulacja warunkowa” (ang. *conditional simulation*).

Różnice między geostatystycznymi estymacjami a symulacjami w kontekście analizowanych danych przedstawiono na przykładzie rocznych MSDO z roku 1976 (tab. 29, ryc. 126-128). Tabela 29 i rycina 126 pokazują relacje między statystykami opisowymi i rozkładem danych empirycznych a danymi modeli estymowanych i symulowanych. Zbiór estymacji ma znacząco mniejszy zakres niż oryginalne dane pomiarowe (wyższe minimum i niższe maksimum) i mniejszą zmienność reprezentowaną przez odchylenie standardowe. Porównanie kwantyli rozkładu (ryc. 126) wskazuje, że estymacje generalnie zawyżają wartości mniejsze od średniej i zaniżają od niej większe. Brak dobrego odwzorowania charakterystyk zmienności cechy jest szczególnie dokuczliwy, jeśli, jak to jest w tym przypadku, analizuje się zmienność przestrzenną wartości ekstremalnych, a nie przeciętnych. Rozkład wartości symulowanych jest praktycznie tych wad pozbawiony. Nie odwzorowują one jednak całkowicie wiernie odpowiednich charakterystyk statystycznych danych pomiarowych, lecz jest to cecha raczej pożądana<sup>55</sup>. Dane pomiarowe to przecież niewielka próba z całej populacji<sup>56</sup> i są także obciążone błędami<sup>57</sup>, a więc wyliczone z nich statystyki stanowią jedynie przybliżenie. Zbyt ściśle ich honorowanie mogłoby prowadzić do fałszywego wyobrażenia o pewności i podobieństwie symulowanych wartości. Równie ważne są różnice modeli estymowanych i symulowanych w ujęciu przestrzennym (ryc. 127 i 128). Semiwariogram uzyskany z próby danych estymowanych różni się zasadniczo od wyjściowego semiwariogramu empirycznego – brakuje nuggetu, początkowy odcinek ma charakter paraboliczny, semiwariancje wyliczone dla każdego odstępów są o połowę mniejsze. Estymowany obraz jest zatem sztucznie wygładzony, tworząc mylne wrażenie ciągłości przestrzennej analizowanej cechy. Daje to co prawda bardziej klarowny, zgeneralizowany obraz zróżnicowania geograficznego, ale należy pamiętać, że owa generalizacja nie jest jednolita. Zależy ona bowiem od lo-

---

<sup>55</sup> Odstępstwa dystrybuanty i semiwariogramu danych symulowanych od wartości wyliczonych z oryginalnych pomiarów nazywane są fluktuacjami ergodycznymi (ang. *ergodic fluctuations*, Deutsch, Journal 1998, Goovaerts 1997).

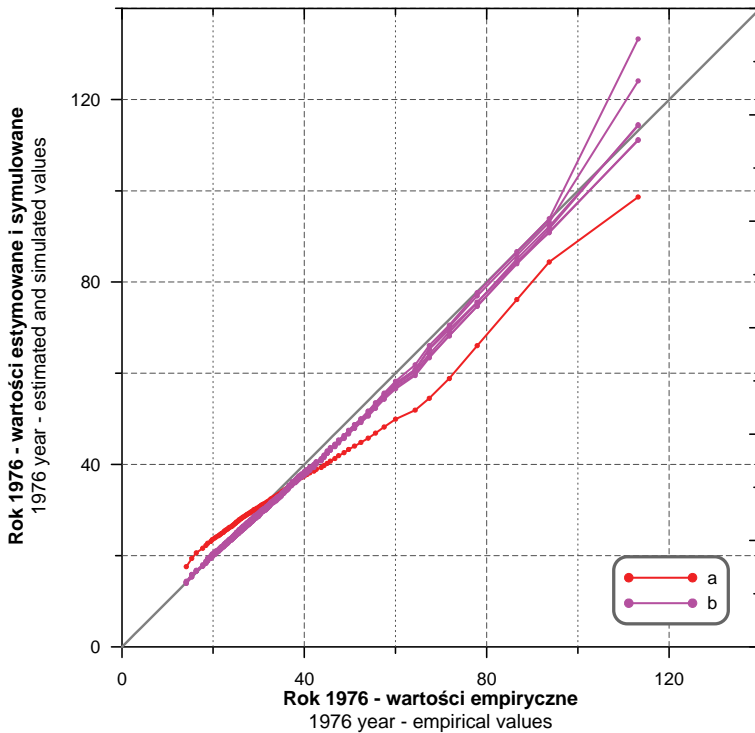
<sup>56</sup> Patrz przypis 10 (s. 39).


<sup>57</sup> Patrz podrozdział V.2 (s. 64).

kalnej konfiguracji danych: wygładzanie jest minimalne w bezpośrednim sąsiedztwie lokalizacji pomiarów, a zwiększa się wraz ze wzrostem odległości punktu estymacji od stanowisk pomiarowych. Estymacje krigingowe są zatem bardziej zróżnicowane na gęściej opróbowanych obszarach, mogąc tworzyć pozorne struktury przestrzenne. Wygładzone mapy interpolowane nie powinny być zatem używane w tych zastosowaniach, gdzie końcowy wynik jest czuły na obecność wartości ekstremalnych i na charakter ich ciągłości przestrzennej. Obrazy symulowane są tych wad w znacznym stopniu pozbawione (ryc. 128). Semiwariogramy danych symulowanych oscylują wokół semiwariogramu wyliczonego z pomiarów, a początkowa nieciągłość (nugget) jest wiernie odwzorowana (ryc. 127). Każdy obraz symulowany jest unikatowy, ale honorowanie danych pomiarowych powoduje, że główne cechy zmienności przestrzennej są zachowane. Na mapach widoczna jest „ziarnistość”, wynikająca z bezpośredniego sąsiedztwa węzłów siatki o silnie zróżnicowanych wartościach, co można być interpretowane jako nieciągłości opadów wynikające z ich ograniczonego zasięgu przestrzennego oraz asynchroniczności danych MSDO. To duża zaleta, ale należy pamiętać, że owa właściwość symulowanych map MSDO wynika z użycia jednego, globalnego modelu struktury przestrzennej, który przy tak dużym obszarze i tak genetycznie zmiennym zjawisku nie zawsze jest adekwatny lokalnie. W miejscach występowania i dominacji szczególnie rozległych opadów frontalnych o łagodnej zmienności przestrzennej, symulowany obraz może być sztucznie zróżnicowany. Rozwiązaniem tego problemu mogłoby być zastosowanie lokalnych modeli, co w wypadku estymacji rzeczywiście ma już miejsce (Haas 1990, Minasny i in. 2005, Whelan i in. 2001).

**Tabela 29.** Porównanie statystyk oryginalnego zbioru danych pomiarowych rocznych MSDO z roku 1976 oraz wyników jego estymacji (średnia oczekiwana IK: *E-type mean*) i symulacji (trzy pierwsze realizacje) w siatce 1 na 1 km

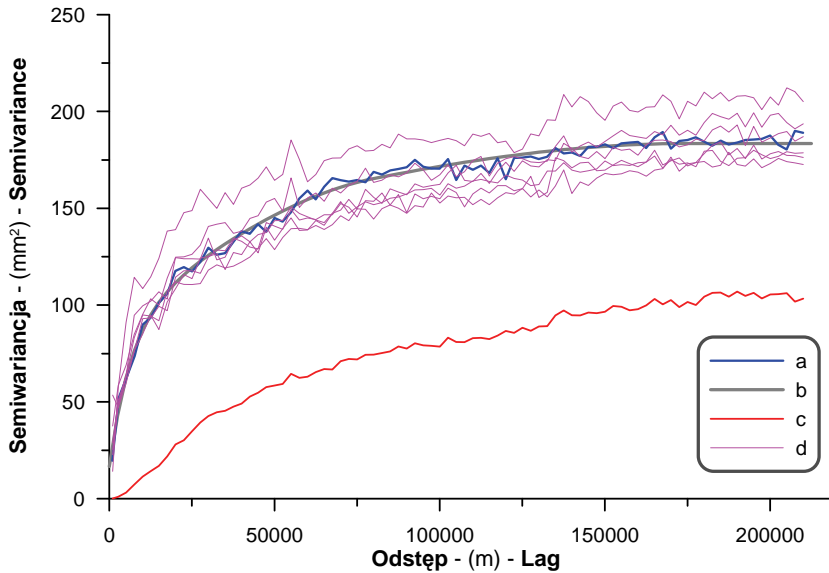
Parametr	Parameter	1976 dane / data	1976 est. E-mean	1976 sim. 01	1976 sim. 02	1976 sim. 03
Liczebność	Count	2678	319 114	319 114	319 114	319 114
Minimum	Minimum	13,0	16,5	13,0	13,0	13,0
1 kwartył	First quartile	25,1	27,4	24,6	24,6	24,6
Mediana	Median	30,8	31,2	29,6	29,6	29,7
3 kwartył	Third quartile	39,4	37,1	37,4	37,2	37,6
Maksimum	Maximum	166,1	109,0	166,1	166,1	166,1
Średnia	Mean	34,61	33,71	33,44	33,31	33,40
Odch. stand.	Standard deviation	14,389	10,199	14,125	13,917	13,720
Skośność	Skewness	2,040	2,363	2,406	2,381	2,190
Kurtoza	Kurtosis	9,918	11,592	12,539	12,416	10,880




**Ryc. 126.** Porównanie kwartyli (0,001, 0,005, 0,01, 0,02, 0,03, ..., 0,97, 0,98, 0,99, 0,995, 0,999) rozkładu danych empirycznych rocznych MSDO z roku 1976 z (a) wynikami jego estymacji (średnia oczekiwana IK: *E-type mean*) i (b) symulacji (trzy pierwsze realizacje) w siatce 1 na 1 km 

Podsumowując, mimo że najbardziej dokładne szacunki wartości punktowych uzyskuje się metodami estymacji, do ilustracji rozważań na temat zmienności struktury przestrzennej w ujęciu obszarowym najlepiej nadają się obrazy symulowane. Z tego względu były one wykorzystywane w rozdziale VI niniejszej rozprawy.

Ostatnim, choć niemniej ważnym, zastosowaniem geostatystyki jest optymalizacja próbkowania. Intuicyjnie wiadomo, że stopień pożądanego zagęszczenia pomiarów powinien być funkcją zmienności czasowej i/lub przestrzennej mierzonego zjawiska. Tam i wtedy gdzie/kiedy zmienność jest niewielka, pomiarów może być mało, gdy zaś coś zmienia się szybko i na niedużym obszarze, powinniśmy sieć pomiarową zagęścić i zwiększyć ich frekwencję. Geostatystyka oferuje szerokie spektrum narzędzi umożliwiających optymalizację próbkowania przy zakładanym stosunku kosztów do efektów i danym zbiorze lokalizacji potencjalnych obiektów badań (Par-do-Igúzquiza, Dowd 2005, van Groeningen i in. 1997, 1999).

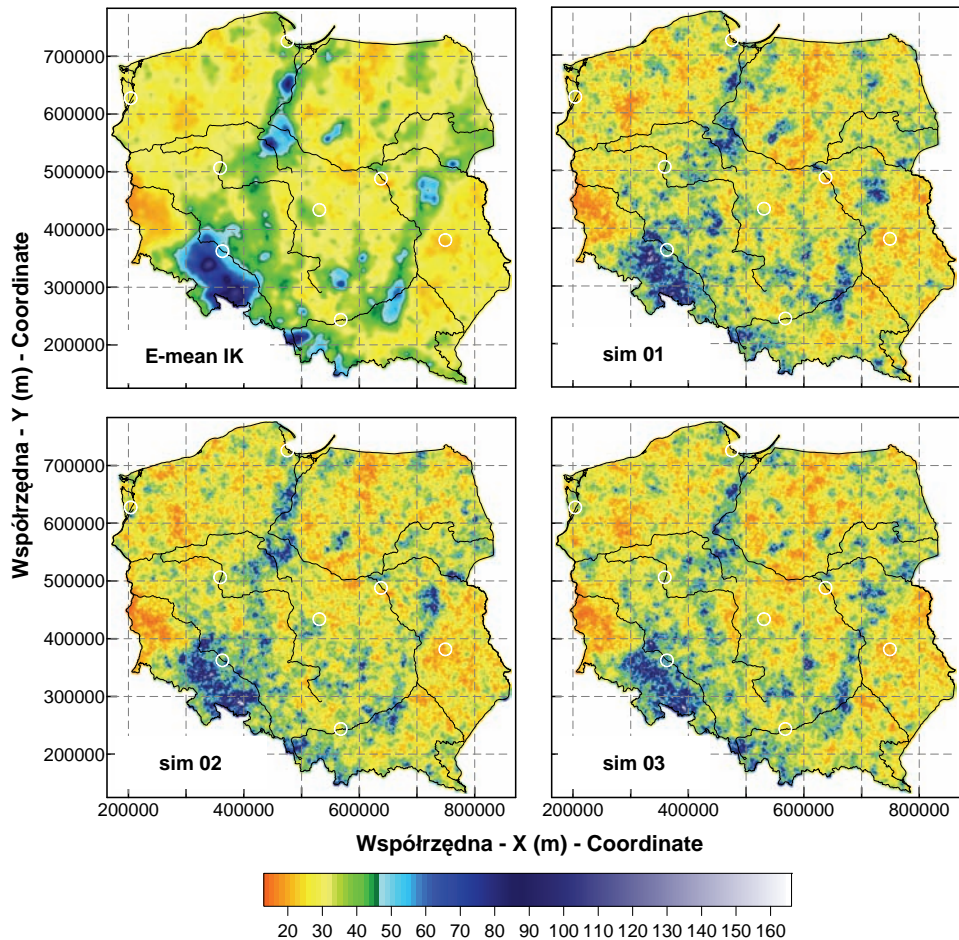


**Ryc. 127.** Porównanie semiwariogramu empirycznego (a) i jego modelu (b) danych pomiarowych rocznych MSDO z roku 1976 z semiwariogramami empirycznymi (c) estymacji (średnia oczekiwana IK: *E-type mean*) i (d) symulacji (trzy pierwsze realizacje). Semiwariogramy (c) i (d) wyliczono dla 2700 danych pobranych losowo z siatki o rozdzielczości 1 na 1 km (319 114 węzłów) 

Najczęściej stosowane metody symulacji przestrzennej, tak zwane „sekwencyjne” (ang. *sequential simulation*), mają istotną słabość, która dyskwalifikowała je do zastosowania w niniejszej pracy. Nie można bowiem, albo z dużymi „problemami”, z ich pomocą w wiarygodny sposób odwzorowywać struktur wielkoskalowych symulowanego zjawiska (Deutsch, Journel 1998, Goovaerts 1997). MSDO są zaś efektem kilku procesów działających w różnych skalach przestrzennych: opadów konwekcyjnych, frontalnych i orograficznych. Symulacje metodami sekwencyjnymi są także dość czasochłonne, co przy 325 zbiorach danych jest bardzo istotne (300 zbiorów miesięcznych i 25 rocznych MSD z lat 1956-1980). Zdecydowano się zatem skorzystać z algorytmu symulacji pola prawdopodobieństwa (ang. *p-field simulation*, Srivastava 1992, Froidevaux 1993), który wyżej wymienionych wad jest pozbawiony. Inicjalne pole prawdopodobieństwa o strukturze przestrzennej zgodnej z modelem semiwariogramu analizowanego zbioru danych generowano za pomocą bardzo szybkiej, bezwarunkowej symulacji spektralnej (Yao 1998a, 1998b). Wartości wszystkich węzłów siatki symulacyjnej były następnie jednocześnie losowane z odpowiednich warunkowych

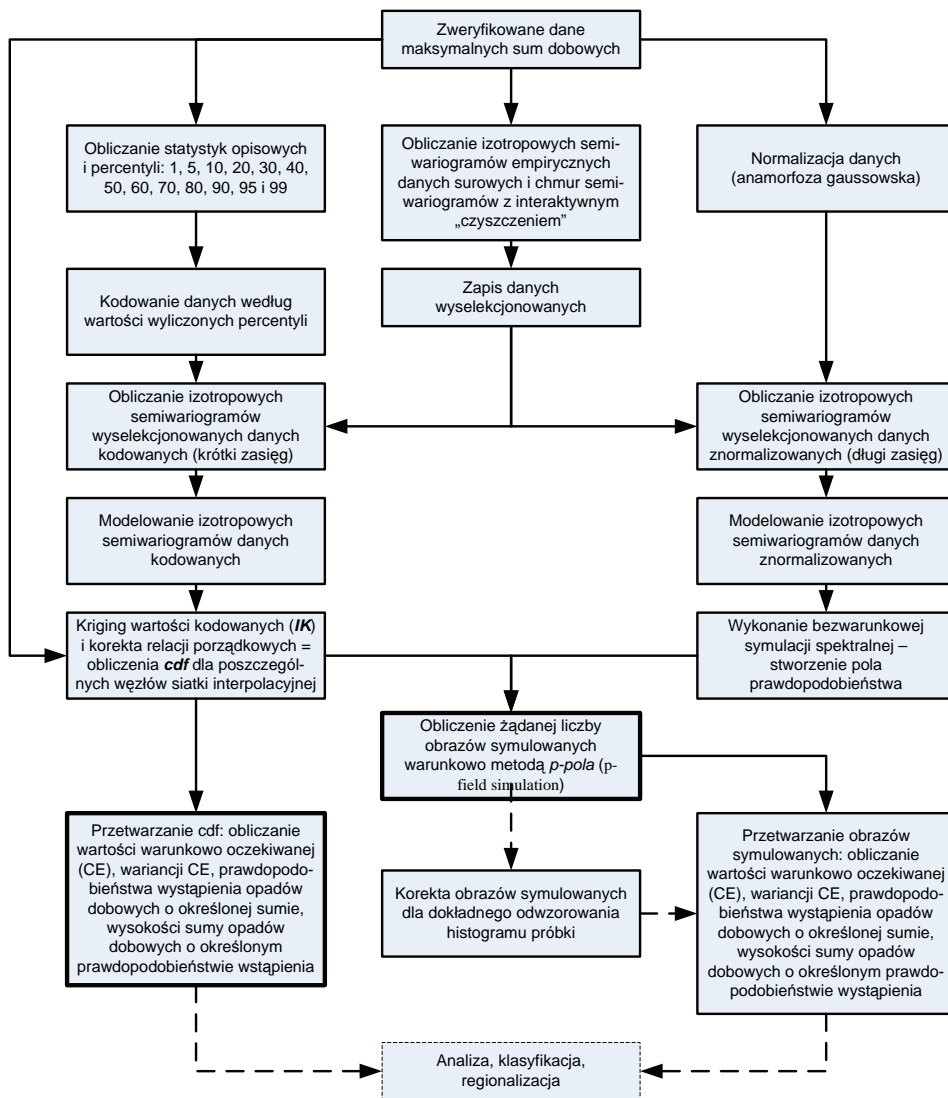
kumulacyjnych funkcji rozkładu prawdopodobieństwa (ang. *conditional cumulative distribution function - ccdf*), uzyskanych za pomocą metody krigingu wartości kodowanych.

Kompletny schemat wykonywanych obliczeń struktury przestrzennej, estymacji i symulacji MSDO przedstawiono na rycinie 129. Zaznaczono na nim, jaka część całego projektu badawczego została wykorzystana w niniejszej rozprawie. Normalizację zmiennych oraz obliczenia i modelowanie semiwariancji wykonano za pomocą programu ISATIS (Bleinès i in. 2006), natomiast pozostałą część procedury – korzystając z programu IKSIM (Ying 2000).



Ryc. 128. Estymowana (średnia oczekiwana IK: *E-type mean*) i symulowana w siatce 1 na 1 km (trzy pierwsze realizacje) zmienność pola rocznych MSDO wyrażonych w mm w roku 1976





**Ryc. 129.** Schemat całej procedury analizy i modelowania struktury przestrzennej oraz estymacji i symulacji pola prawdopodobieństwa MSDO dla pojedynczego zbioru danych miesięcznych lub rocznych. Operacje opcjonalne zaznaczono linią przerywaną. Kroki uwzględnione w niniejszej pracy zaznaczono szarym tłem

Decyzje o zastosowanej metodyce estymacji i symulacji pola prawdopodobieństwa MSDO pociągnęły za sobą istotne konsekwencje w zakresie sposobu analizy i modelowania ich struktury przestrzennej.

W szczególności, użycie do estymacji pola prawdopodobieństwa MSDO krigingu danych kodowanych spowodowało konieczność kodowania binarnego danych względem takiej ilości progów, która w zadowalająco dokładny sposób pozwoli modelować lokalne dystrybuanty. Z danych binarnych dla każdej wartości progowej obliczane były semiwariogramy empiryczne, do których następnie dopasowywano modele. Ta ścieżka metodyczna, choć skomplikowana i czasochłonna, dała unikatową możliwość analizy zróżnicowania struktury przestrzennej MSDO w różnych klasach ich wysokości.

Inicjalne bezwarunkowe pole prawdopodobieństwa, używane przy symulacji metodą *p-pola*, było generowane w przestrzeni danych znormalizowanych. Wymagało to opracowania modeli konwersji poszczególnych podzbiorów MSDO do postaci standardowego rozkładu normalnego, a następnie wyliczenie z nich empirycznych miar struktury przestrzennej i opracowanie ich matematycznych modeli. Jak szczegółowo opisano w dodatku X.4, to pozornie kolejne „utrudnienie” metodyczne miało dwie istotne zalety. Po pierwsze – umożliwiała bezpośrednie porównywanie semiwariogramów poszczególnych zbiorów MSDO, a po drugie – na etapie modelowania redukowało destrukcyjny wpływ danych anomalnych, nie zmieniając w żaden sposób autentycznych prawidłowości autokorelacji przestrzennej analizowanych danych.

### 3. Anizotropia struktury przestrzennej MSDO

Podobieństwo/zróżnicowanie zmienia się w przestrzeni nie tylko wraz z odległością, ale także najczęściej również z kierunkiem. Właściwość ta jest nazywana anizotropią struktury przestrzennej (Goovaerts 1997, Gringarten, Deutsch 2001, Olea 2006). Wyróżnia się dwie jej odmiany:

- geometryczną, kiedy semiwariogramy kierunkowe mają taki sam kształt i wariancję progową, a różnią się zasięgiem; róża zasięgów dla poszczególnych kierunków tworzy kształt elipsy;
- strefową, kiedy semiwariogramy dla poszczególnych kierunków różnią się wariancją progową.

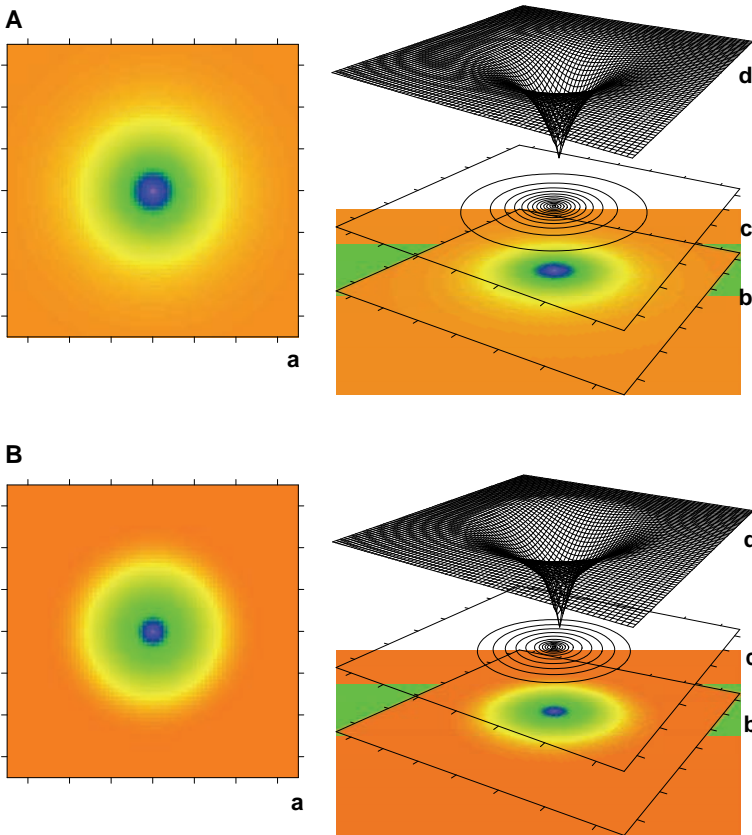
Często spotykane są także kombinacje obu typów.

Charakter anizotropii identyfikuje się na podstawie tzw. powierzchni semiwariogramu (ang. *semivariogram surface*, Deutsch, Journel 1998, Pannatier 1996, ryc. 130 i 131). Oblicza się ją dla ustalonej siatki przedziałów odległości i kierunku – najczęściej w układzie współrzędnych prostokątnych<sup>58</sup>.

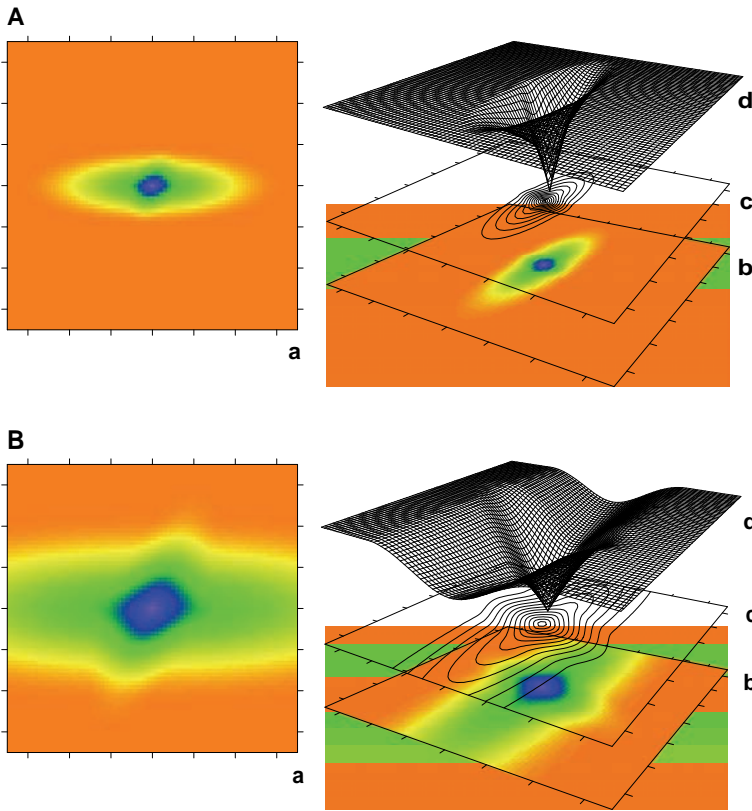
---


<sup>58</sup> Kierunek definiuje się poprzez podanie wielkości współrzędnych X i Y (tangens kąta).

Ponieważ  $\gamma(\mathbf{h}) = \gamma(-\mathbf{h})$ , powierzchnia taka jest w każdym kierunku symetryczna względem początku układu. Jeśli analizowana cecha wykazuje identyczny spadek podobieństwa w każdym kierunku (struktura izotropowa) i ustaloną wariancję progową, to geometrycznym tego obrazem jest płaszczyzna o kształcie symetrycznego „lejka” z kołnierzem (ryc. 130A). Jego przekrój w każdej płaszczyźnie równoległej do osi odległości jest okręgiem. Kiedy izotropowy semiwariogram ma charakter zagnieżdżony (złożony), wówczas odzwierciedla się to, oczywiście, w kształcie lejka, w postaci segmentów o różnym nachyleniu (ryc. 130B).



**Ryc. 130.** Schemat powierzchni modeli semiwariancji prostych (A) i złożonych (B) przy izotropowej autokorelacji danych: a – mapa rastrowa w projekcji płaskiej, b, c, d – obraz w rzucie perspektywicznym, odpowiednio: mapa rastrowa, izoliniowa i siatkowy model trójwymiarowy 🗄

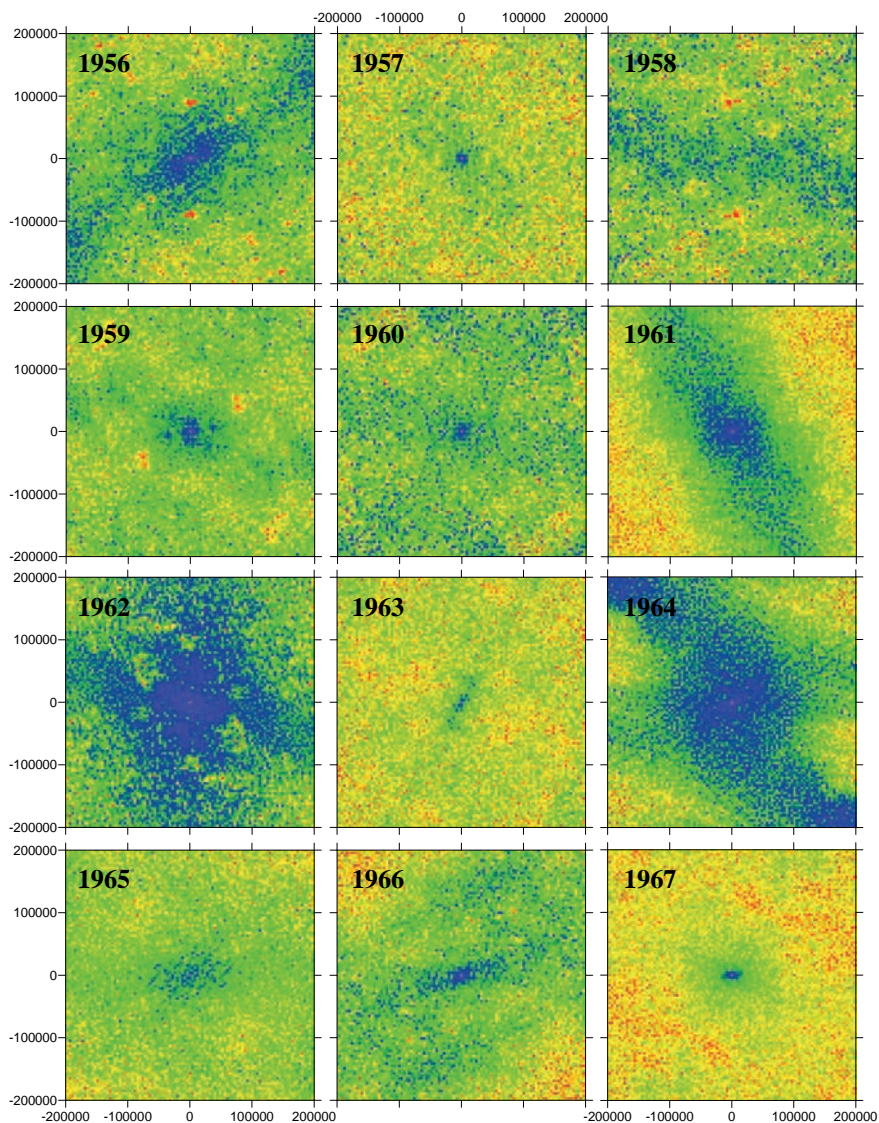



**Ryc. 131.** Schemat powierzchni złożonych modeli semiwariancji anizotropowych przy geometrycznym (A) i geometryczno-strefowym (B) typie anizotropii: a - mapa rastrowa w projekcji płaskiej, b, c, d - obraz w rzucie perspektywicznym, odpowiednio: mapa rastrowa, izoliniowa i siatkowy model trójwymiarowy 

Przy anizotropii geometrycznej „lejek” jest spłaszczony – ma w każdej płaszczyźnie równoległej do osi odległości przekrój elipsy. Każdy element<sup>59</sup> anizotropowego semiwariogramu złożonego może mieć inny kierunek i stosunek anizotropii – elipsy będą się wtedy różnić orientacją i mimośrodem (ryc. 131A). Kiedy semiwariogramy kierunkowe różnią się wariancją progową, lecz mają ten sam zasięg (anizotropia strefowa), wtedy kołnierz lejka niepodobieństwa nie stanowi płaszczyzny równoległej od osi odległości, lecz tworzy sfalowaną powierzchnię o kształcie siodła, a jego przekroje są okręgami. Najbardziej skomplikowany kształt występuje przy złożonej anizotropii geometryczno-strefowej – lejek ma segmenty różniące się kierunkiem i/lub stopniem spłaszczenia oraz sfalowany kołnierz (ryc. 131B). Dla

<sup>59</sup> Poza składową nuggetową.

uproszczenia powierzchnię semiwariogramu przedstawia się najczęściej na płaszczyźnie w postaci mapy izolinowej lub pikselowej (rastrowej). W tej drugiej kolorem lub natężeniem szarości obrazuje się wartość semiwariancji dla danego przedziału odległości wzdłuż osi X i Y.



**Ryc. 132.** Powierzchnie (mapy) semiwariogramów dla lat 1956-1967. Zakres wartości przedstawionych na każdej mapie jest różny. Każdy piksel ma 4 na 4 km. Szczegółowe objaśnienia w tekście 

Jedną z najistotniejszych decyzji metodologicznych, jaką musiano podjąć, dotyczyła uwzględniania lub pominięcia anizotropii przy analizie struktury przestrzennej MSDO. Indywidualne opady wykazują najczęściej bardzo silną anizotropię (stosunek zasięgów zbliżony lub mniejszy niż 0,5), a jej azymut związany jest z kierunkiem adwekcji mas powietrza (Bacchi, Kottogoda 1995, Berne i in. 2006, Moszkowicz 2000). Wraz ze wzrostem długości czasu kumulacji opadów (doba, dekada, miesiąc, rok), i powierzchni analizowanego obszaru, maleje rola pojedynczych opadów, a pojawia się i zaczyna dominować wpływ kierunków oraz natężenia ogólnej cyrkulacji strefowej. Takiej prawidłowości można się spodziewać prowadząc analizę „pełnych” serii pomiarowych. Dane MSDO również w tym kontekście są specyficzne. Miesięczne i roczne MSDO to połączone, bardzo zmienne pod względem proporcji w stosunku do całości, podzbiory wielu sum dobowych zarejestrowanych na obszarze Polski w różnych terminach. Każdy z tych opadów charakteryzował się odmienną anizotropią wynikającą z jego genezy i przebiegu. Można było się spodziewać, że będzie to skutkowało, szczególnie w wypadku MSDO rocznych, występowaniem mało zróżnicowanej „mieszanki”. Z drugiej strony, losowy charakter wysokich opadów mógł spowodować, że w przypadku danych miesięcznych pojedynczy, silnie anizotropowy, opad zanotowany w wielu stanowiskach zdominuje obraz struktury przestrzennej uzyskany dla całej Polski. Wszystko to przemawiało za pominięciem w analizie aspektu zróżnicowania kierunkowego MSDO.

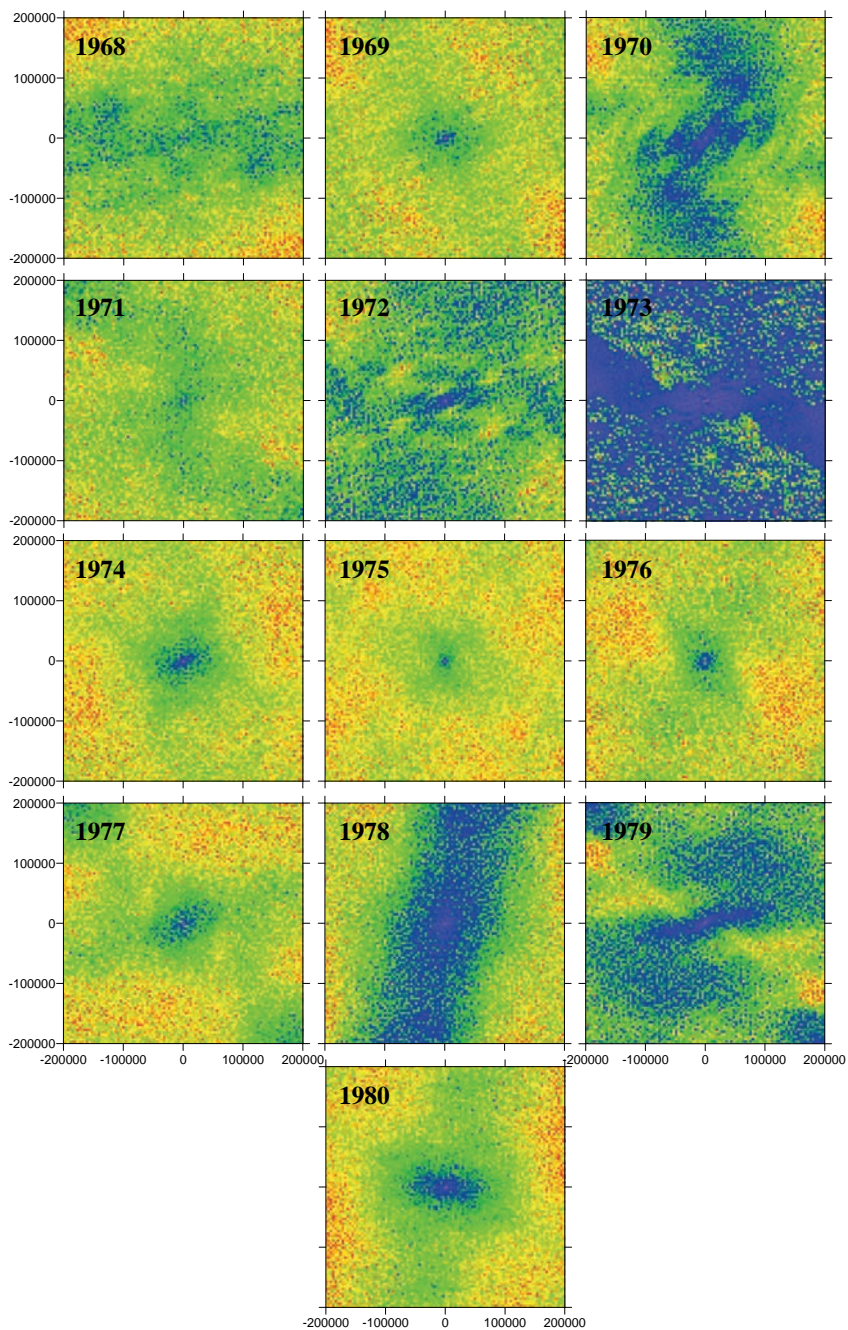
Potwierdzeniem powyższych spekulacji są mapy rastrowe powierzchni semiwariogramów wykonane dla wszystkich rocznych MSDO oraz miesięcznych z jednego przykładowego roku z analizowanego wielolecia (ryc. 132-134). Wartości semiwariancji obliczono dla 100 odstępów po 4 km wzdłuż osi W - E (X) i N - S (Y). Każda wartość semiwariancji w komórce mapy jest średnią z około 50-510 par danych<sup>60</sup>. Przeciętna liczba par, z których obliczona była wartość semiwariancji przekraczała 100, co zapewnia dostateczny poziom reprezentatywności statystycznej (Webster, Oliver 1992). W zastosowanej standardowej, tęczowej palecie barw kolorem niebieskim zaznaczono niskie wartości semiwariancji, a więc wysokie podobieństwo danych, kolorem czerwonym - maksymalne jej wartości, czyli najmniejsze podobieństwo.


Anizotropia struktury przestrzennej rocznych MSDO (ryc. 132 i 133) w 9 latach z analizowanego wielolecia w ogóle się nie zaznacza (lata 1957, 1959, 1960, 1963, 1965, 1967, 1969, 1974, 1975), a w kolejnych 9 jest bardzo słabo zarysowana (lata 1958, 1962, 1966, 1968, 1971, 1972, 1973, 1976, 1977). W pozostałych 7 (lata 1956, 1961, 1964, 1970, 1978, 1979, 1980) jest lepiej widoczna, ale ma raczej charakter mocno „rozmażany”, świadcząc o złożonym,

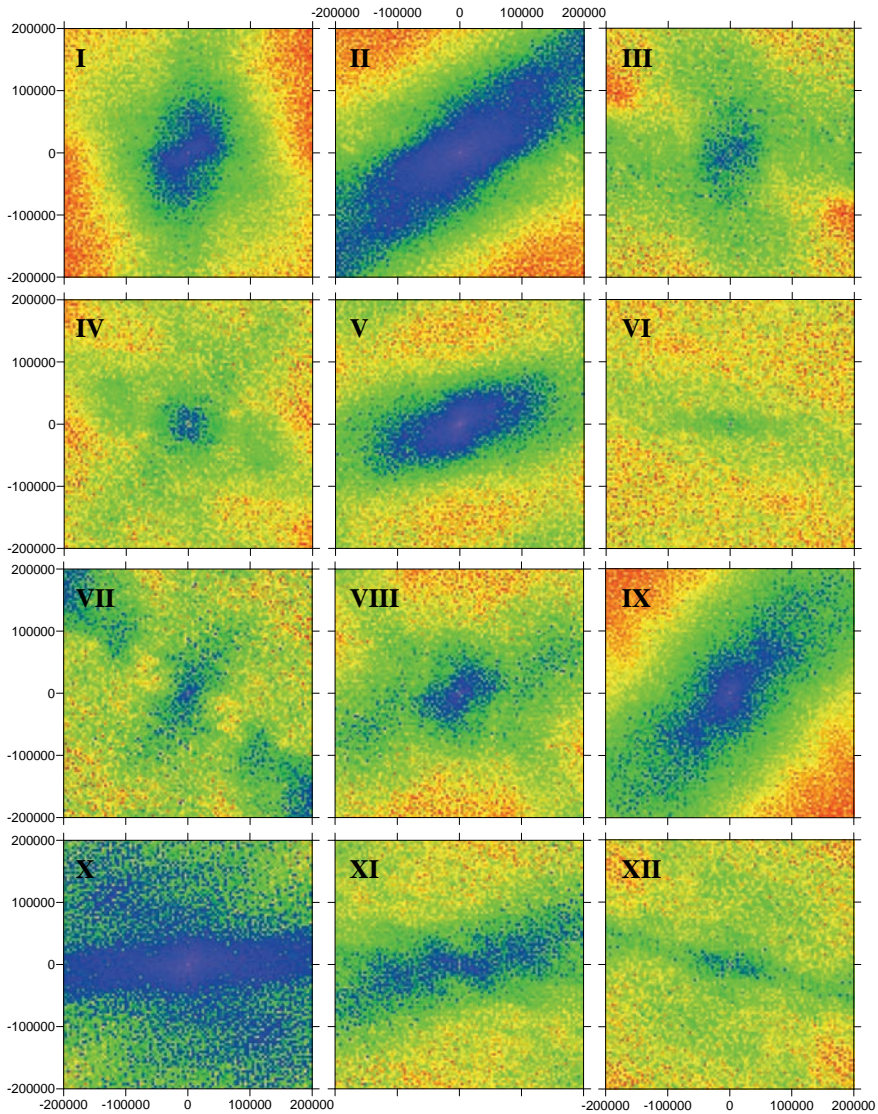
---


<sup>60</sup> Z wyjątkiem komórki centralnej, w której zawsze było ponad 2000 par.





**Ryc. 133.** Powierzchnie (mapy) semiwariogramów dla lat 1968-1980. Zakres wartości przedstawionych na każdej mapie jest różny. Każdy piksel ma 4 na 4 km. Szczegółowe objaśnienia w tekście 



**Ryc. 134.** Powierzchnie (mapy) semiwariancji dla poszczególnych miesięcy roku 1977. Zakres wartości przedstawionych na każdej mapie jest różny. Każdy piksel ma 4 na 4 km. Szczegółowe objaśnienia w tekście 

heterogenicznym charakterze analizowanego pola. Wyraźnie jednak dominują układy południkowe NW – SE (lata 1961, 1964, 1979, 1978, 1979, 1980). W drugiej grupie lat anizotropia ma głównie charakter geometryczny, w trzeciej zaś – geometryczno-strefowy.



Na mapach semiwariogramów miesięcznych MSDO z roku 1977 obraz jest nieco inny (ryc. 134). Na większości z nich anizotropia zaznacza się bardzo wyraźnie, ale ma zazwyczaj charakter „przecinających” się elips (anizotropia geometryczna) bądź pasów (anizotropia strefowa). Świadczy to o występowaniu o obrębie analizowanego obszaru podzbiorów danych pomiarowych o różnym ukierunkowaniu rozkładu zmierzonych MSDO i różnym zasięgu. Jedynie w lutym 1977 rozkład przestrzenny maksymalnych opadów dobowych charakteryzował się silną i w miarę jednorodną anizotropią strefową. Przykład losowo wybranego z całego wielolecia roku potwierdza przypuszczenie o skomplikowanym charakterze kierunkowych relacji podobieństwa miesięcznych MSDO, wynikającym z ich niesynchroniczności i ograniczonego, do niedużej części terytorium Polski, zasięgu.

Nie można również pominąć bardzo ważnego względu praktycznego. O ile obliczanie empirycznych kierunkowych miar struktury przestrzennej nie jest specjalnie kłopotliwe i nie wydłużyłoby w sposób znaczący czasu pracy, to etap modelowania byłby radykalnie odmienny. Tworzenie, za pomocą stosowanej w niniejszej rozprawie metody półautomatycznej (por. podrozdz. III.2.5.3 – III.2.5.6), złożonego modelu anizotropowego trwa kilka razy dłużej niż modelu izotropowego. Ponieważ w sumie wykonano 4550 modeli izotropowych i zajęło to ponad pół roku bardzo intensywnej pracy, uwzględnienie anizotropii uniemożliwiłoby zakończenie opracowania w okresie wynikającym z realizowanego projektu badawczego. Problem ten wymaga jednak dalszej uwagi i będzie tematem osobnego opracowania przy zastosowaniu odmiennej, nieparametrycznej metodyki (Yao, Journal 1998, Ma, Yao 2001).

## 4. Normalizacja danych

### 4.1. Ogólna charakterystyka procedur normalizacji danych wykorzystywanych w geostatystyce

Istotnym etapem procedury symulacji metodą pola prawdopodobieństwa (ang. *p-field*)<sup>61</sup> była normalizacja danych. Używane one były do opracowania modelu struktury przestrzennej symulowanego zjawiska (patrz rozdz. III.2.5.7 i VI). Model ten był następnie wykorzystany do wygenerowania inicjalnego, bezwarunkowego pola prawdopodobieństwa<sup>62</sup>. Prze-

---

<sup>61</sup> Element ten, jak wspomniano we wprowadzeniu (rozdz. I) i omówieniu metodyki (rozdz. III.2, dodatek X.2), zostanie zaprezentowany w odrębnej publikacji.

<sup>62</sup> Bezwarunkowa symulacja przestrzenna oparta jest jedynie na modelu struktury przestrzennej analizowanej cechy. Nie są brane pod uwagę wyniki ewentualnych pomiarów.

strzennie skorelowane liczby losowe były potrzebne do próbkowania uży- skanych metodą krigingu wartości kodowanych warunkowych krzywych rozkładu prawdopodobieństwa (patrz dodatek X.2). Taka procedura za- pewniała uzyskanie dwóch warunków nakładanych na wszystkie metody symulacji przestrzennej: a) zapewnienie zgodności rozkładu statystycznego danych symulowanych i empirycznych; b) zapewnienie zgodności struktury przestrzennej danych symulowanych i empirycznych. Normalizacja danych ma jednak w geostatystyce znacznie szersze zastosowania niż tylko w kon- tekście wykorzystywanej w niniejszej pracy metodyki (Chilés, Delfiner 1999, Goovaerts 1997). Nie należy jej mylić ze standaryzacją i/lub normalizacją w sensie statystycznym (StatSoft Inc., 2004). W naukach o Ziemi zmienne po- miarowe wykazują często asymetryczny rozkład z niewielką ilością bardzo dużych wartości (skośność dodatnia, ryc. 138A). W wielu procedurach geo- statystycznych wymagane jest jednakże, albo bardzo korzystne, aby prze- twarzane dane miały rozkład normalny. Normalizacja danych to nieliniowa odwracalna transformacja jakiegokolwiek ciągłej skumulowanej funkcji roz- kładu (ang. *cumulative distribution function* = *cdf*) w standaryzowany rozkład gaussowski o średniej równej zero i odchyleniu standardowym równym je- den<sup>63</sup>. Odwracalność transformacji zakłada możliwość uzyskania z powro- tem z danych znormalizowanych oryginalnego, nieobciążonego *cdf*<sup>64</sup>.

W geostatystyce stosowane są w praktyce dwa sposoby normalizacji da- nych: nieparametryczny i parametryczny. Dają one w zasadzie wyniki rów- noważne. Metoda nieparametryczna, zwana najczęściej *normal score trans- form* (Deutsch, Journel 1998, Goovaerts 1997, Olea 1999), polega na prostym, „graficznym” przekodowaniu kwantyli empirycznego rozkładu posiada- nych danych na odpowiednie kwantyle standaryzowanego rozkładu nor- malnego. Zachowana tablica przekodowania umożliwi dokonanie w razie potrzeby transformacji odwrotnej. Taki sposób normalizacji danych jest do- stępny w pakiecie GSLIB (Deutsch, Journel 1998). Słabością metody niepa- rametrycznej jest konieczność osobnego przeprowadzenia wygładzania roz- kładu empirycznego, kiedy próbka jest zbyt mała i/lub ze względu na duże błędy pomiarowe jest on bardzo chaotyczny (szczególnie przy ogonach roz- kładu). Parametryczna alternatywa jest tej wady pozbawiona. Wygładzanie rozkładu empirycznego jest jej integralną składową. Z tego względu właśnie ona została użyta w niniejszym opracowaniu.

---

<sup>63</sup> W ujęciu bardziej ogólnym, transformację jakiegokolwiek ciągłego *cdf* w dowolny inny *cdf*.

<sup>64</sup> „Normalizację” prawoskośnych danych bardzo często uzyskuje się przez logarytmowa- nie. Rzadko jednak jest to procedura w pełni skuteczna, to jest całkowicie „likwidująca” sko- śność. Poza tym, ma dwie poważne wady: a) może być zastosowana jedynie do danych do- datnich (wymaga przeskalowania wartości, jeśli są w nich równe lub mniejsze od zera); b) transformacja odwrotna (antylogarytmowanie) jest obciążona (Goovaerts 1997).

## 4.2. Normalizacja parametryczna z użyciem wielomianów Hermite

Normalizacja parametryczna zwana jest często w literaturze „anamorfozą gaussowską” (ang. *Gaussian anamorphosis*, Chilés, Delfiner 1999, Webster, Oliver 2001, Wackernagel 2003). Dokonuje się ona z użyciem serii wielomianów Hermite (Weisstein 2006). Punktem wyjścia jest chęć znalezienia takiej funkcji  $\Phi$ , która umożliwiłaby transformację dowolnego rozkładu empirycznego  $Z(\mathbf{x})$  w standardowy rozkład normalny  $Y(\mathbf{x})$ :

$$Y(\mathbf{x}) = \Phi[Z(\mathbf{x})] \quad [33]$$

Jej odwrotność z kolei  $\Phi^{-1}$  umożliwia dokonanie konwersji w drugą stronę:

$$Z(\mathbf{x}) = \Phi^{-1}[Y(\mathbf{x})] \quad [34]$$

Wzór na funkcję gęstości prawdopodobieństwa standardowego rozkładu normalnego jest następujący:

$$g(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \quad [35]$$

Wielomiany Hermite są powiązane z powyższą funkcją i mogą być przedstawione w postaci:

$$H_k(y) = \frac{1}{\sqrt{k!g(y)}} \cdot \frac{d^k g(y)}{dy^k} \quad [36]$$

gdzie  $k$  jest stopniem wielomianu mogącym przybrać wartości  $0, 1, 2, \dots$ , a  $1/\sqrt{k!}$  – współczynnikiem standaryzacji. Dwa pierwsze wielomiany Hermite, to jest dla  $k = 0$  i  $k = 1$  są następujące:

$$H_0(y) = 1, \quad [37]$$

$$H_1(y) = -y, \quad [38]$$

i dlatego wielomiany wyższego stopnia stosują się do zasady rekurencyjnej:

$$H_k(y) = -\frac{1}{\sqrt{k}} y H_{k-1}(y) - \sqrt{\frac{k-1}{k}} H_{k-2}(y). \quad [39]$$

Tak więc, wielomiany te mogą być dla standardowego rozkładu normalnego obliczane do dowolnego stopnia.

Wielomiany Hermite są ortogonalne w odniesieniu do funkcji wazącej  $\exp(-y^2/2)$  w przedziale od  $-\infty$  do  $+\infty$ . Są one zatem niezależnymi składnikami rozkładu normalnego o ciągle rosnącej, wraz ze stopniem wielomianu, złożoności. Praktycznie każda funkcja  $Y(\mathbf{x})$  może być przedstawiona jako ich suma:

$$f\{Y(\mathbf{x})\} = f_0 H_0\{Y(\mathbf{x})\} + f_1 H_1\{Y(\mathbf{x})\} + f_2 H_2\{Y(\mathbf{x})\} + \dots, \quad [40]$$

a ponieważ wielomiany Hermite są ortogonalne:

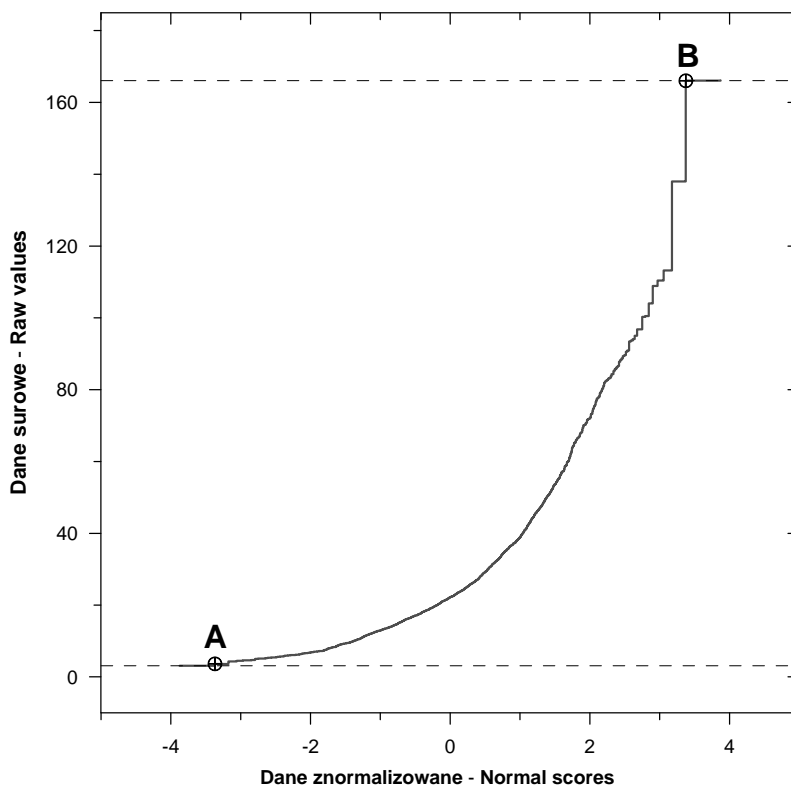
$$\begin{aligned} E\left[f\{Y(\mathbf{x})\} H_k\{Y(\mathbf{x})\}\right] &= E\left[H_k\{Y(\mathbf{x})\} \sum_{l=0}^{\infty} f_l H_l\{Y(\mathbf{x})\}\right] \\ &= \sum_{l=0}^{\infty} f_l E\left[H_l\{Y(\mathbf{x})\} H_k\{Y(\mathbf{x})\}\right] \\ &= f_k \end{aligned} \quad [41]$$

Pozwala to obliczyć współczynniki  $\phi_k$  funkcji  $\Phi[Y(\mathbf{x})]$  z równania [33] jako:

$$\begin{aligned} Z(x) &= \Phi[Y(\mathbf{x})] \\ &= \phi_0 H_0\{Y(\mathbf{x})\} + \phi_1 H_1\{Y(\mathbf{x})\} + \phi_2 H_2\{Y(\mathbf{x})\} + \dots \\ &= \sum_{k=0}^{\infty} \phi_k H_k\{Y(\mathbf{x})\} \end{aligned} \quad [42]$$

Transformacja dokonana z użyciem wielomianów Hermite jest oczywiście w pełni odwracalna [wzór 34], co oznacza, że wyniki obliczeń na danych znormalizowanych mogą być w każdym momencie przekonwertowane do oryginalnej skali pomiarowej.

Szczegóły algorytmu wyznaczania współczynników wielomianów Hermite opisane są przez Webstera i Oliver (2001) oraz Wackernagla (2003). W niniejszym opracowaniu anamorfozę gaussowską wykonano w programie ISATIS (Bleinès i in. 2007). Modelowanie anamorfozy rozpoczyna się od konstrukcji dyskretnej krzywej rozkładu danych surowych (pomiarowych, ryc. 135). Użytkownik ma możliwość modyfikowania kształtu modelu na jego krawędziach poprzez podanie wartości 2 punktów kontrolnych (A i B na ryc. 135). Ta opcja ma znaczenie tylko wtedy, kiedy liczba pomiarów jest niewielka. Poza tym, można podać interwał danych pomiarowych (ang. *Authorized Interval on the Raw Variable*), zdefiniowany jako różnica pomiędzy wartością minimum  $Z_{\min}$  i maksimum  $Z_{\max}$ , oraz rząd rozwinięcia wielomianów Hermite (liczba wielomianów). Domyślnymi wartościami interwału są minimum i maksimum danych empirycznych. Przy takiej konfiguracji dwa wymienione wyżej punkty kontrolne nie powodują zmiany krzywej empirycznej.

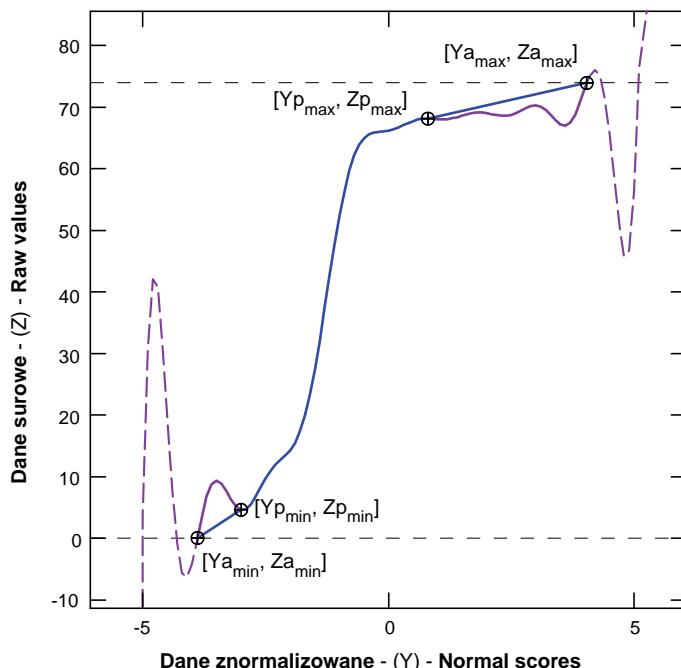


Ryc. 135. Usytuowanie dwóch punktów kontrolnych anamorfozy gaussowskiej z użyciem wielomianów Hermite (Bleinès i in. 2007). Szersze objaśnienia w tekście

Po zdefiniowaniu warunków anamorfozy dyskretnej program oblicza współczynniki  $\psi_i$  rozwinięcia wielomianów Hermite. Wykreślana jest krzywa i obliczany „praktyczny” oraz „absolutny interwał definicji” (ang. *Practical Interval of Definition*, *Absolute Interval of Definition*, ryc. 136). Stąd, wiadomo, że granicami:

- „praktycznego interwału definicji” są punkty  $[Y_{p_{\min}}, Z_{p_{\min}}]$  i  $[Y_{p_{\max}}, Z_{p_{\max}}]$  (ryc. 136). Są to miejsca, gdzie modelowana krzywa przecina górny i dolny kres interwału danych pomiarowych ( $Z_{a_{\min}}, Z_{a_{\max}}$ ) lub punkty, w których wartość na krzywej empirycznej nie zmienia się przy zmianach  $Y$ ;

- „absolutnego interwału definicji” są punkty  $[Y_{a_{\min}}, Z_{a_{\min}}]$  i  $[Y_{a_{\max}}, Z_{a_{\max}}]$  (ryc. 136). Zlokalizowane są one na przecięciu modelowanej krzywej z horyzontalnymi liniami określającymi interwał danych pomiarowych. Wartości generowane przez funkcję anamorfozy poza „absolutnym interwałem definicji” nie powinny być w żadnym przypadku uwzględniane. Krzywa wykazuje bowiem tam najczęściej chaotyczne fluktuacje (ryc. 136 i 137).



Ryc. 136. Schemat określający „praktyczny” i „absolutny interwał definicji” modelu anamorfozy gaussowskiej (Bleinès i in. 2007). Dokładny opis w tekście

Może to stanowić istotne ograniczenie w sytuacji, kiedy symulowane w przestrzeni danych znormalizowanych wartości wykraczają poza „absolutny interwał definicji”. Próba konwersji do oryginalnej skali pomiarowej daje wtedy nierealistyczne wyniki. Istnieje jednak podawane przez Wackernagla (2003) proste rozwiązanie tego problemu.

W praktycznej realizacji anamorfoza jest zdefiniowana jako funkcja  $Y: Z(\mathbf{x}) = \Phi[Y(\mathbf{x})]$  (czyli odwrotnie niż we wzorach [33] i [34]), więc aby dokonać transformacji danych surowych w gaussowskie, należy dokonać jej inwersji do postaci  $Y(\mathbf{x}) = \Phi^{-1}[Z(\mathbf{x})]$ . W programie ISATIS można dokonać tego trzema sposobami (Bleinès i in. 2007), z których zalecana jest inwersja na bazie frekwencji (ang. *frequency inversion*). Na początek program sortuje wartości pomiarowe. Dla każdej próbki  $FC_i$ , poczynając od najmniejszej, obliczana jest skumulowana frekwencja [43]:

$$FC_i = FC_{i-1} + W_i \quad [43]$$

Zmienna wążąca  $W_i$  jest podawana przez operatora lub obliczana jako  $W_i = 1/N$ . Dzięki temu dwie próbki o identycznych wartościach empirycznych

nych będą miały inne skumulowane frekwencje. Finalne obliczenie danej znormalizowanej polega na wyznaczeniu średniej arytmetycznej dwóch kolejnych wyrazów szeregu:

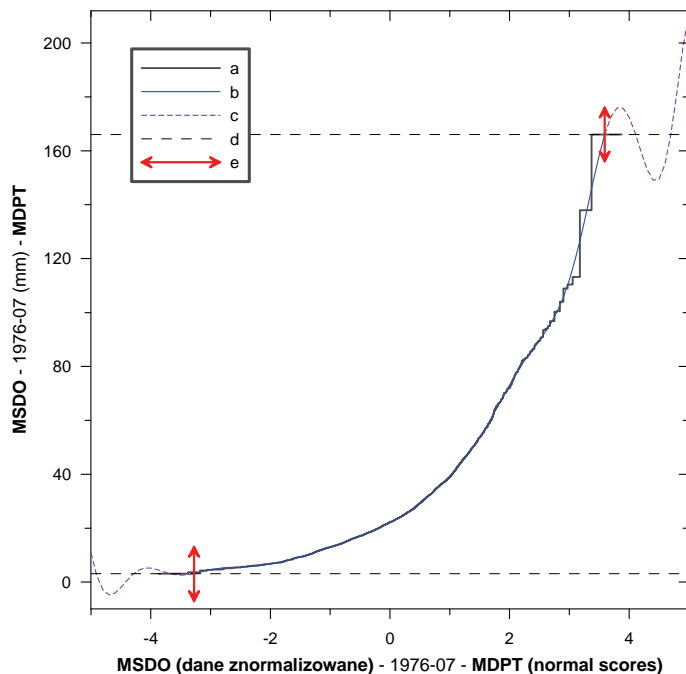
$$Y_i = \frac{\Phi^{-1}(FC_i) + \Phi^{-1}(FC_{i-1})}{2} \quad [44]$$


Konsekwentnie zatem dwie identyczne dane pomiarowe będą miały odmienne wartości znormalizowane.

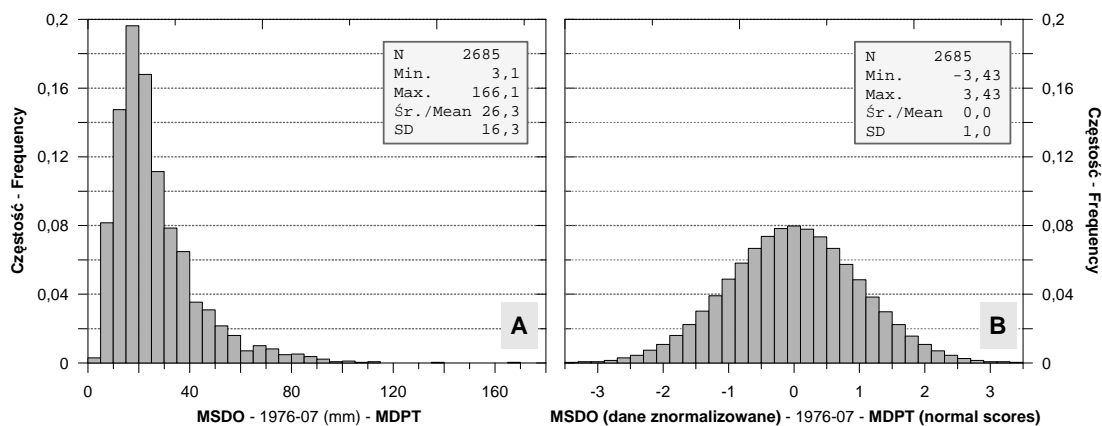
### 4.3. Przykład normalizacji danych MSDO

Do zilustrowania zagadnienia normalizacji danych użyto zbioru miesięcznych MSDO z lipca 1976. Charakteryzowały się one typowym skośnym rozkładem, z wysoką koncentracją wartości niskich i średnich (62,3% w przedziale 10-30 mm) i maksimum 166,1 mm (ryc. 138A, tab. 30). Anamorfozę przeprowadzono posługując się standardowym ustawieniem opcji, to jest z interwałem danych pomiarowych (ang. *Authorized Interval on the Raw Variable*) zgodnym z rzeczywistymi (empirycznymi) wartościami minimum i maksimum analizowanego zbioru danych. Przyjęto również 30 rząd rozwinięcia wielomianów Hermite, zalecany, w programie ISATIS, jako najlepszy w większości przypadków. W praktyce optymalne rozwiązanie mieści się zazwyczaj w zakresie 15-50 rzędu (Wackernagel 2003). Możliwe jest iteracyjne poszukiwanie najlepszego wariantu, lecz nie ma to dużego wpływu na jakość transformacji, a w znaczący sposób wydłuża obliczenia. Uzyskane znormalizowane *cdf* (ryc. 137) rozciąga się w przedziale od -3,43 do +3,43. Dopasowanie modelu do danych empirycznych, do granicy około 90 mm, jest prawie idealne. Powyżej nieciągłe, „schodkowe” *cdf* empiryczne zostało przez funkcję anamorfozy wygładzone.

O skuteczności transformacji świadczy nie tylko uzyskanie zbioru danych o średniej równej zero i odchyleniu standardowym równym jeden oraz symetrycznego histogramu (ryc. 138B, tab. 30). Bardziej wymowne są inne charakterystyki ilościowe – skośność, która w wyjściowej próbce MSDO z lipca 1976 wynosiła 1,905 po transformacji równa się dokładnie 0. Kurtosis zmniejszyła się z 5,874 w oryginalnym zbiorze danych do -0,021. Prawdopodobieństwo normalności rozkładu oszacowane z pomocą testu Shapiro-Wilka wynosiło dla wyjściowych danych mniej niż  $2,2 \cdot 10^{-16}$ , a po ich transformacji – dokładnie 1.



**Ryc. 137.** Przykład normalizacji danych (ang. *Gaussian anamorphosis*) MSDO z lipca 1976: a – skumulowana empiryczna funkcja rozkładu (*cdf*), b – krzywa rozkładu danych znormalizowanych uzyskana z użyciem 30 wielomianów Hermita, c – nierealistyczna ekstrapolacja rozkładu poza zasięgiem danych empirycznych, d – granice rozkładu danych empirycznych (3,1-166,1 mm), e – granice przedziału danych znormalizowanych (od -3,43 do 3,43) 



**Ryc. 138.** Histogramy i statystyki opisowe MSDO z lipca 1976 roku: A – surowe dane pomiarowe w mm, B – dane znormalizowane według metodyki opisanej w tekście



**Tabela 30.** Statystyki opisowe danych MSDO z lipca 1976: VII-76 – oryginalne dane pomiarowe, VII-76n – dane znormalizowane, VII-76bt – dane po konwersji zwrotnej

Parametr - Parameter	Zmienna - Variable		
	VII-76	VII-76n	VII-76bt
Średnia - Mean	26,269	0,000	26,266
Błąd standardowy - Standard Error	0,314	0,019	0,314
Mediana - Median	22,2	0,0	22,0
Odchylenie standardowe - Standard Deviation	16,283	1,000	16,251
Kurtoza - Kurtosis	5,874	-0,021	5,309
Skośność - Skewness	1,905	0,000	1,863
Zakres - Range	163,0	6,86	148,8
Minimum - Minimum	3,1	-3,43	3,1
Maksimum - Maximum	166,1	3,43	151,9

Sprawdzono również jakość konwersji odwrotnej, to jest z powrotem z danych znormalizowanych do oryginalnej skali pomiarowej. Statystyki opisowe wszystkich trzech zbiorów danych (oryginalne, znormalizowane, po konwersji zwrotnej) zestawiono w tabeli 30. Wygładzenie górnej części rozkładu empirycznego poskutkowało niewielkim obniżeniem odchylenia standardowego, skośności i kurtozy przy zachowaniu praktycznie idealnym średniej. Maksimum zmalało o 14,2 mm. Przekształcenie do oryginalnej skali jest rzeczywiście nieobciążone – średnie odchylenie tylko minimalnie różni się od zera (-0,003 mm). Średnie absolutne odchylenie (RMSE) wynosi 0,383 mm, a współczynnik korelacji równa się 1.

#### 4.4. Wpływ normalizacji danych na ocenę ich struktury przestrzennej

Normalizacja danych nie zmienia w najmniejszym stopniu ich struktury przestrzennej reprezentowanej przez semiwariogram. Wręcz przeciwnie, taka transformacja korzystnie zmniejsza czułość semiwariogramu na dane ekstremalne, potencjalnie błędne, lub należące do osobnej populacji. Właściwość ta została zilustrowana na rycinie 139. Przedstawia ona cztery semiwariogramy empiryczne policzone dla MSDO z lipca roku 1976. Krzywe „a” i „e” zostały wyliczone z całego zbioru danych (2685 przypadków), odpowiednio dla danych surowych i znormalizowanych. Krzywe „c” i „g” to z kolei semiwariogramy jak wyżej, ale uzyskane tym razem z danych wyselekcjonowanych (2672 przypadków) po procedurze interaktywnego czyszczenia (patrz dodatek X.5). Układ obu skal rzędnych (Y) został tak dobrany, aby uzyskać jak najlepszą zgodność przebiegu krzywych „c” i „g”. Takie zestawienie wykresów pokazuje dobitnie, że sama normalizacja całego zbioru danych (krzywa „e”) daje już ob-

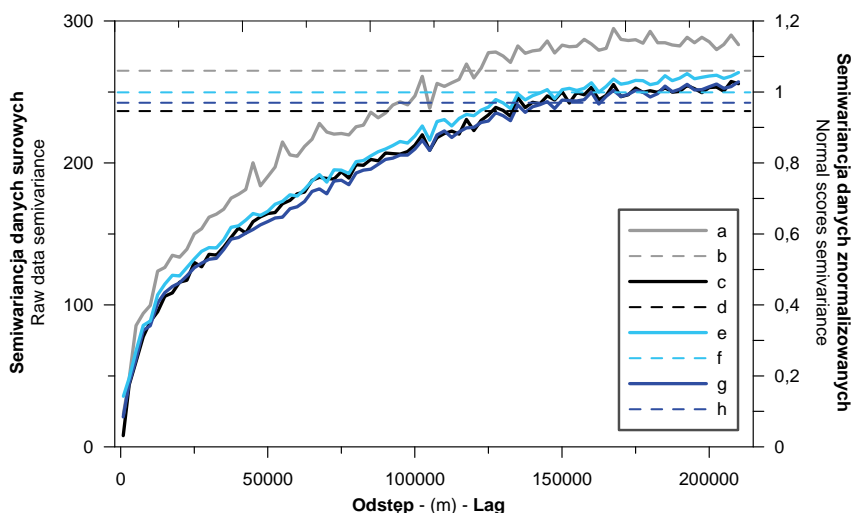
raz struktury przestrzennej bardzo zbliżony do uzyskanego po wyeliminowaniu 13 odstających przypadków (krzywe „c” i „g”). Różni się on znacznie od wyjściowego semiwariogramu pełnego zbioru danych surowych (krzywa „a”).

Zaletą normalizacji danych empirycznych jest także możliwość bezpośredniego porównywania struktury przestrzennej wielu zbiorów danych różniących się skalą pomiarów i/lub zakresem ich zmienności. Ominięcie tego etapu przygotowania danych spowodowałoby konieczność wprowadzenia innej procedury – standaryzacji semiwariogramów empirycznych i ich modeli<sup>65</sup>.

Cenną właściwością anamorfozy gaussowskiej z użyciem wielomianów Hermite jest możliwość obliczenia, na podstawie współczynników wielomianów ( $\psi_i$ ), wariancji danych surowych (Bleinès i in. 2007). Dokonuje się tego na podstawie prostej relacji pomiędzy kowariancją znormalizowanych danych a kowariancją danych surowych:

$$C(\mathbf{h}) = \sum_{i=1}^{\infty} \psi_i^2 \rho^i(\mathbf{h}) \quad [45]$$

gdzie:  $\rho(\mathbf{h})$  – kowariancja danych znormalizowanych, a  $C(\mathbf{h})$  – kowariancja danych surowych.



**Ryc. 139.** Struktura przestrzenna danych MSDO z lipca 1976: a, b – semiwariogram i wariancja całego zbioru danych surowych (2685 przypadków), c, d – semiwariogram i wariancja danych surowych wyselekcjonowanych (2672 przypadków), e, f – semiwariogram i wariancja całego zbioru danych po normalizacji, g, h – semiwariogram i wariancja wyselekcjonowanych danych znormalizowanych

<sup>65</sup> Wykonuje się to poprzez podzielenie wartości semiwariogramów empirycznych i ich modeli przez wariancję próby.

Relacja ta jest spełniona wówczas, kiedy para zmiennych ( $Z(x)$ ,  $Z(x + \mathbf{h})$ ) może być uważana za dwuzmienną normalną (ang. *bivariate normal*, Goovaerts 1997). Na podstawie zależności istniejącej pomiędzy kowariancją  $C(\mathbf{h})$  a semiwariancją  $\gamma(\mathbf{h})$ :

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h}) \quad [46]$$

gdzie  $C(0)$  oznacza wariancję danych, a zatem można uzyskać formułę na konwersję semiwariogramów pomiędzy przestrzenią danych znormalizowanych a wyjściową skalą pomiarową. Jest to cenna właściwość w kontekście zademonstrowanej powyżej mniejszej czułości semiwariogramu zmiennej gaussowskiej na wartości odstające i ekstremalne, a co za tym idzie – prostszym jego modelowaniu. W niniejszej pracy nie było jednakże potrzeby korzystania z tej opcji. Zachowane parametry anamorfozy gaussowskiej (współczynniki wielomianów Hermite) umożliwiają jednak jej zastosowanie w razie potrzeby w przyszłości.

## 5. Maskowanie danych w trakcie analizy struktury przestrzennej

### 5.1. Wpływ danych ekstremalnych i odstających na określenie struktury przestrzennej zjawiska

Podobnie jak inne statystyki będące pochodną wariancji, semiwariancja jest czuła na wartości ekstremalne (ze względu na operację potęgowania, Kerry, Oliver 2007). W wypadku semiwariancji negatywny wpływ na ustalenie prawidłowości struktury przestrzennej mają także wartości lokalnie odstające<sup>66</sup>. Pojedyncze wyniki pomiarów mogą zniekształcić wynik analizy dla setek, czy tysięcy danych. Jest to szczególnie znaczące dla 1-3 pierwszych odstępów, które zazwyczaj zawierają znacznie mniejszą liczbę par punktów niż kolejne. Aby uniknąć takich sytuacji, po upewnieniu się, że wartości owe nie są po prostu efektem błędów pomiarowych, stosuje się różne zabiegi (Goovaerts 1997).

Jeśli wartości ekstremalne wykazują sąsiedztwo w przestrzeni, można je uznać za osobną populację i wyłączyć z analizy. Stosuje się również transformację danych (np. logarytmiczną) zmniejszającą skalę zmienności, a przez to redukującą wpływ wartości ekstremalnych. Kolejnym wariantem jest wykorzystanie takich miar struktury przestrzennej mniej czułych na obecność wartości odstających i ekstremalnych, jak madogram czy rodogram (Deutsch, Journel 1998, Goovaerts 1997). Ostatnią możliwością stanowi

---

<sup>66</sup> Są to takie wartości, które nie stanowiąc ekstremów analizowanej próby, wyraźnie odbiegają od sąsiadujących w przestrzeni i/lub w czasie wyników pomiarów.

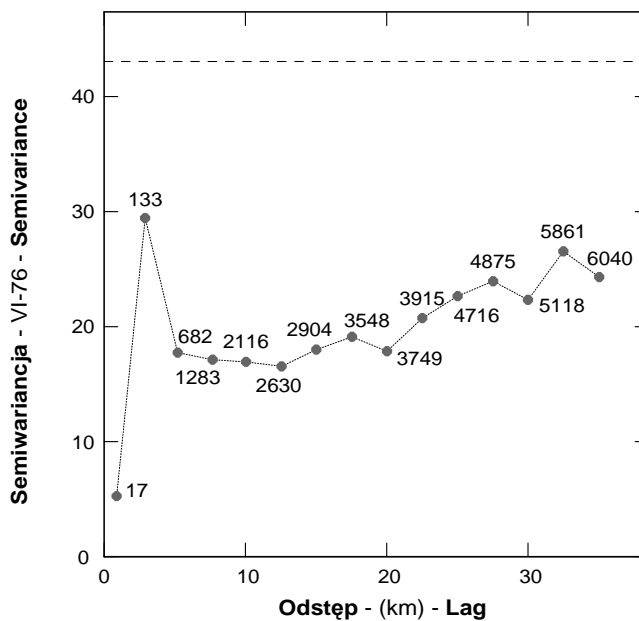
tw. interaktywne czyszczenie semiwariogramu (Goovaerts 1997, Pannatier 1996). Wykorzystuje się do tego celu wykresy tzw. chmur semiwariogramu (ang. *variogram cloud*) lub rozrzutu z przesunięciem (ang. *h-scatterplots*). Pierwszy z nich przedstawia wszystkie wyliczone wartości kwadratów różnic między parami pomiarów jako funkcję dzielącą je odległości, natomiast drugi – relację między wartościami poszczególnych par pomiarów dla jednego określonego przedziału dzielących je odległości. Procedura polega na wizualnej identyfikacji na wymienionych powyżej wykresach odstających punktów – czyli par pomiarów. Są one następnie przez operatora zaznaczane przy wykorzystywaniu graficznego interfejsu programu komputerowego i maskowane. Maskowanie polega na tymczasowym ich usuwaniu z obliczeń semiwariogramu (przeliczaniu ponownie wartości semiwariogramu z wyłączeniem zamaskowanych wartości). Usunięcie pary danych z obliczeń konkretnej wartości semiwariogramu nie oznacza ich całkowitej eliminacji z analizowanego zbioru. Te dwa wyniki pomiarów wciąż mogą być użyte do obliczeń w innych parach tego samego przedziału odległości lub innych przedziałów. Tego typu procedura ma liczne zalety, szczególnie przy analizowaniu małych zbiorów danych. W niniejszym opracowaniu dysponowano jednak bardzo licznymi, jak na typową analizę geostatystyczną, zbiorami pomiarów. Dlatego, wykorzystując program ISATIS (Bleinès i in. 2007), czyszczenie semiwariogramu wykonano drugim sposobem (na podstawie chmury semiwariogramu), eliminując całkowicie z obliczeń i modelowania struktury przestrzennej MSDO sprawiające „problemy” dane.

## 5.2. Metodyka procedury interaktywnego czyszczenia semiwariogramu

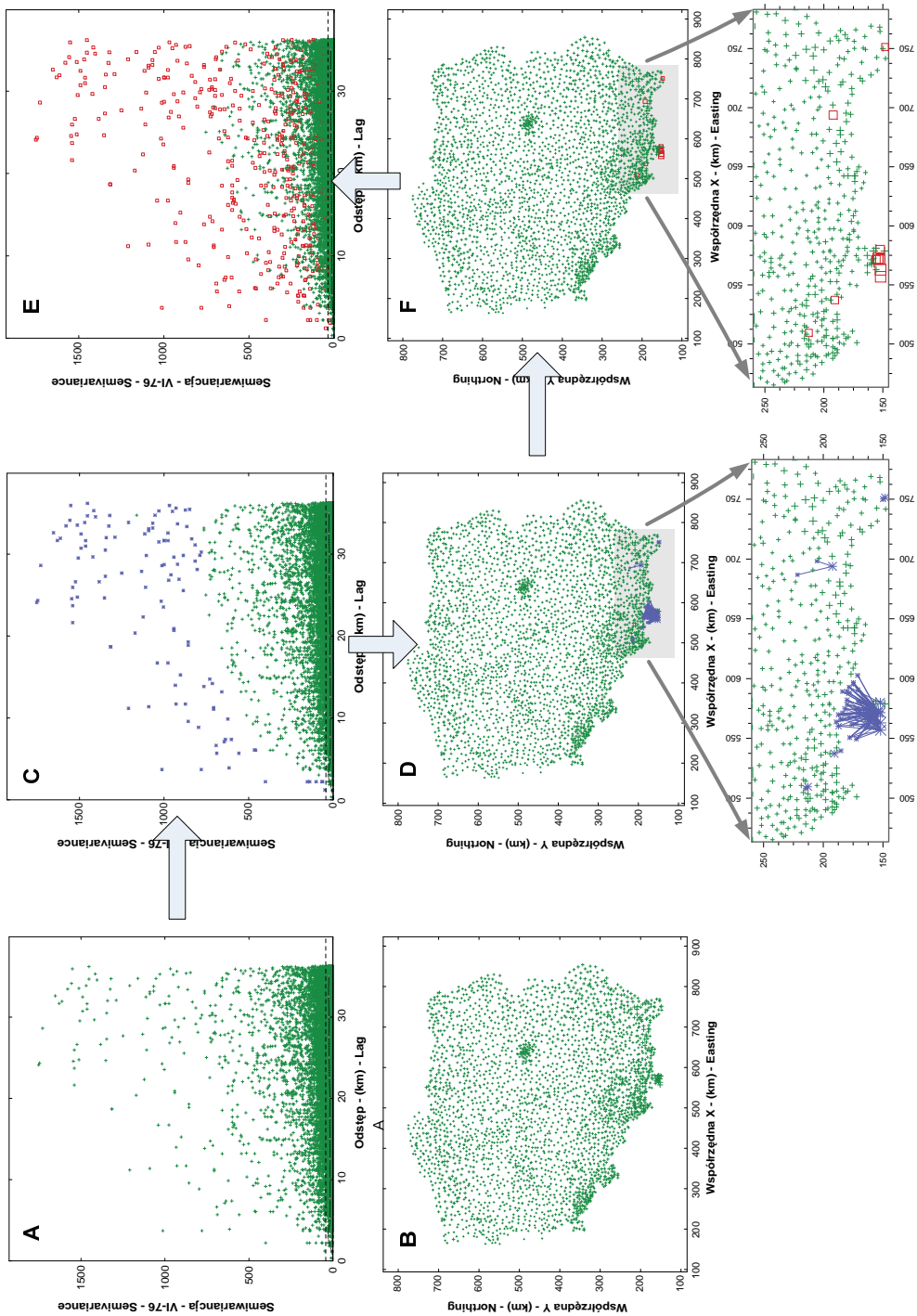
Kolejne etapy zastosowanej procedury interaktywnego czyszczenia semiwariogramu przedstawiono na rycinach 140-142. W przykładzie wykorzystano surowe dane MSDO (w mm) zanotowane w czerwcu 1976. Dla uproszczenia obliczenia wykonywano jedynie dla 15 odstępów o szerokości 2,5 km.

Wykres semiwariogramu MSDO z czerwca 1976 roku, wykonany dla wszystkich dostępnych danych, przedstawiono na rycinie 140. Pokazano na niej klarowny przykład struktury przestrzennej zjawiska zaburzonej przez występowanie odstających (być może ekstremalnych) danych. Wartości semiwariancji jako miary niepodobieństwa powinny rosnać wraz ze wzrostem odległości pomiędzy porównywanymi wynikami pomiarów. Tymczasem, na wykresie jest widoczne, że po gwałtownym wzroście w pierwszym odstępnie (od ok. 5<sup>2</sup> do blisko 30<sup>2</sup> mm) wartości semiwariancji dla następnych czterech odstępów (do ok. 12,5 km) zamiast rosnać – maleją. Aby zidentyfikować przyczynę tej anomalii, wykonuje się wykres chmury semiwariogramu (A na ryc. 141). Wyraźnie rysuje się na nim szereg, rozproszonych, od-

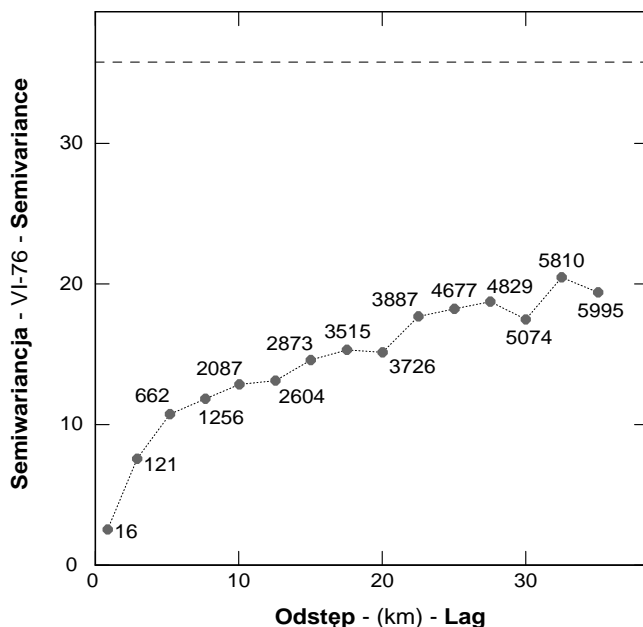
stających od głównej chmury, bardzo wysokich wartości, podniesionych do kwadratu, różnic wyników pomiarów. Identyfikacja miejsc, gdzie znajdują się punkty pomiarowe, na których te opady zostały zmierzone, i jakie konkretnie są to wartości, dokonywana jest za pomocą narzędzia połączonych okien (ang. *linking windows*). Dzięki temu zaznaczone na jednym wykresie lub mapie dane są również widoczne w innym oknie programu zawierającym odmienną ich reprezentację graficzną. W odrębnym oknie wyświetlana jest mapa lokalizacyjna z symbolami proporcjonalnymi do zmierzonych w czerwcu 1976 MSDO (B na ryc. 141). Na wykresie chmury semiwariogramu zaznacza się, za pomocą kursora myszki, wszystkie budzące wątpliwość wysokie wartości semiwariancji (niebieskie gwiazdki na części C ryc. 141). Wykonanie tej czynności skutkuje natychmiast zaznaczeniem na powiązanej z wykresem mapie wszystkich par lokalizacji, które brały udział w obliczeniach zaznaczonych wartości semiwariancji (połączone liniami niebieskie gwiazdki na części D ryc. 141, widoczne dokładniej na powiększonym fragmencie). Widać dzięki temu wyraźnie, że występują one wyłącznie w Karpatach, a zwłaszcza Tatrach i na ich przedpolu. Jest to zatem efekt bardzo wysokich opadów orograficznych, które zdecydowanie pod względem relacji



**Ryc. 140.** Bez kierunkowy (izotropowy) semiwariogram empiryczny surowych danych pomiarowych MSDO z czerwca 1976 roku. Obliczono go dla 15 odstępów o szerokości 2,5 km. Przy każdej wartości semiwariancji podano liczbę par danych, z których została wyliczona



**Ryc. 141.** Schemat procedury interaktywnego czyszczenia semiwariogramu oparty na wykresie chmury semiwariogramu (A) i mapie lokalizacyjnej z symbolami proporcjonalnymi do wyników pomiarów (B). Szczegółowe omówienie w tekście [14]



Ryc. 142. Skorygowany bezkierunkowy (izotropowy) semiwariogram empiryczny surowych danych pomiarowych MSDO z czerwca 1976 roku - porównaj z ryciną 140. Przy każdej wartości semiwariancji podano liczbę par danych, z których została wyliczona; porównanie z ryciną 140 daje informację o liczbie usuniętych (zamaskowanych) par

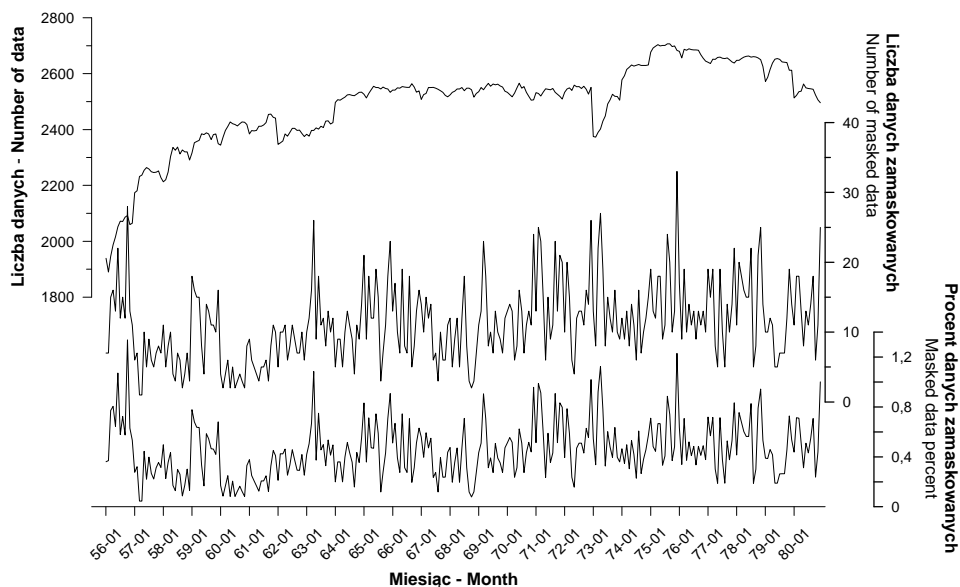
przestrzennych „nie pasują” do reszty danych z całej Polski. Dokładna inspekcja mapy w dużym powiększeniu umożliwiła precyzyjne zlokalizowanie miejsc, w których wystąpiły anomalne opady, i podjęcie decyzji, które z nich zostaną „zamaskowane”, a zatem usunięte czasowo z dalszych obliczeń.

W wypadku czerwca 1976 roku zdecydowano o maskowaniu danych z 11 stanowisk (F na ryc. 141), co wobec całego, liczącego 2686 lokalizacji, zbioru danych stanowiło nieco ponad 0,4%. Na owe 11 punktów 7 zlokalizowane było w Tatrach. Należały do nich m.in. posterunki Hala Ornak, Polana Chochołowska, Hala Gąsienicowa, Kasprowy Wierch, Myślenickie Turnie, gdzie maksymalne opady dobowe przekraczały 60 mm i były najwyższe w tym miesiącu w całej Polsce. Do grupy tej włączono jednak także Babią Górę, Magurkę i Szyndzielnię oraz Duklę i Wetlinę, na których notowano opady o wysokościach wyraźnie niższych (w kolejności od maksymalnego opadu pozycje 16, 20, 50, 10 i 22). O „anomalności” wartości zmierzonego opadu nie decyduje bowiem jego bezwzględna wysokość, tylko relacja względem pomiarów na stacjach sąsiednich. Eliminacja z obliczeń semiwariancji 11 wyjściowych wartości (surowych pomiarów) spowodowała, że liczba porównywanych par danych na badanym dystansie zmalała o 455

z wyjściowych 47 587 (0,96%, E na ryc. 141). Semiwariancje dla kolejnych odstępów zostały obliczone z pominięciem od jednej do maksymalnie 51 par danych (ryc. 142). Redukcja ta zawierała się w przedziale 0,6-9,0%, ale uzyskany skorygowany semiwariogram daje klarowny, reprezentatywny dla całej Polski, obraz struktury przestrzennej MSDO z czerwca 1976.

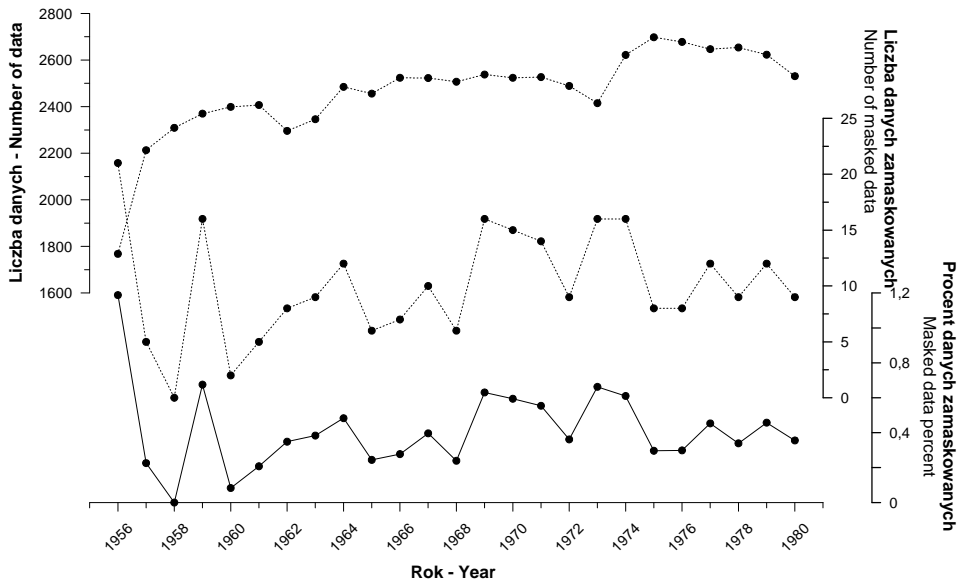
### 5.3. Statystyki procedury czyszczenia semiwariogramu dla wszystkich analizowanych zbiorów danych MSDO

Czyszczenie semiwariogramu było konieczne w wypadku wszystkich analizowanych zbiorów danych oprócz rocznych MSDO z roku 1958 (ryc. 143 i 144). W trakcie obliczania semiwariancji miesięcznych MSDO maskowano 1-33 stanowisk (średnio ok. 11), co stanowiło 0,05-1,34% całości (średnio 0,45%). Te same charakterystyki w odniesieniu do rocznych zbiorów danych wynosiły odpowiednio: 0-21 (średnio ok. 10) i 0-1,19% (średnio 0,4%). W przebiegu czasowym zaznacza się okres lat 1957-1962, gdzie maskowanych danych było nieco mniej. W pozostałej części wielolecia liczba takich przypadków zmienia się dość chaotycznie z miesiąca na miesiąc, bez żadnej wyraźnej sezonowości czy trendu.



Ryc. 143. Liczba i odsetek danych zamaskowanych w trakcie analizy struktury przestrzennej w zestawieniu z całkowitą liczbą danych poszczególnych miesięcznych zbiorów MSDO

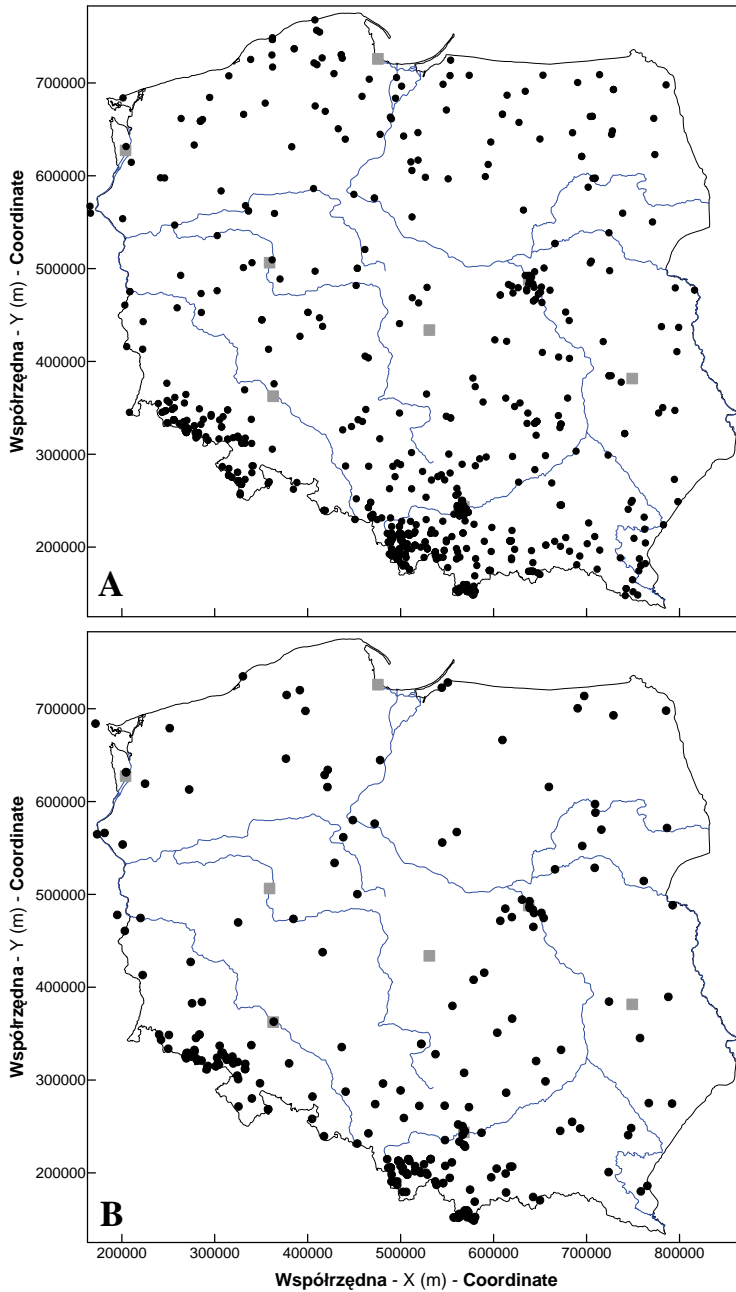




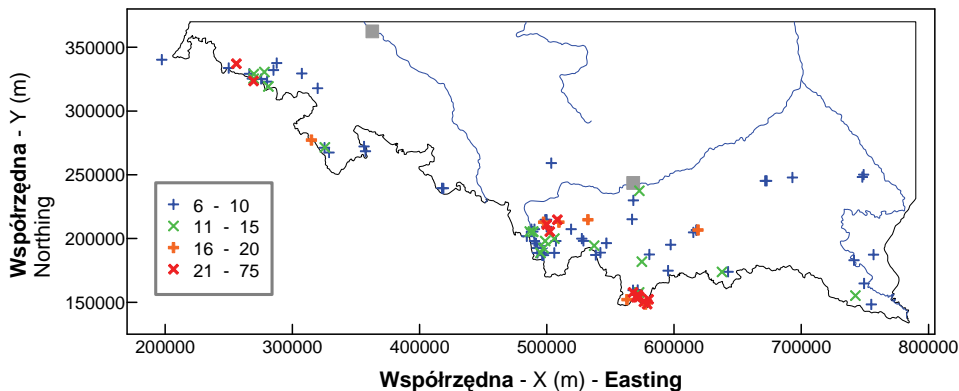
**Ryc. 144.** Liczba i odsetek danych zamaskowanych w trakcie analizy struktury przestrzennej w zestawieniu z całkowitą liczbą danych poszczególnych rocznych zbiorów MSDO

W przykładzie przedstawionym w poprzednim podrozdziale wartości zaburzające globalny obraz struktury przestrzennej dla całego kraju były ewidentnie związane z lokalnymi opadami orograficznymi. I jakkolwiek w innych zbiorach danych takie sytuacje były bardzo częste, to jednak anomalne w sensie przestrzennej struktury podobieństwa opady notowane były na terenie całego kraju (ryc. 145). Jakaś ich część, raczej niewielka, to z pewnością wartości błędne, nie zidentyfikowane na etapie archiwizacji danych. Ponieważ jednak w zdecydowanej większości przypadków nie było jednoznacznych podstaw do podjęcia decyzji o ich całkowitej eliminacji z bazy danych, wszystkie wyniki pomiarów wykluczone na etapie opracowania modelu struktury przestrzennej były używane w trakcie obliczeń estymacji i symulacji (por. dodatek X.2).

Zamieszczone mapy (ryc. 145) pokazują, że wykluczone z obliczeń semiwariancji MSDO występowały głównie w Sudetach i Karpatach (zwłaszcza Zachodnich i Tatrach). Na pozostałym obszarze kraju przypadki takich opadów są raczej rozmieszczone losowo, jakkolwiek widać także skupienia związane z miejskimi sieciami pomiarowymi w Krakowie i Warszawie, czy relatywnie nieco większą ich gęstość na Wyżynie Krakowsko-Częstochowskiej oraz w północno-wschodniej części Pojezierza Pomorskiego.



**Ryc. 145.** Rozmieszczenie posterunków opadowych, na których dokonano pomiarów MSDO zamaskowanych w trakcie analizy struktury przestrzennej: A - zbiory miesięczne, B - zbiory roczne. Dla uproszczenia na mapie A zaznaczono punkty pomiarowe jedynie z 5 wybranych lat: 1956, 1960, 1965, 1970 i 1975



**Ryc. 146.** Rozmieszczenie posterunków opadowych, na których dokonano więcej niż 5 pomiarów MSDO zamaskowanych w trakcie analizy struktury przestrzennej zbiorów miesięcznych. W pozostałej, nie przedstawionej na mapie części Polski, takich lokalizacji było tylko 4: Szczecin – Warszewo (6 razy), Warszawa – Muzeum Narodowe (8), Wyszaków nad Bugiem (7), Klusy na południowo-wschodnim krańcu Krainy Wielkich Jezior Mazurskich (7) 🗺

Należy również zwrócić uwagę, że „niepasujące”, ze względu na relacje przestrzenne, MSDO zdarzały się w tych samych lokalizacjach w analizowanym wieloleciu często więcej niż jeden raz (ryc. 146). Na przykład, wśród 251 zakwestionowanych przypadków maksymalnych rocznych sum dobowych było takich sytuacji 26. Sześciokrotnie w danym miejscu zanotowano dwa takie przypadki, trzykrotnie – trzy, a raz – aż pięć (stanowisko Równica w Beskidzie Śląskim – bądź „schronisko”, bądź „wieś”). W zbiorach miesięcznych takie sytuacje powtarzały się zdecydowanie częściej. Na 3399 zamaskowanych wartości MSDO 1274 lokalizacji pojawiło się tylko raz. Można przypuszczać, że wśród nich znaczący odsetek stanowią pojedyncze przypadki zanotowanych tylko w jednej lokalizacji, głównie na Niżu, wysokich opadów konwekcyjnych oraz zapewne w jakiejś części pomiary błędne. Pozostałe 2125 usuniętych z analizy struktury przestrzennej powtarzało się w 564 lokalizacjach. W 467 przypadkach powtórzeń było 2-5; 38 razy w jednym stanowisku takich przypadków było więcej niż 10. Maksymalna liczba powtórzeń, bo aż 75, związana jest z pomiarami wykonanymi na Kasprowym Wierchu – najwyższym stałym stanowisku pomiarów opadów atmosferycznych w Polsce (1991 m n.p.m.). Kolejne 4 punkty stanowią następne w kolejności malejącej wysokości: Dolina Pięciu Stawów (1670 m n.p.m. – 53 powtórzeń), Śnieżka (1603 m n.p.m. – 45 powtórzeń), Hala Gąsienicowa (1520 m n.p.m. – 45 powtórzeń), Myślenickie Turnie (1360 m n.p.m. – 39 powtórzeń). Tylko 4 z 97 przypadków, kiedy powtórzeń było więcej niż 5, zlokalizowano na Niżu (ryc. 146). Pozostałe usytuowane były w Sudetach i Karpatach oraz

na ich przedgórzu. Największym zagęszczeniem takich stanowisk charakteryzowały się Karkonosze, Beskid Śląski i Tatry.

Omówione w niniejszym podrozdziale zagadnienie pozwala spojrzeć na kwestię „ekstremalnych” opadów w trochę innym świetle. Zazwyczaj termin ten rezerwuje się dla zdarzeń, w trakcie których zanotowano bardzo wysokie sumy lub intensywności, których częstość występowania jest bardzo niska. Decyduje tu zatem kryterium czasowe. Z przeprowadzonej analizy wynika jednakże, że notowane są w Polsce opady o bardzo ograniczonym zasięgu przestrzennym, niekoniecznie skrajnie wysokie, które są wyraźnymi anomaliami z punktu widzenia struktury globalnego ich pola. Zagadnienie to zarówno w aspekcie czasowym, jak i przestrzennym będzie tematem osobnego opracowania.

## **6. Rozdzielczość modeli przestrzennych**

### **6.1. Ogólne kryteria wyboru rozdzielczości rastrowych modeli przestrzennych**

Analizy GIS, których efektem są rastrowe modele zmienności przestrzennej wymagają na etapie planowania metodyki podjęcia decyzji o ich rozdzielczości, czyli inaczej mówiąc – o wymiarach oczka siatki (ang. *grid*). Bardzo często problem ten pomijany jest milczeniem, a decyzja – podejmowana bez uwzględniania jakichkolwiek obiektywnych kryteriów. Jasne jest, że najważniejsze z nich muszą dotyczyć zbioru danych, na podstawie którego budowany jest model. Decydujące znaczenie ma typ próbkowania (punktowe bądź obszarowe; losowe, regularne, preferencyjne, profilowe itp.) i zagęszczenie danych (liczba na jednostkę powierzchni) oraz charakterystyki zmienności przestrzennej analizowanej cechy (statystyki lokalne, a w tym także parametry autokorelacji). Powinno się również uwzględnić rozkład statystyczny błędów zarówno samych pomiarów, jak i określenia lokalizacji stanowisk pomiarowych. Można stwierdzić intuicyjnie, że tworzenie modelu o rozdzielczości 100 m, kiedy lokalizacje danych są od siebie odległe o kilkadziesiąt kilometrów, ma niewielki sens.

Nie można również pominąć aspektów technicznych przy podejmowaniu decyzji o rozdzielczości modelu. Pierwszy jest związany ze skalą docelowej mapy. Te dwa elementy powinny być ze sobą zharmonizowane, a skala tak dobrana, aby można było przy danej rozdzielczości zobaczyć wszystkie istotne szczegóły modelu. I odwrotnie, nie ma sensu tworzyć modelu o wysokiej rozdzielczości, jeśli przy założonej z góry skali mapy finalny efekt będzie musiał ulec generalizacji. Drugi aspekt techniczny stanowi

ograniczenia pojemności pamięci i szybkości przetwarzania wykorzystywanego komputera. Parametry te zmieniają się z miesiąca na miesiąc, ale wciąż nie można ich ignorować. W tym momencie (styczeń 2008) górną granicę<sup>67</sup> dla przeciętnego komputera osobistego o pamięci rzędu 1 GB stanowi dla obszarów o powierzchni Polski rozdzielczość rzędu 10-50 m ( $x \cdot 10^8$ - $x \cdot 10^9$  węzłów).

Trzecim elementem, który należy wziąć pod uwagę jest przeznaczenie modelu. Jeśli ma posłużyć jedynie do celów naukowych – charakterystyki i wyjaśniania zmienności przestrzennej zjawiska – zazwyczaj ważne są raczej proporcje i relacje niż bezwzględne wartości. Często też ignoruje się wówczas, przy analizach obejmujących cały kraj, zmienność zachodzącą w małej skali przestrzennej. Zastosowania praktyczne stawiają w wielu aspektach wyższe wymagania odnośnie rozdzielczości, precyzji i dokładności modelu. Różnią się one oczywiście znacznie w zależności od celu. Inne muszą być przy obliczeniach bilansów wodnych zlewni, inne przy projektowaniu sieci kanalizacji burzowej w mieście itp.

## **6.2. Ocena optymalnej rozdzielczości modeli z wykorzystaniem charakterystyk analizowanego zbioru danych**

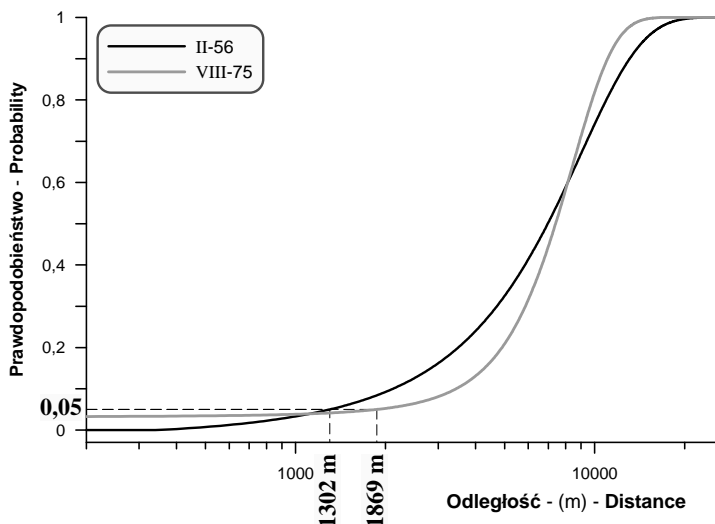
Przy konstruowaniu modeli przestrzennych maksymalnych sum dobowych opadów na obszarze Polski, tworzonych w niniejszej rozprawie, stosowano zalecenia zawarte w publikacji Hengla (2006), a dotyczące optymalizacji ich rozdzielczości. Nie są to jakieś sztywne reguły pozwalające w sposób jednoznaczny określić optymalną wartość. Autor cytowanej pracy zestawiał dla kryteriów związanych z właściwościami analizowanego zbioru danych, a także z docelową skalą mapy, formuły pozwalające oszacować optymalny zakres rozdzielczości, a także jej wartość zalecaną. Przy każdym z kryteriów są to jednakże nieco inne liczby.

Ocenę zaczęto od parametru najmniej ważnego – docelowej skali map. Niniejsza publikacja wydrukowana jest w formacie B5. Oznacza to, że maksymalny wymiar mapy<sup>68</sup>, biorąc pod uwagę kształt Polski, może na kierunku W-E wynosić 12,5 cm. Żeby przedstawić całe terytorium, należy zastosować skalę około 1 : 6 000 000. Według podanych przez Hengla (2006) i omówionych szczegółowo reguł, optymalna rozdzielczość dla takiej skali powinna się zawierać w przedziale 600-15 000 m (od  $SM \times 0,0001$  do  $SM \times 0,0025$ ). Wartość zalecana wynosi zaś 3000 m ( $SM \times 0,0005$ ).

---

<sup>67</sup> Spore znaczenie ma optymalizacja algorytmów, czyli oprogramowanie.

<sup>68</sup> Wykluczono, ze względu na liczbę map, opcję druku wkładek lub załączników o większym formacie.



**Ryc. 147.** Krzywe prawdopodobieństwa znalezienia jednego punktu danych w określonym promieniu (odległości) od innego punktu dla miesięcy z najmniejszą (luty 1956) i największą (sierpień 1975, szczegóły w rozdz. V) liczbą stanowisk. Dla obu zbiorów danych zaznaczono odległość 95% prawdopodobieństwa, że ów punkt znajduje się dalej

Drugie kryterium związane jest z gęstością danych (ang. *inspection density*), to jest ich liczbą ( $N$ ) względem powierzchni analizowanego obszaru ( $A$ ), niezależnie od ich rozkładu przestrzennego. Zgodnie z informacjami podanymi w rozdziale V, dysponowano dla analizowanego wielolecia średnio 2492 wynikami pomiarów miesięcznych MSDO i 2478 – rocznych. Ponieważ liczby te różnią się niewiele, zdecydowano, że w obliczeniach użyta zostanie wartość 2480. Tworzone modele obejmowały powierzchnię<sup>69</sup> 319 114 km<sup>2</sup>. Według Hengla (2006) zakres rozdzielnoczości ( $p$ )<sup>70</sup> tworzonego modelu powinien mieścić się w przedziale

$\left( 0,05 \cdot \sqrt{\frac{A}{N}} \leq p \leq 0,1 \cdot \sqrt{\frac{A}{N}} \right)$ , a zalecana wartość wynosić  $p = 0,0791 \cdot \sqrt{\frac{A}{N}}$ . Użycie podanych powyżej danych we wzorach daje zakres 567-1134 m i wartość zalecaną 897 m.

Kryterium zagęszczenia danych jest bardzo proste, bo nie uwzględnia w ogóle ich rozkładu przestrzennego. Bardziej wyrafinowana metoda oceny

\_\_\_\_\_

<sup>69</sup> Terytorium państwa plus Zalew Szczeciński, Zatoka Pucka i Zalew Wiślany oraz strefa o szerokości 2 km wokół granic.

<sup>70</sup> W niniejszym podrozdziale zastosowano notację używaną przez Hengla (2006), co oznacza, że niektóre symbole literowe mają inne znaczenie niż w pozostałej części pracy. Są one jednak w tekście dokładnie objaśnione.

optymalnej rozdzielczości modelu opiera się na charakterystykach statystycznych rozkładu punktów (ang. *point pattern analysis*). W zależności od generalnego typu rozmieszczenia danych – regularnego lub losowego – używa się nieco innych formuł. Obliczenia, których wyniki przedstawiono w podrozdziale V.4 dają podstawy do stwierdzenia, że przez cały okres analizowanego wielolecia, niezależnie od zmiany liczby i rozmieszczenia stanowisk pomiarowych, rozkład ich może być uważany za losowy (wartości wskaźnika Clarka-Evansa oscylujące wokół 1). W tej sytuacji potrzebne są dwie charakterystyki statystyczne analizowanego zbioru danych (Hengl 2006): średnia odległość do najbliższego sąsiada ( $\bar{h}_{ij}$ ) i odległość od dowolnego punktu danych, w której z prawdopodobieństwem 5% znaleziony zostanie inny punkt ( $h_{ij(p=5\%)}$ ). Odpowiednie wzory wyznaczające zakres optymalnych rozdzielczości modelu są wówczas następujące:

$$h_{ij(p=5\%)} \leq p \leq \frac{\bar{h}_{ij}}{2}, \quad [47]$$

a wartość zalecaną określa się z poniższej formuły:

$$p = \frac{\sqrt{\frac{A}{N}}}{4} \quad [48]$$

Średnia odległość do najbliższego sąsiada ( $\bar{h}_{ij}$ ) oscylowała w analizowanych zbiorach danych MSDO w zakresie 6,8-7,8 km, wynosząc przeciętnie 7,3 km (por. podrozdz. V.4). Drugi potrzebny parametr ( $h_{ij(p=5\%)}$ ) określono tylko dla miesięcy o skrajnej liczbie i rozkładzie stanowisk pomiarowych (por. podrozdz. V.4): lutego 1956 roku i sierpnia 1975 (ryc. 147). Wynosił on odpowiednio 1302 i 1869 m. Tak więc, biorąc, z przytoczonych powyżej skrajne wartości, optymalny zakres rozdzielczości siatek estymacji i symulacji powinien, ze względu na rozkład przestrzenny danych, zawierać się w przedziale 1302-3900 m, a zalecany rozmiar – 2836 m.

Ostatnie kryterium dotyczy zasięgu autokorelacji danych – odległości podobieństwa wyników pomiarów (ang. *spatial dependence structure*). Daje ono możliwość oceny ich nadmiarowości (redundancji w języku teorii informacji), a poprzez to ocenę skali dokładności interpolacji. Do obliczeń potrzebny jest zasięg autokorelacji ( $h_R$ ), liczba ( $m$ ) par punktów danych znajdujących się w jego zasięgu, a także zdefiniowany powyżej parametr rozkładu przestrzennego stanowisk pomiarowych  $h_{ij(p=5\%)}$ . Zakres optymalnych rozdzielczości modelu określa się przy użyciu następujących wzorów:

$$h_{ij(p=5\%)} \leq p \leq \frac{h_R}{2}, \quad [49]$$

a wartość zalecaną – z poniższej formuły:

$$p = h_R \cdot m^{\frac{1}{3}} \quad [50]$$

Określenie zakresu  $p$  w tym przypadku nastęczyło nieco trudności. Struktura przestrzenna dominującej większości analizowanych zbiorów miesięcznych i rocznych MSDO jest złożona i bardzo zmienna (patrz rozdz. VI). Ponadto, w ponad 63% miesięcy i 44% lat definitywny zasięg autokorelacji nie został określony – przekraczał maksymalny uwzględniony w analizie odstęp danych, czyli 212,5 km. Zdecydowano zatem przyjąć do obliczeń zasięg autokorelacji najczęściej występującej struktury krótkodystansowej istniejącej w 285 analizowanych zborach miesięcznych i wszystkich (25) – zborach rocznych. Wynosiła ona średnio 15 500 m (8000-40 000 m, SD = 6370 m). Taką i mniejszą odległość między sobą miały 7770-9690 par punktów danych. Podstawienie do wyżej podanych wzorów średnich wartości obu potrzebnych parametrów ( $h_R = 15\,500$  i  $m = 8700$  oraz  $h_{ij(p=5\%)} = 1560$  m) dało przedział 1560-7750 m i wartość zalecaną 753 m. Jest to oczywiście wynik sprzeczny – wartość zalecana znajduje się poza zakresem optymalnym. Wynika on z omówionych powyżej trudności w parametryzacji obliczeń według kryterium zasięgu autokorelacji.

Zestawienie zalecanych rozdzielczości modeli rastrowych estymacji i symulacji danych MSDO uzyskanych dla czterech kryteriów – skali mapy, zagęszczenia danych, ich rozkładu przestrzennego i zasięgu autokorelacji – daje odpowiednio wartości około 3000, 900, 2840 i 750 m. Wszystkie, przy objęciu analizą całego terytorium kraju, są w zakresie aktualnych możliwości przetwarzania i gromadzenia danych. Biorąc pod uwagę prostotę struktury modelu, a także zakres błędów lokalizacji stanowisk pomiarowych wynoszący od około 600 do 900 m (por. podrozdz. V.2), zdecydowano się na przyjęcie konsekwentnie we wszystkich przypadkach kilometrowej rozdzielczości (1 × 1 km) tworzonych map rastrowych.



## Literatura

- Bacchi B., Kottegoda N., 1995: *Identification and calibration of spatial correlation patterns of rainfall*. Journal of Hydrology, 165, 311-348.
- Barabesi L., Marcheselli M., 2002: *Species abundance estimation using point-to-plant sampling in a design-based setting*. Environmental and Ecological Statistics, 9, 393-403.
- Bell J.L., Sloan L.C., Snyder M.A., 2004: *Regional changes in extreme climatic events: a future climate scenario*. Journal of Climate, 17, 81-87.
- Berne A., Boudevillain B., Kirstetter P.-E., Delrieu G., 2006: *Spatial structure of intense Mediterranean precipitation*. [W:] Proceedings of ERAD 2006, Fourth European Conference on Radar in Meteorology and Hydrology, Barcelona, 18-22.09.2006. ([www.grahi.upc.es/ERAD2006/proceedingsMask/00123.pdf](http://www.grahi.upc.es/ERAD2006/proceedingsMask/00123.pdf))
- Bleinès C., Deraisme J., Geffroy F., Jeannée N., Perseval S., Rambert F., Renard D., Torres O., Touffait Y., 2007: *Isatis technical references, version 7.0.0*. Geovariances and Ecole des Mines de Paris, 1-132.
- Bogdanowicz E., Stachý J., 1998: *Maksymalne opady deszczu w Polsce. Charakterystyki projektowe* [sum.: Heavy rainfalls in Poland. A design approach]. Instytut Meteorologii i Gospodarki Wodnej, Materiały Badawcze, Seria: Hydrologia i Oceanologia 23, 1-85.
- Brunsdon C., 1995: *Estimating probability surfaces for geographical point data: an adaptive kernel algorithm*. Computers and Geosciences, 21, 7, 877-894.
- Carr J.R., Deng E.D., 1987: *Comparison of two techniques for applying disjunctive kriging: the Gaussian anamorphosis model versus the direct statistical inference of the bivariate distributions*. Mathematical Geology, 19, 1, 57-68.
- Chilès J.-P., Delfiner P., 1999: *Geostatistics. Modeling spatial uncertainty*. John Wiley and Sons, New York, 1-695.
- Christakos G., 2000: *Modern spatiotemporal geostatistics*. International Association for Mathematical Geology, Studies in Mathematical Geology 6, Oxford University Press, New York, 1-288.
- Christensen O.B., Christensen J.H., 2004: *Intensification of extreme European summer precipitation in a warmer climate*. Global and Planetary Change, 44, 107-117.
- Ciach G.J., Krajewski W.F., 2006: *Analysis and modeling of spatial structure in small-scale rainfall in Central Oklahoma*. Advances in Water Resources, 29, 1450-1463.
- Cressie N., 1985: *Fitting variogram models by weighted last squares*. Mathematical Geology, 17, 5, 563-585.
- Cressie N., 1990: *The origins of kriging*. Mathematical Geology, 22, 3, 239-252.

- Cressie N.A.C., 1993: *Statistics for spatial data*. Revised edition. John Wiley and Sons, New York, 1-900.
- CrimeStat III, 2004: *CrimeStat III version 3: A spatial statistics program for the analysis of crime incident locations*. Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington DC., November 2004, 1-987.
- Davis B.M., 1987: *Uses and abuses of cross-validation in geostatistics*. *Mathematical Geology*, 19, 3, 241-248.
- Deutsch C.V., Journel A.G., 1992: *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 1-340.
- Deutsch C.V., Journel A.G., 1998: *GSLIB: Geostatistical software library and user's guide*. Edition 2, Oxford University Press, New York, 1-369.
- Dietrich C.R., Osborne M.R., 1991: *Estimation of covariance parameters in kriging via restricted maximum likelihood*. *Mathematical Geology*, 23 (7), 655-672.
- Dobesch H., Tveito O.E., Bessemoulin P., 2001: *Geographic Information Systems in climatological application*. Norwegian Meteorological Institute, Oslo, DNMI – Report, Klima, 13/01, 1-49.
- Dore M.H.I., 2005: *Climate change and changes in global precipitation patterns: What do we know*. *Environment International*, 31, 1167-1181.
- Douglas E.M., Barros A.P., 2003: *Probable maximum precipitation estimation using multifractals: application in the eastern United States*. *Journal of Hydrometeorology*, 4, 1012-1024.
- Dubois G., 2003: *Spatial Interpolation Comparison 97. Introduction and description of the data set*. [W:] *Mapping radioactivity in the environment*. Spatial Interpolation Comparison 97. Dubois G., Malczewski J., De Cort M. (red.), European Commission, Joint Research Centre, EUR 20667 EN, Luxembourg, 39-44.
- Dubois G., Shibli S.A.R., 2003: *Monitoring of environmental radioactivity: automatic mapping or expert-dependent systems?* [W:] *Mapping radioactivity in the environment*. Spatial Interpolation Comparison 1997. Dubois G., Malczewski J., De Cort M., (red.), EUR 20667 EN, 253-268.
- Durrans S.R., Kirby J.T., 2004: *Regionalization of extreme precipitation estimates for the Alabama rainfall atlas*. *Journal of Hydrology*, 295, 101-107.
- Fal B., 1987: *Maksymalne opady dobowe*. [W:] *Atlas hydrologiczny Polski*, Wydawnictwa Geologiczne, Warszawa.
- Fal B., Plenzler W., 1981: *Intensywność zasilania opadowego w Polsce (mapa maksymalnych sum opadów dobowych o prawdopodobieństwie  $p = 1\%$ )*. *Wiadomości Instytutu Meteorologii i Gospodarki Wodnej*, 7, (1-2), 55-60.
- Faulkner D.S., Prudhomme C., 1998: *Mapping an index of extreme rainfall across the UK*. *Hydrology and Earth Systems Sciences*, 2, 2-3, 183-194.
- Ferraris L., Gabellani S., Parodi U., Rebora N., 2003: *Revisiting multifractality in rainfall fields*. *Journal of Hydrometeorology*, 4, 544-551.
- Fisher N.L., 1993. *Statistical analysis of circular data*. Cambridge University Press, Cambridge, 1-277.
- Fowler A.M., Hennessy K.J., 1995: *Potential impacts of global warming on the frequency and magnitude of heavy precipitation*. *Natural Hazards*, 11, 283-303.
- Fowler H.J., Ekström M., Kilsby C.G., Jones P.D., 2005: *New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate*. *Journal of Hydrology*, 300, 212-233.
- Frei C., Davies H.C., Joachim Gurtz J., Schär C., 2000: *Climate dynamics and extreme precipitation and flood events in Central Europe*. *Integrated Assessment*, 1, 281-299.

- Frei C., Schär C., Lüthi D., Davies H.C., 1998: *Heavy precipitation processes in a warmer climate*. Geophysical Research Letters, 25, 9, 1431-1434.
- Frich P., Alexander L.V., Della-Marta P., Gleason B., Haylock M., Klein Tank A.M.G., Peterson T., 2002: *Observed coherent changes in climatic extremes during the second half of the twentieth century*. Climate Research, 19, 193-212.
- Froidevaux R., 1993: *Probability field simulation*. [W:] Geostatistics Troia 1992, Soares A. (red.), vol. 1, Kluwer, New York, 73-84.
- Gabrosek J., Cressie N., 2002: *The effect on attribute prediction of location uncertainty of spatial data*. Geographical Analysis, 34, 3, 262-285.
- Goodrich D.C., Faurès J.-M., Woolhiser D.A., Lane L.L., Sorooshian S., 1995: *Measurement and analysis of small-scale convective storm rainfall variability*. Journal of Hydrology, 173, 283-308.
- Goovaerts P., 1997: *Geostatistics for natural resources evaluation*. Oxford University Press, New York, 1-483.
- Górska-Zabielska M., Stach A., 2008: *Analiza struktury przestrzennej i estymacja składu petrograficznego osadów fluwioglacjalnych wistulianu w strefie marginalnej lobu Odry i na obszarach przyległych*. Przegląd Geograficzny, 80, 1, 101-130.
- Grimes D.I.F., Pardo-Igúzquiza E., Bonifacio R., 1999: *Optimal areal rainfall estimation using rain gauges and satellite data*. Journal of Hydrology, 222, 93-108.
- Gringarten E., Deutsch C.V., 2001: *Teacher's aide. Variogram interpretation and modeling*. Mathematical Geology, 33, 4, 507-534.
- Groisman P.Ya., Knight R.W., Easterling D.R., Karl T.R., Hegerl G.C., Razuvaev V.N., 2005: *Trends in intense precipitation in the climate record*. Journal of Climate, 18, 1326-1350.
- Groisman P.Y., Legates D.R., 1995: *Documenting and detecting long-term precipitation trends: where we are and what should be done*. Climatic Change, 31, 601-622.
- Haas T.C., 1990. *Kriging and automated variogram modeling within a moving window*. Atmospheric Environment 24A, 1759-1769.
- Haberlandt U., 2007: *Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event*. Journal of Hydrology, 332, 144-157.
- Hamilton J.D. 1994: *Time Series Analysis*. Princeton University Press, Princeton, New Jersey, 1-779.
- Harada Y., Shimada T., 2006: *Examining the impact of the precision of address geocoding on estimated density of crime locations*. Computers and Geosciences, 32, 1096-1107.
- Hauf T., Theusner M., 2003: *On the small scale structure of convective precipitation*. [W:] 83<sup>rd</sup> Annual Meeting American Meteorological Society, 17<sup>th</sup> Conference on Hydrology, Observing and Understanding the Variability of Water in Weather and Climate, Long Beach, CA, 8-13.02.2003. (<http://ams.confex.com/ams/pdfpapers/54660.pdf>)
- Heino R., Brázdil R., Førland E., Tuomenvirta H., Alexandersson H., Beniston M., Pfister C., Rebetez M., Rosenhagen G., Rösner S., Wibig J., 1999: *Progress in the study of climatic extremes in Northern and Central Europe*. Climatic Change, 42, 151-181.
- Hengl T., 2006: *Finding the right pixel size*. Computers and Geosciences, 32, 1283-1298.
- Hennessy K.J., Gregory J.M., Mitchell J.F.B., 1997: *Changes in daily precipitation under enhanced greenhouse conditions*. Climate Dynamics, 13, 667-680.
- Hense A., Friederichs P., 2006: *Wind and precipitation extremes in the Earth's atmosphere*. [W:] Extreme Events in Nature and Society, Albeverio S., Jentsch V., Kantz H. (red.), Springer, Berlin, Heidelberg, 169-187.

- Hirsch R.M., Slack J.R., Smith R.A., 1982: *Techniques of trend analysis for monthly water quality data*. Water Resources Research, 18(1), 107-121.
- Isaaks E.H., Srivastava R.M., 1989: *Applied geostatistics. An Introduction*. Oxford University Press, New York, 1-551.
- Jakubiak J., 1984: *Błędy systematyczne pomiaru opadów atmosferycznych ciektych i ich redukcja z poziomu standardowego do poziomu gruntu [sum.: Systematic errors in measurements of liquid atmospheric precipitations and their reduction from the standard to the ground level]*. Zeszyty Problemowe Postępów Nauk Rolniczych, 288, 11-20.
- Jian X., Olea R.A., Yu Y.-S., 1996: *Semivariogram modeling by weighted least squares*. Computers and Geosciences, 22 (3), 387-397.
- Joseph R., Ting M., Kumar P., 2000: *Multiple-scale spatio-temporal variability of precipitation over coterminous United States*. Journal of Hydrometeorology, 1, 373-392.
- Jurczyk A., Osródkka K., Moszkowicz S., Mazzetti C., Szturc J., 2004: *Precipitation field estimation based on radar and raingauge data*. [W:] Sixth International Symposium on Hydrological Applications of Weather Radar, Melbourne, Australia, 2-4 February 2004, 6 str.
- Jurczyk A., Osródkka K., Szturc J., 2007: *Research studies on improvement in real-time estimation of radar-based precipitation in Poland*. Meteorology and Atmosphere Physics, DOI 10.1007/s00703-007-0266-3, 15 str.
- Kemenade van C.H.M., La Poutré H., Mokken R.J., 1999: *Unsupervised class detection by adaptive sampling and density estimation*. [W:] Stein A., van der Meer F., Gorte B., Spatial Statistics for Remote Sensing, Kluwer Academic Publishers, 165-183.
- Kennedy P., 1998: *A guide to econometrics*. Fourth Edition. The MIT Press, Cambridge, Massachusetts, 1-468.
- Kerry R., Oliver M.A., 2007: *Determining the effect of asymmetric data on the variogram. II. Outliers*. Computers and Geosciences, 33(10), 1233-1260.
- Kharin V.V., Zwiers F.W., Zhang X., Hegerl G.C., 2007: *Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations*. Journal of Climate, 20, 1419-1444.
- Kharoufeh J.P., Goulias K.G., 2002: *Nonparametric identification of daily activity durations using kernel density estimators*. Transportation Research, Part B, 36, 59-82.
- Kiktev D., Sexton D.M.H., Alexander L., Folland C.K., 2003: *Comparison of modeled and observed trends in indices of daily climate extremes*. Journal of Climate, 16, 3560-3571.
- Klein Tank A.M.G., Können G.P., 2003: *Trends in indices of daily temperature and precipitation extremes in Europe, 1946-99*. Journal of Climate, 16, 3665-3680.
- Kłysik K., Fortuniak K., 1993: *Maksymalne opady dobowe w środkowej Polsce [sum.: The maximum daily precipitation in central Poland]*. Przegląd Geograficzny IGI PAN, 65, 1-2, 97-110.
- Kondracki J., 2000: *Geografia regionalna Polski*. Państwowe Wydawnictwo Naukowe, Warszawa.
- Kondracki J., Richling A., 1994: *Mapa 53.3 Regiony fizycznogeograficzne*. [W:] Atlas Rzeczypospolitej Polskiej. Główny Geodeta Kraju.
- Koning A.J., Franses P.H., 2005: *Are precipitation levels getting higher? Statistical evidence for the Netherlands*. Journal of Climate, 18, 4701-4714.
- Kothavala Z., 1997: *Extreme precipitation events and the applicability of global climate models to the study of floods and droughts*. Mathematics and Computers in Simulation, 43, 261-268.
- Kożuchowski K., 1985: *Zmienność opadów atmosferycznych w Polsce w stuleciu 1881-1980*. Acta Geographica Lodziensia, 48, Ossolineum, Wrocław, 1-158.

- Krige D.G., 1951: *A statistical approach to some basic mine valuation problems on the Witwatersrand*. Journal of the Chemical, Metallurgical and Mining Society of South Africa, December, 119-139.
- Krige D.G., 1952: *A statistical analysis of some of the borehole values in the Orange Free State goldfield*. Journal of the Chemical, Metallurgical and Mining Society of South Africa, September, 47-64.
- Kundzewicz Z.W., 2005: *Intense precipitation and high river flows in Europe – observations and projections*. Acta Geophysica Polonica, 53, 4, 385-400.
- Kundzewicz Z.W., Ulbrich U., Brücher T., Graczyk D., Krüger A., Leckebusch G.C., Menzel L., Pińskwar I., Radziejewski M., Szwed M., 2005: *Summer floods in central Europe – climate change track?* Natural Hazards, 36, 165-189.
- Kunkel K.E., 2003: *North American trends in extreme precipitation*. Natural Hazards, 29: 291-305.
- Kunkel K.E., Andsager K., Easterling D.R., 1999: *Long-term trends in extreme precipitation events over the conterminous United States and Canada*. Journal of Climate, 12, 2515-2527.
- Lajaunie C., 1990: *Comparing some approximate methods for building local confidence intervals for predicting regionalized variables*. Mathematical Geology, 22, 1, 123-144.
- Lam N., 1983: *Spatial Interpolation Methods: A Review*. The American Cartographer, 10, 2, 129-149.
- Lark R.M., Ferguson R.B., 2004: *Mapping risk of soil nutrient deficiency or excess by disjunctive and indicator kriging*. Geoderma, 118, 39-53.
- Lenart W., 1984: *Porównanie metod oceny opadów rzeczywistych w krajach zlewiska Morza Bałtyckiego [sum.: Comparison of estimation methods of real precipitations in countries of the Baltic Sea catchment area]*. Zeszyty Problemowe Postępów Nauk Rolniczych, 288, 149-158.
- Libiseller C., Grimvall A., 2002: *Performance of partial Mann-Kendall test for trend detection in the presence of covariates*. Environmetrics, 13, 71-84.
- Lovejoy S., 1982: *Area perimeter relations for rain and cloud areas*. Science, 187, 1035-1037.
- Lovejoy S., Schertzer D., 1995: *Multifractals and rain*. [W:] New Uncertainty Concepts in Hydrology and Water Resources, Kundzewicz Z.W. (red.), Cambridge University Press, Cambridge, 62-103.
- Lovejoy S., Schertzer D., Allaire V.C., 2008: *The remarkable wide range spatial scaling of TRMM precipitation*. Atmospheric Research, 90, Issue 1, 10-32.
- Lynch S., 2007: *Dynamical Systems with Applications using Mathematica®*. Birkhäuser, Boston, 1-484.
- Ma X., Yao T., 2001: *A program for 2D modeling (cross) correlogram tables using fast Fourier transform*. Computers and Geosciences, 27, 763-774.
- Mao S., Journel A.G., 1999: *Conditional 3D simulation of lithofacies with 2D seismic data*. Computers and Geosciences, 25, 845-862.
- Mardia K.V., Jupp P.E., 2000: *Statistics of directional data*. 2nd Edition. John Wiley and Sons, Chichester, 1-429.
- Minasny B., McBratney A.B., Whelan B.M., 2005: *VESPER version 1.62*. Australian Centre for Precision Agriculture, McMillan Building A05, The University of Sydney, NSW 2006. (<http://www.usyd.edu.au/su/agric/acpa>)
- Moszkowicz S., 2000: *Small-scale structure of rain field – preliminary results basing on a digital gauge network and on MRL-5 Legionowo radar*. Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 25, 10-12, 933-938.
- Moyeed R.A., Papritz A., 1999: *An empirical comparison of kriging methods for nonlinear spatial point prediction*. Mathematical Geology, 34, 4, 365-386.

- Namysłowska-Wilczyńska B., 2006: *Geostatystyka. Teoria i zastosowania*. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 1-356.
- Nesbitt S.W., Zipser E.J., Kummerow C.D., 2004: *An examination of version-5 rainfall estimates from the TRMM microwave imager, precipitation radar, and rain gauges on global, regional, and storm scales*. *Journal of Applied Meteorology*, 43, 1016-1036.
- New M., Todd M., Hulme M., Jones P., 2001: *Precipitation measurements and trends in the twentieth century*. *International Journal of Climatology*, 21, 1899-1922.
- Olea R.A., 1999: *Geostatistics for engineers and earth scientists*. Kluwer Academic Publishers, Boston, 1-303.
- Olea R.A., 2006: *A six-step practical approach to semivariogram modeling*. *Stochastic Environmental Research and Risk Assessment*, 20, 307-318.
- Palmer T.N., Räisänen J., 2002: *Quantifying the risk of extreme seasonal precipitation events in changing climate*. *NATURE*, 415, January 2002, 512-514.
- Pannatier Y., 1996, *VARIOWIN: Software for spatial data analysis in 2D*. Springer-Verlag, New York, 1-91.
- Panorska A.K., Gershunov A., Kozubowski T.J., 2007: *From Diversity to Volatility: Probability of Daily Precipitation Extremes*. [W:] *Nonlinear Dynamics in Geosciences*, Tsionis A.A. i Elsner J.B. (red.), Springer, New York, 465-484.
- Papritz A., Dubois J.-P., 1999: *Mapping heavy metals in soil by (non-)linear kriging: an empirical validation*. [W:] *GeoENV II - Geostatistics for Environmental Applications*, Gómez-Hernández J., Soares A., Froidevaux R. (red.), Kluwer Academic Publishing, Dordrecht, 429-440.
- Pardo-Igúzquiza E., 1997: *MLREML: a computer program for the inference of spatial covariance parameters by maximum likelihood and restricted maximum likelihood*. *Computers and Geosciences*, 23 (2), 153-162.
- Pardo-Igúzquiza E., 1998: *Maximum likelihood estimation of spatial covariance parameters*. *Mathematical Geology*, 30 (1), 95-108.
- Pardo-Igúzquiza E., 1999: *VARFIT: a fortran-77 program for fitting variogram models by weighted least squares*. *Computers and Geosciences*, 25, 251-261.
- Pardo-Igúzquiza E., Dowd P.A., 2005: *Multiple indicator cokriging with application to optimal sampling for environmental monitoring*. *Computers and Geosciences*, 31, 1-13.
- Peitgen H.-O., Jürgens H., Saupe D., 1997: *Granice chaosu. Fraktale*. Część 1. Wydanie 2 poprawione, PWN, Warszawa, 1-543.
- Prokop P., 2007: *Maksymalne opady w funkcji czasu trwania obserwowane w Polsce [sum.: Maximum precipitation as a function of duration observed in Poland]*. [W:] *Funkcjonowanie geoekosystemów zlewni rzecznych t. 4, Procesy ekstremalne w środowisku geograficznym*, Kostrzewski A., Szpikowski J. (red.), Bogucki Wydawnictwo Naukowe, Poznań, 189-197.
- Pruchnicki J., 1987: *Metody opracowań klimatologicznych*. Państwowe Wydawnictwo Naukowe, Warszawa, 1-203.
- Ratajczak W., 1998: *Metodologiczne aspekty fraktalnego modelowania rzeczywistości*. Instytut Geografii Społ.-Ekon. UAM, Poznań, 1-73.
- Rezacova D., Pesice P., Sokol Z., 2005: *An estimation of the probable maximum precipitation for river basins in the Czech Republic*. *Atmospheric Research*, 77, 407-421.
- Richmond A., 2002: *An alternative implementation of indicator kriging*. *Computers and Geosciences*, 28, 555-565.
- Riva de la J., Pérez-Cabello F., Lana-Renault N., Koutsias N., 2004: *Mapping wildfire occurrence at regional scale*. *Remote Sensing of Environment*, 92, 288-294.

- Salvatore D., Reagle D., 2002: *Statistics and econometrics*. Second Edition. McGraw-Hill Inc., 1-328.
- Semmler T., Jacob D., 2004: *Modeling extreme precipitation events – a climate change simulation for Europe*. *Global and Planetary Change*, 44, 119-127.
- Skaugen T., 1997: *Classification of rainfall into small- and large-scale events by statistical pattern recognition*. *Journal of Hydrology*, 200, 40-57.
- Solow A.R., 1986: *Mapping by simple indicator kriging*. *Mathematical Geology*, 18, 3, 335-352.
- Sornette D., 2006: *Critical Phenomena in Natural Sciences. Chaos, Fractals, Selforganization and Disorder: Concepts and Tools*. Second Edition, Springer, Berlin, Heidelberg, 1-529.
- Srivastava R.M., 1992: *Reservoir characterization with probability field simulation*. *SPE Formation Evaluation*, 7, 4, 927-937.
- Stach A., Tamulewicz J., 2005a: *Analiza struktury pola opadów atmosferycznych Polski metodą składowych głównych powierzchni trendu*. [W:] *Zintegrowany Monitoring Środowiska, Funkcjonowanie geosystemów Polski w warunkach zmian klimatu i różnokierunkowej antropopresji*, Biblioteka Monitoringu Środowiska, 523-540.
- Stach A., Tamulewicz J., 2005b: *Rastrowa baza danych opadów atmosferycznych Polski i jej zastosowania*. [W:] *Zintegrowany Monitoring Środowiska, Funkcjonowanie geosystemów Polski w warunkach zmian klimatu i różnokierunkowej antropopresji*, Biblioteka Monitoringu Środowiska, 541-561.
- Stach A., 1998: *Zmienność przestrzenna właściwości warstwy ornej na niejednorodnym litologicznie stoku morenowym*. *Bibliotheca Fragmenta Agronomica*, 4B, 125-142.
- Stach A., 1999: *Geostatystyczna analiza pola opadów atmosferycznych w dorzeczu Parsęty*. [W:] *Funkcjonowanie geosystemów zlewni rzecznych*, t. 2, Kostrzewski A. (red.), 181-193.
- Stach A., 2002a: *Geostatystyczna identyfikacja mechanizmów transportu roztworów w ciekach*. *Przegląd Naukowy, Inżynieria i Kształtowanie Środowiska*, XI, 2 (25), 130-143.
- Stach A., 2002b: *Metodyka estymacji stężeń roztworów w odplywie rzeczny*. *Przegląd Naukowy, Inżynieria i Kształtowanie Środowiska*, XI, 2 (25), 112-129.
- Stach A., 2002c: *Struktura czasowa wybranych serii pomiarowych Zintegrowanego Monitoringu Środowiska Przyrodniczego*. *Regionalny Monitoring Środowiska Przyrodniczego 3/02*, Kieleckie Towarzystwo Naukowe, Kielce 61-68.
- Stach A., 2005: *A different basis for deriving best estimate chemical fluxes for rivers*. *Sediment Budgets Symposium – VII IAHS Scientific Assembly*, Foz do Iguaçu, PR, Brazil, 4-9 April 2005, Papers for poster presentation, CD-ROM, 1-8.
- Stach A., 2006a: *Analiza przestrzenna danych jakościowych. Aplikacje geologiczne*. *Polskie Towarzystwo Geologiczne i Uniwersytet im. A. Mickiewicza*, Streszczenia referatów ogłoszonych w 2005 roku, Skoczył J. (red.), XIV, 121-132.
- Stach A., 2006b: *Estimation and simulation of the probability field of maximum daily precipitation*. [W:] *Extreme hydrometeorological events in Poland and their impacts. European context*. *International Conference*, Warsaw, Poland, 7-9 December 2006, Book of abstracts, Faculty of Earth Sciences, University of Silesia, Sosnowiec – Warszawa, 67-70.
- Stach A., 2006c: *Spatial structure of maximum daily precipitation in Poland*. [W:] *Extreme hydrometeorological events in Poland and their impacts. European context*. *International Conference*, Warsaw, Poland, 7-9 December 2006, Book of abstracts, Faculty of Earth Sciences, University of Silesia, Sosnowiec – Warszawa, 71-73.
- Stach A., 2006d: *Estymacja i symulacja pola prawdopodobieństwa maksymalnych opadów dobowych*. [W:] *Funkcjonowanie geosystemów zlewni rzecznych 4*, *Procesy ekstre-*

- malne w środowisku geograficznym, konferencja naukowa, Kołobrzeg, 20-22 IX 2006, Streszczenia referatów, Uniwersytet im. A. Mickiewicza, Poznań – Storkowo, 25-29.
- Stach A., 2006e: *Struktura przestrzenna maksymalnych opadów dobowych w strefie Pojezierzy*. [W:] Ogólnopolska konferencja naukowa „Przemiany środowiska geograficznego Polski północno-zachodniej”, 20-21 marca 2006, Poznań, 163-165.
- Stach A., 2007a: *Estymacja i symulacja pola prawdopodobieństwa maksymalnych opadów dobowych* [sum.: Estimation and simulation of the probability field of maximum daily precipitation]. [W:] Funkcjonowanie geosystemów zlewni rzecznych 4, Procesy ekstremalne w środowisku geograficznym, Kostrzewski A., Szpikowski J. (red.), UAM Poznań, Bogucki Wydawnictwo Naukowe Poznań, 51-65.
- Stach A., 2007b: *Spatial structure of maximum daily precipitation in Poland*. *Quaestiones Geographicae* 26A, Adam Mickiewicz University Press, Poznań, 53-59.
- Stach A., 2007c: *Analiza autokorelacji przestrzennej i czasowej danych paleoekologicznych – perspektywy zastosowań*. [W:] *Studia interdyscyplinarne nad środowiskiem i kulturą w Polsce*, Makohonienko M., Makowiecki D., Kurnatowska Z. (red.), Środowisko – Człowiek – Cywilizacja, tom 1. Seria Wydawnicza Stowarzyszenia Archeologii Środowiskowej. Bogucki Wydawnictwo Naukowe, Poznań, 277-299.
- Stach A., Kostrzewski A., Mazurek M., Michalska G.M., Zwoliński Z., 2003a: *Types of stream alimentation in lowland areas of NW Poland. A geostatistical analysis*. *Technical Documents in Hydrology*, 67, UNESCO, Paris 237-243.
- Stach A., Kostrzewski A., Mazurek M., Michalska G.M., Zwoliński Z., 2003b: *Spatial patterns of stream alimentation in lowland areas of NW Poland. A hydrochemical and geostatistical analysis*. *Journal of Hydrology and Hydromechanics*, 51, 3, Bratislava – Praga, 221-240.
- Stach A., Podsiadłowski S., 2000: *Pulverizing and wind erosion as influenced by spatial variability of light morainic soils*. [W:] XV Conference of International Soil Tillage Organization, Fort Worth, Texas, 2-7 VII 2000, CD-ROM, 1-10.
- Stach A., Podsiadłowski S., 2001: *Wpływ erozji eolicznej na teksturę lekkich gleb Niziny Wielkopolskiej*. [W:] *Geneza, litologia i stratygrafia utworów czwartorzędowych*, III, Seria Geografia 64, Wyd. Nauk. UAM, Poznań, 359-379.
- Stach A., Podsiadłowski S., 2002: *Pulverizing and wind erosion as influenced by spatial variability of soils texture*. *Quaestiones Geographicae* 22, Adam Mickiewicz University Press, Poznań, 67-78.
- Stach A., Tamulewicz J., 2003: *Wstępna ocena przydatności wybranych algorytmów przestrzennej estymacji miesięcznych i rocznych sum opadów na obszarze Polski*. [W:] *Funkcjonowanie geosystemów zlewni rzecznych*, t. 3, Obieg wody – uwarunkowania i skutki w środowisku przyrodniczym, Kostrzewski A., Szpikowski J. (red.), Instytut Badań Czwartorzędu i Geokologii UAM, Bogucki Wydawnictwo Naukowe, Poznań, 87-111.
- Stachy J., 1999: *Polska służba hydrologiczna w 25-leciu 1973-1997*. *Gospodarka Wodna*, 8, 277-283.
- StatSoft, Inc. 2004: *Electronic Statistics Textbook*. Tulsa, OK: StatSoft. WEB: <http://www.statsoft.com/textbook/stathome.html>.
- Stępniewska-Podraźka M., 1991: *Kalendarz typów cyrkulacji atmosferycznej (1951-1990)*. IMGW, Warszawa.
- Sun X., Manton M.J., Ebert E.E., 2003: *Regional rainfall estimation using double-kriging of raingauge and satellite observations*. Australian Bureau of Meteorology Research Centre, Research Report, 93, 1-40.



- Trenberth K.E., Dai A., Rasmussen R.M., Parsons D.B., 2003: *The changing character of precipitation*. Bulletin of American Meteorological Society, 84, 1205-1217.
- Trömel S., Schönwiese C.-D., 2007: *Probability change of extreme precipitation observed from 1901 to 2000 in Germany*. Theoretical and Applied Climatology, 87, 29-39.
- Tveito O.E. i zespół 2006: *Spatialisation of climatological and meteorological information by the support of GIS*. COST 719 Final Report, WG2 Spatialisation, Chapter II, Draft Version, Oslo, 26.07.2006, 1-142. ([http://cost719.met.no/wg2\\_final\\_report\\_DRAFT260706.pdf](http://cost719.met.no/wg2_final_report_DRAFT260706.pdf))
- Tveito O.E., Schöner W., 2002: *Applications of spatial interpolation of climatological and meteorological elements by the use of geographical information systems (GIS)*. Norwegian Meteorological Institute, Oslo, met.no, Klima, 28/02, 1-43.
- Ulbrich U., Brücher T., Fink A.H., Leckebusch G.C., Krüger A., Pinto J.G., 2003: *The central European floods of August 2002: Part 2 – Synoptic causes and considerations with respect to climatic change*. Weather, 58, 434-443.
- Ungersböck M., Rubel F., Fuchs T., Rudolf B., 2001: *Bias correction of global daily rain gauge measurements*. Physics and Chemistry of the Earth, B, 26, 411-414.
- Ustrnul Z., Czekierda D., 2000: *Cyrkulacyjne uwarunkowania opadów atmosferycznych w Polsce* [sum.: Circular conditioning of the atmospheric precipitation in Poland]. Wiadomości Instytutu Meteorologii i Gospodarki Wodnej, XXIII (XLIV), 2, 3-22.
- Van Groeningen J.W., Siderius W., i Stein A., 1999: *Constrained optimisation of soil sampling for minimisation of the kriging variance*. Geoderma, 87, 239-259.
- Van Groeningen J.W., Stein A., Zuurbier R., 1997: *Optimization of environmental sampling using interactive GIS*. Soil Technology, 10, 83-97.
- Wackernagel H., 2003: *Multivariate geostatistics. An introduction with applications*. Third, completely revised edition. Springer-Verlag, Heidelberg, 1-387.
- Weber D., Englund E.J., 1992: *Evaluation and comparison of spatial interpolators*. Mathematical Geology, 24, 4, 381-391.
- Weber D., Englund E.J., 1994: *Evaluation and comparison of spatial interpolators II*. Mathematical Geology, 26, 5, 589-603.
- Webster R., Oliver M.A., 1992: *Sample adequately to estimate variograms of soils properties*. Journal of Soil Sciences, 43, 177-192.
- Webster R., McBratney A.B., 1989: *On the Akaike information criterion for choosing models for variograms of soil properties*. Journal of Soil Science, 40, 493-496.
- Webster R., Oliver M., 2001: *Geostatistics for environmental scientists*. John Wiley and Sons, Chichester, 1-271.
- Wehner M.F., 2004: *Predicted twenty-first-century changes in seasonal extreme precipitation events in the parallel climate model*. Journal of Climate, 17, 4281-4290.
- Weisstein E.W., 2006: *Hermite Polynomial*. From MathWorld – Wolfram Web Resource. <http://mathworld.wolfram.com/HermitePolynomial.html>.
- Whelan B.M., McBratney A.B., Minasny B., 2001: *Vesper - Spatial Prediction Software for Precision Agriculture*. [W:] ECPA 2001, Grenier G., Blackmore S. (red.), Proceedings of the 3rd European Conference on Precision Agriculture, agro-Montpellier, Montpellier, France, 139-144.
- Yao T., 1998a: *Conditional spectral simulation with phase identification*. Mathematical Geology 30 (3), 285-308.
- Yao T., 1998b: *SPECSIM: A FORTRAN-77 program for conditional spectral simulation in 3D*. Computers and Geosciences 24, 911-921.
- Yao T., Journel A.G., 1998: *Automatic modeling of (cross) correlogram tables using fast Fourier transform*. Mathematical Geology, 30, 6, 589-615.

- Ying Z., 2000: *IKSIM: A fast algorithm for indicator kriging and simulation in the presence of inequality constraints, hard and soft data*. *Computers and Geosciences*, 26, 493-507.
- Yoo C., Ha E., 2007: *Effect of zero measurements on the spatial correlation structure of rainfall*. *Stochastic Environmental Research and Risk Assessment*, 21, 287-297.
- Zawadzki J., 2005: *Wykorzystanie metod geostatystycznych w badaniach środowiska przyrodniczego [sum.: Using geostatistical methods for environmental studies]*. *Politechnika Warszawska, Prace Naukowe, Inżynieria Środowiska*, 49, 1-134.
- Zimmerman D.L., 1989: *Computationally efficient restricted maximum likelihood estimation of generalized covariance functions*. *Mathematical Geology*, 21 (7), 655-672.
- Zimmerman D., Pavlik C., Ruggles A., Armstrong M.P., 1999: *An experimental comparison of Ordinary and Universal Kriging and Inverse Distance Weighting*. *Mathematical Geology*, 31, 4, 375-390.
- Zolina O., Kapala A., Simmer C., Gulev S.K., 2004: *Analysis of extreme precipitation over Europe from different reanalyses: a comparative assessment*. *Global and Planetary Change*, 44, 129-161.
- Zolina O., Simmer C., Kapala A., Gulev S., 2005: *On the robustness of the estimates of centennial-scale variability in heavy precipitation from station data over Europe*. *Geophysical Research Letters*, 32, L14707.

# XII

## Załączniki

1. Zestawienie stacji meteorologicznych i posterunków opadowych zlokalizowanych na terytorium Niemiec, z których dane sum dobowych opadów zostały wykorzystane w niniejszej pracy

Nr stacji	Nazwa stacji	$\phi$	$\lambda$	Hs m n.p.m.	Uwzględniany okres	Okresy braku danych
11205	JONSDORF, KURORT	50°51'	14°42'	460	01.01.1956 -31.12.1980	
11110	OLBERSDORF	50°52'	14°46'	305	01.01.1969 -31.12.1980	
11115	NIEDERODERWITZ	50°57'	14°44'	308	01.01.1969 -31.12.1980	
11315	HERRNHUT	51°01'	14°45'	344	01.01.1956 -31.12.1980	31.12.1966 -01.01.1969
11305	KEMNITZ	51°04'	14°47'	290	01.01.1969 -31.12.1980	
46236	LOEBAU	51°06'	14°41'	249	01.01.1956 -31.12.1980	31.03.1976 -01.05.1976 30.11.1977 -01.01.1978 30.04.1979 -01.06.1979
11310	GOERLITZ-KUNNERWITZ	51°07'	14°56'	271	01.01.1969 -31.12.1980	
11405	GOERLITZ (WST)	51°10'	14°57'	238	01.01.1956 -31.12.1980	
11410	NEISSEAUE	51°14'	15°01'	180	01.01.1969 -31.12.1980	
46272	HAEHNICHEN	51°22'	14°52'	155	01.01.1956 -31.12.1980	
11415	KRAUSCHWITZ- KL.PRIEBUS	51°28'	14°57'	135	01.01.1969 -31.12.1980	
11445	WEISSKEISSEL	51°30'	14°43'	125	01.01.1969 -31.12.1980	

Nr stacji	Nazwa stacji	$\phi$	$\lambda$	Hs m n.p.m.	Uwzględniany okres	Okresy braku danych
11420	MUSKAU,BAD	51°33'	14°43'	125	01.01.1956 -31.12.1980	
46325	DOEBERN	51°37'	14°36'	135	01.01.1969 -31.12.1980	30.09.1980 -01.11.1980
11425	KLEIN BADEMEUSEL	51°41'	14°45'	87	01.01.1969 -31.12.1980	28.02.1970 -01.04.1970
46320	COTTBUS (WST)	51°46'	14°19'	69	01.01.1956 -31.12.1980	
46330	GROETSCH	51°48'	14°31'	64	01.01.1970 -31.12.1980	
11430	GRIESSEN	51°51'	14°36'	98	01.04.1969 -31.12.1980	31.12.1978 -01.06.1979
11455	GUBEN	51°56'	14°42'	46	01.01.1956 -31.12.1980	31.05.1980 -01.07.1980
12045	RATZDORF	52°04'	14°45'	35	01.01.1969 -31.12.1980	
12050	TREPELN	52°05'	14°33'	110	01.06.1969 -31.12.1980	31.12.1979 -01.02.1980
12005	EISENHUETTENSTADT	52°08'	14°38'	55	01.01.1974 -31.12.1980	
46460	BEESKOW	52°10'	14°15'	43	01.01.1956 -31.12.1980	30.06.1980 -01.08.1980
12040	POHLITZ (WASSERWERK)	52°11'	14°35'	46	01.01.1969 -31.05.1973	
12035	BRIESKOW-FINKENHEERD	52°15'	14°34'	35	01.01.1969 -31.12.1980	31.07.1969 -01.11.1969 31.12.1969 -01.02.1970 30.11.1978 -01.04.1979
12010	FRANKFURT/ODER	52°22'	14°32'	49	01.01.1969 -31.12.1980	
12015	LEBUS	52°26'	14°32'	20	01.01.1969 -31.12.1980	30.06.1972 -01.08.1972
12020	REITWEIN	52°29'	14°37'	15	01.01.1969 -31.12.1980	
13020	SEELOW	52°32'	14°23'	55	01.01.1956 -31.12.1980	
13005	KUESTRIN-KIETZ	52°34'	14°37'	13	01.06.1970 -30.11.1979	
13050	ZECHIN-WOLLUP	52°38'	14°25'	9	01.01.1969 -31.12.1980	
13045	GROSS NEUENDORF	52°42'	14°24'	10	01.01.1969 -31.10.1980	31.01.1969 -01.03.1969
13035	HASELBERG	52°43'	14°02'	107	01.01.1956 -31.12.1980	31.01.1977 -01.03.1977

Nr stacji	Nazwa stacji	$\phi$	$\lambda$	Hs m n.p.m.	Uwzględniany okres	Okresy braku danych
13060	ALTREETZ	52°46'	14°10'	6	01.01.1969 -31.12.1980	
13055	WUSTROW	52°46'	14°13'	6	01.01.1969 -31.12.1980	30.04.1976 -01.06.1976
13075	SCHIFFMUEHLE- NEUTORNOW	52°49'	14°04'	4	01.01.1956 -31.12.1980	31.07.1977 -01.09.1977 31.07.1979 -01.09.1979
14005	ODERBERG/MARK	52°53'	14°03'	30	01.01.1969 -31.12.1980	
14010	LUNOW	52°56'	14°07'	7	01.01.1969 -31.12.1980	
14015	CRIEWEN	53°01'	14°13'	13	01.01.1969 -31.12.1980	29.02.1972 -01.04.1972 31.07.1972 -01.09.1972
14040	ANGERMUENDE (WST)	53°02'	14°00'	54	01.01.1956 -31.12.1980	
14020	SCHWEDT/ODER	53°04'	14°20'	6	01.01.1969 -30.11.1980	30.09.1971 -01.11.1971
14035	VIERRADEN	53°06'	14°19'	4	01.01.1969 -31.12.1980	31.10.1978 -01.05.1979
15005	CASEKOW	53°13'	14°13'	26	01.01.1956 31.12.1980	31.10.1978 -01.01.1979
15010	TANTOW	53°16'	14°21'	31	01.01.1956 -31.12.1980	
21215	PENKUN	53°18'	14°14'	32	01.01.1956 -31.12.1980	
21220	GRAMBOW	53°25'	14°20'	39	01.01.1956 -31.12.1980	
21205	ROTHENKLEMPENOW	53°31'	14°12'	14	01.01.1956 -31.12.1980	
21310	TORGELOW	53°38'	14°01'	10	01.01.1956 -31.12.1980	31.12.1972 -01.04.1973
20010	AHLBECK	53°40'	14°11'	5	01.01.1956 -31.12.1980	
20005	UECKERMUENDE (MNS)	53°45'	14°04'	1	01.01.1956 -31.12.1980	
20205	USEDOM	53°52'	13°55'	2	01.01.1956 -31.12.1980	
20510	MELLENTHIN	53°55'	14°00'	18	01.01.1969 -31.12.1980	31.01.1969 -01.08.1969 31.03.1975 -01.08.1975
22005	SEEBAD HERINGS DORF	53°57'	14°10'	10	01.01.1956 -31.12.1980	31.12.1968 -01.01.1970

## 2. Opis zawartości załączonego dysku DVD

### 2.1. Wstęp

Jednym z zadań projektu badawczego PBZ-KBN-086/P04/2003: „Ekstremalne zdarzenia meteorologiczne i hydrologiczne w Polsce (ocena zdarzeń oraz prognozowanie ich skutków dla środowiska życia człowieka)”, w ramach którego wykonano niniejszą rozprawę było:

- kompilacja, ujednoczenie, weryfikacja i uzupełnienie istniejących baz danych dotyczących zdarzeń ekstremalnych oraz
- ich udostępnienie publiczne do celów naukowych, edukacyjnych i praktycznych.

Dołączona do niniejszej publikacji płyta DVD, zawierająca zarówno dane źródłowe, na których się opierano, jak i najważniejsze wyniki numeryczne, stanowi owego zadania częściową realizację. W trakcie pracy nad analizą struktury przestrzennej MSDO, ich estymacją i symulacją wygenerowano ponad 30 GB danych, z czego na DVD zostało zapisane 1,56 GB (1 684 990 913 bajtów) w 120 folderach zawierających 6151 plików. Oprócz tego, w celu ułatwienia użytkownikowi wizualizacji niektórych wyników, dodano również pliki instalacyjne trzech programów działających w 32-bitowych systemach Windows. Zawartość płyty po zakończeniu całości projektu będzie w drugiej połowie roku 2009 również udostępniona sieciowo poprzez Internet, pod adresem <https://geo.wnoz.us.edu.pl/>.

Niniejszy załącznik zawiera podstawowe informacje o zawartości załączonej płyty DVD oraz wskazówki dotyczące korzystania z niektórych, bardziej skomplikowanych składników tej zawartości. Folderom, plikom i danym w tabelach nadawano takie nazwy, aby po przeczytaniu niniejszej rozprawy użytkownik nie miał większego problemu w rozpoznaniu, jakie dane są w nich zapisane. Brano jednakże również pod uwagę możliwość, że baza danych umieszczona na płycie będzie wykorzystywana oddzielnie, bez książki. Dlatego, w niektórych folderach zamieszczono dodatkowe pliki tekstowe z bardziej szczegółowym opisem ich zawartości. Jakkolwiek autor dołożył maksymalnej staranności, aby materiały te były wolne od błędów, w konsekwentny sposób uporządkowane i opisane, to zdaje sobie sprawę, że przy tak dużej „objętości” danych pewne niedociągnięcia są nieuniknione. Dlatego, zwraca się z uprzejmą prośbą do czytelników niniejszej rozprawy i jednocześnie użytkowników dołączonej do niej bazy danych o przesyłanie na adres [frdstach@amu.edu.pl](mailto:frdstach@amu.edu.pl) wszystkich uwag dotyczących ewentualnych błędów, nieścisłości, czy niejasności.

### 2.2. Używane formaty plików

Większość danych numerycznych, zarówno źródłowych jak i wyników przeprowadzonych analiz, zostało zapisane w plikach skoroszytów MS Excela, wersja 2003 (XLS). W wielu przypadkach tabele były umieszczane w kilku, kilkunastu, a nawet kilkudziesięciu arkuszach tego samego pliku. Nazwy zakładki arkuszy powinny w sposób jednoznaczny identyfikować ich zawartość. Tabele nie mają tytułów, ale stosowane „długie” nazwy poszczególnych zmiennych (kolumn) i przypadków (wierszy) są raczej wystarczające do objaśnienia ich treści.

Znaczna część wyników obliczeń, głównie dotyczących analizy i modelowania struktury przestrzennej MSDO oraz normalizacji danych, jest zawarta w plikach tekstowych (TXT). Są to wyłącznie raporty wygenerowane przez program ISATIS (Bleinès i in. 2007)\*.

---

\* Poszczególne serie danych miesięcznych w obrębie rocznych zbiorów danych MSDO w ISATIS były opisane liczebnikami rzymskimi (I, II, ..., XI, XII). W trakcie eksportu raportów

Pliki te można otworzyć w dowolnym edytorze tekstowym, poczynając od Notatnika dołączanego do systemów Windows. Wszystkie opisy tekstowe w tych dokumentach są w języku angielskim. Zrozumienie ich zawartości nie powinno sprawić trudności osobie, która zapoznała się z treścią niniejszej rozprawy. O ile pliki tekstowe zawierające obliczone semiwariogramy empiryczne i parametry wielomianów Hermite (modeli normalizacji danych) są jedynym źródłem informacji odnośnie tych zagadnień, to modele struktury przestrzennej, ze względu na ich podstawowe znaczenie, zamieszczono również w zbiorczych tabelach zapisanych w skoroszytach Excela.

Za pomocą Irfan View można również przeglądać, modyfikować i drukować pliki JPG (Joint Photographic Experts Group). W formacie tym są to zapisane za pomocą algorytmu kompresji stratnej rastrowe reprezentacje kolorowych rycin i tabel z oryginału rozprawy. W niniejszej publikacji zostały one wydrukowane w odcieniach szarości. Utrudnia to, a w niektórych przypadkach wręcz uniemożliwia pełną interpretację ich treści. Zamieszczenie elektronicznej wersji kolorowych rycin w załączonym DVD ma na celu częściowe zrehabilitowanie tej niedogodności. Kolor został w każdym przypadku zakodowany w 24-bitowym standardzie RGB. Ponieważ na wykorzystanym nośniku nie miało znaczenia, ważne w innych sytuacjach, ograniczenie wielkości plików, zastosowano stosunkowo wysoką rozdzielczość – 600 punktów na cal (600 DPI). Można było dzięki temu dobrze odwzorować nawet drobne szczegóły oryginalnych, skomplikowanych wektorowych rycin. Pliki mają duże rozmiary (maks. 3,54 MB; maksymalny rozmiar po dekompresji aż 242,3 MB), ale przy typowych obecnie wielkościach RAM komputerów osobistych, ich przeglądanie i przetwarzanie nie powinno sprawić problemu. W razie potrzeby do zredukowania rozdzielczości, a poprzez to wielkości plików można użyć również programu Irfan View (menu Image\Resize/Resample Image lub Obraz\Zmiana Rozmiaru obrazu po zmianie języka opisów na polski).

Wszystkie opracowane modele struktury przestrzennej (sum opadów: danych surowych, normalizowanych i kodowanych; terminów rocznych MSDO) przedstawiono również w postaci graficznej. Są to wygenerowane w programie ISATIS metapliki Windows (WMF). Zawierają one oprócz wykresu semiwariancji empirycznych i dopasowanej do nich funkcji także informacje o warunkach obliczeń i zapis parametrów modelu. Plików tych jest na płycie najwięcej, bo w sumie 4900. Ponieważ przez blisko 10 lat, w kolejnych wersjach systemu Windows, WMF był standardowym formatem wymiany graficznych danych wektorowych, pliki te można przeglądać i edytować praktycznie w każdym programie do obróbki grafiki wektorowej oraz wielu – grafiki rastrowej. Można je również wstawiać do dokumentów i tam modyfikować we wszystkich wersjach Microsoft Word. Żeby ułatwić użytkownikowi szybkie przeglądanie zawartości tych plików, na płycie w folderze \\programy\Irfan\_View\_4.22 zamieszczono wersję instalacyjną bezpłatnego programu, który to umożliwia ([www.irfanview.com](http://www.irfanview.com)). Jego zaletą oprócz szybkiego wyświetlania zawartości plików większości stosowanych na świecie formatów grafiki rastrowej i wektorowej jest również możliwość tworzenia miniatur ułatwiających porównywanie wielu wykresów (menu File\Thumbnails lub Plik\Widok Miniatur po zmianie języka opisów na polski).

Wyniki analiz przestrzennych dających w efekcie mapy rastrowe można by również przedstawić w postaci tabelarycznej wraz ze współrzędnymi X i Y każdej komórki (węzła

---

tekstowych wyniki obliczeń były porządkowane alfabetycznie. Dlatego też w plikach tych kolejność miesięcy jest nieco kłopotliwa i na pierwszy rzut oka nielogiczna: III, II, IV, IX, I, Rok, VIII, VII, VI, V, XII, XI i X.

siatki). Ponieważ jednak przyjęta w niniejszej rozprawie standardowa siatka interpolacyjna miała rozdzielczość 1 na 1 km i 319 114 węzłów (patrz dodatek X.6), użycie arkuszy MS Excel stwarzałoby utrudnienie w dalszym przetwarzaniu tych danych (ograniczenie 65 536 wierszy). Konieczny byłby zapis tabeli w plikach tekstowych lub wykorzystanie arkuszy trudno dostępnych programów komercyjnych bez takich ograniczeń jak Excel. Obydwa rozwiązania mają oczywiste wady. W pierwszym przypadku stanowi ją przede wszystkim nieefektywne wykorzystanie przestrzeni dyskowej – plik tekstowy zawierający pojedynczą siatkę interpolacyjną ze współrzędnymi zajmuje około 10 MB. Z drugiej strony dane te są najbardziej użyteczne, kiedy można je szybko przedstawić w postaci kartograficznej. Mając na uwadze te wszystkie czynniki zdecydowano się zamieścić mapy rastrowe w formacie binarnym programu Surfer for Windows (**GRD** wersja 7). Jest to co prawda program komercyjny (<http://www.goldensoftware.com/products/surfer/surfer.shtml>), ale ze względu na dużą jego popularność format **GRD** jest „czytany” przez większość oprogramowania GIS i służącego do analizy rzeźby. Jest wśród nich co najmniej kilka bardzo dobrych produktów bezpłatnych, a wśród nich SAGA GIS (<http://sourceforge.net/projects/saga-gis>), GRASS (<http://grass.itc.it/>), ILWIS (<http://www.ilwis.org/index.htm>). Dwa ostatnie wymienione pakiety wykorzystują do importu plików **GRD** bibliotekę GDAL ([http://www.gdal.org/formats\\_list.html](http://www.gdal.org/formats_list.html)), która została również dostosowana do używania w środowisku R w postaci pakietu RGDAL (<http://cran.r-project.org/web/packages/rgdal/index.html>). Aby ułatwić użytkownikowi korzystanie z siatek interpolacyjnych w formacie **GRD**, na płycie w folderze `\\programy\Surfer 8\` umieszczono plik instalacyjny wersji testowej programu Surfer oraz skrypt języka Visual Basic Grd2Arc.bas. Daje to możliwość ich: 1) szybkiej wizualizacji w postaci map rastrowych, izolinowych, blokdiagramów siatkowych i wypełnionych (cieniowanych) oraz 2) konwersji do tabel tekstowych (**DAT**) i plików rastrowych ArcInfo/ArcGIS w formacie ASCII (**ASC**). W niniejszym tekście nie ma miejsca na szczegółowe instrukcje, jak wykonywać poszczególne operacje w programie Surfer. Nieznający tego programu czytelnik musi skorzystać z informacji zawartych w pliku pomocy. Poniżej umieszczono jednak kilka użytecznych wskazówek, które powinny znacząco przyspieszyć opanowanie tej wiedzy.

- Wersja testowa programu Surfer umożliwia tworzenie i wizualizację na ekranie komputera wymienionych wyżej reprezentacji kartograficznych siatek interpolacyjnych, lecz nie można ich drukować ani zapisywać.

- Każda mapa (izoliniowa, rastrowa, blokdiagram) utworzona na podstawie pliku **GRD** może być uzupełniona innymi danymi, na przykład warstwą z punktami pomiarowymi, granicami regionów, siecią rzeczną lub drogową. Całość zapisywana jest w pliku z rozszerzeniem **SRF**. W większości przypadków, na dołączonej płycie DVD, siatki **GRD** są przedstawione w postaci gotowych map rastrowych z nałożonymi granicami Polski, lokalizacją głównych miast i przebiegiem największych rzek, i zapisane do plików **SRF**.

- Aby wyeksportować zawartość pliku **GRD** do tabeli tekstowej, należy użyć opcji menu Grid\Convert, a następnie wskazać 1) dane źródłowe do konwersji oraz 2) nazwę i format (**DAT**) pliku docelowego. Konwersja obejmuje całą prostokątną siatkę, a więc również te węzły leżące poza granicą Polski, które nie mają określonej wartości. Są one opisane kodem „1.70141E+038”.

- Do eksportu zawartości pliku **GRD** do rastrowego formatu ArcInfo/ArcGIS (**ASC**) posłużyć się można zainstalowanym wraz Surferem programem Scriptor. Po jego uru-



chomieniu należy załadować zapisany na płycie w folderze `\\programy\Surfer_8\` plik `Grd2Arc.bas`, a następnie uruchomić zapisaną w nim procedurę, „klikając” na ikonę paska narzędziowego przedstawiającą zielony, zwrócony wierzchołkiem w prawo trójkąt. W kolejnym kroku podaje się lokalizację i nazwę źródłowego pliku **GRD**. Docelowy plik o rozszerzeniu **ASC** tworzony jest automatycznie w tym samym folderze.

- Najwięcej, bo 325, plików **GRD** reprezentuje wyniki symulacji pól MSDO dla poszczególnych miesięcy i lat wielolecia. W tym wypadku utworzono tylko jedną, przykładową, mapę rastrową zapisaną w formacie **SRF** (plik `Mapa_rok-58_sym01.srf` w folderze `\\wyniki\symulacje_MSDO\`). Aby utworzyć własną mapę na podstawie danych dla dowolnego innego okresu, należy w niej „podmienić” siatkę interpolacyjną. Operację tę wykonuje się poprzez wskazanie i dwukrotne kliknięcie warstwy „rok-58\_sym01” w panelu „Object Manager”, umieszczonym standardowo w lewej części okna programu. Otworzy się wówczas okno zatytułowane: „Map: rok-58\_sym01 Properties”. Na zakładce „General” tego okna, w polu „Input Grid”, należy kliknąć ikonę symbolizującą otwieranie teczki z dokumentami, a następnie wskazać nowy plik **GRD**. Należy jednak pamiętać, że skala wartości przedstawianych na mapie zostanie dostosowana do ich zakresu w nowej, utworzonej właśnie siatce. Aby móc bezpośrednio wizualnie porównywać kolejne tworzone mapy, za każdym razem należy skorygować skalę do jednolitego zakresu. W wypadku 325 symulowanych w przestrzeni danych znormalizowanych pól MSDO są to wartości w przedziale od -5,47 do +5,8. Zmiany skali dokonuje się po kliknięciu na pasek kolorów widoczny w zakładce „General” okna „Map: rok-58\_sym01 Properties”.

## 2.3. Spis zawartości dysku

### 2.3.1. Dane źródłowe

Zawartość dysku jest uporządkowana w czterech głównych folderach: `\\dane`, `\\wyniki`, `\\ryciny_kolorowe` i `\\programy`. W pierwszym z nich zostały zapisane oryginalne dane źródłowe MSDO, a także cyfrowe modele rzeźby terenu Polski. W folderze `\\dane\MSDO_wszystkie\` znajdują się dwa pliki **XLS** zawierające każdy po 25 arkuszy – po jednym na każdy rok 25-lecia 1956-1980. W jednym znajdują się surowe dane pomiarowe; w drugim dane znormalizowane (patrz dodatek X.4). Arkusz z danymi surowymi zawiera najczęściej 22 kolumny: nr punktu pomiarowego, jego nazwę, współrzędne X i Y w układzie 1992, wysokość punktu w m n.p.m. podawaną w Rocznikach oraz średnią wysokość oczka siatki cyfrowego modelu rzeźby, w którym był on zlokalizowany, odczytaną przy 500 m i 1000 m rozdzielczości DEM\*. W kolejnych 13 kolumnach zapisano miesięczne i roczne wartości MSDO w wyrażone w mm. W pozostałych dwóch, wyjątkowo trzech, kolumnach podano daty wystąpienia rocznego MSDO. Dane pozyskane z DWD, a pochodzące z obszaru dawnego NRD znajdują się zawsze w ostatnich wierszach tabeli (por rozdz. V.1 i załącznik XII.1). W skrótych z danymi znormalizowanymi zapisano jedynie numer każdego punktu i jego współrzędne X i Y.

W folderze `\\dane\MSDO_wszystkie\kodowane\` znajduje się 25 plików **XLS**, po jednym dla każdego roku wielolecia, zawierających przekodowane na formę binarną

---

\* Pochodzenie i sposób przygotowania owego DEM omówiono w dalszej części dodatku X.3.

względem 13 wartości progowych wartości pomiarowe MSDO (patrz rozdz. III.2.3 i III.2.4). Punkty pomiarowe identyfikują ich numery i współrzędne X i Y. W następnych 169 kolumnach zapisano dla kolejnych miesięcy i całego roku po 13 wektorów binarnych (0/1). W nagłówkach kolumn podany jest identyfikator okresu i wartość progowa opadu w mm.

W folderze \\dane\MSDO\_roczne\ znajduje się jeden plik skoroszytu Excela z tabelą zawierającą 61 938 wierszy. Są to uporządkowane chronologicznie wszystkie przypadki rocznych MSDO z analizowanego 25-lecia. Oprócz omówionych już poprzednio kolumn zawierają one numer dnia w roku kalendarzowym (Dzień) i numer dnia w roku juliańskim (Dzień 365), w którym wystąpił roczny MSDO.

W folderze \\dane\hipsometria\ zapisano, w plikach GRD programu Surfer, cyfrowe modele rzeźby (DEM) Polski o rozdzielczości 500 i 1000 m. Pierwszy z nich jest również przygotowany do wizualizacji w wyżej wymienionym programie w postaci pliku SRF, zawierającym także wektorowe warstwy granicy państwa, głównych rzek i miast oraz linii brzegowej Bałtyku. Z pomocą owych modeli odczytano średnie wysokości oczek siatki, wewnątrz których znajdowały się punkty pomiaru opadów (zmiennie h500 i h1000 w tabelach z danymi źródłowymi). Z poprzednio wykonanych analiz (Stach, Tamulewicz 2005a) wynika bowiem, że współrzędna Z pobrana dla oczka 1×1 km GTOPO30 (<http://edcdaac.usgs.gov/gtopo30/gtopo30.html>), w którym zlokalizowano posterunek opadowy jest w wypadku analiz obejmujących terytorium całego kraju lepszym predyktorem opadów niż rzędna wysokości podawana dla posterunków w rocznikach Opady Atmosferyczne. W niniejszym opracowaniu wykorzystano inne dane źródłowe do stworzenia obu modeli rzeźby. Był to przede wszystkim DETD2 udostępniony przez koordynatora projektu (Uniwersytet Śląski). Dla obszarów będących poza jego zasięgiem użyto drugiej wersji modelu SRTM ([e0srp01u.ecs.nasa.gov/srtm/version2/SRTM3/Eurasia/](http://e0srp01u.ecs.nasa.gov/srtm/version2/SRTM3/Eurasia/)). Danych hipsometrycznych w analizach zawartych w niniejszej rozprawie nie wykorzystano. Są one jednak potencjalnie ważnym elementem badania struktur czasoprzestrzennych pól opadowych. Mogą być zatem często potrzebne przyszłym użytkownikom przygotowanej bazy danych MSDO.

### 2.3.2. Numeryczne i graficzne wyniki analiz

Ponad 87% objętości załączonego DVD zajmuje zawartość folderu \\wyniki (6024 pliki w 102 folderach zajmujące 1 395 981 999 bajtów).

Wyniki analizy statystycznej i kartograficznej danych MSDO, opisane w rozdziale V niniejszej rozprawy, zamieszczono w folderze \\wyniki\statystyki\. Są to zarówno tabele umieszczone w skoroszytach plików XLS, jak i mapy punktowe oraz rastrowe (pliki GRD i SRF). W osobnych plikach Excela zapisano wyniki obliczeń statystyk globalnych (Statystyki\_MSDO\_klasyczne.xls) i statystyk lokalnych w stosunku do najbliższego sąsiada (Statystyki\_MSDO\_NN.xls). Te drugie dotyczą zarówno odległości pomiędzy punktami pomiarowymi, jak i różnic zmierzonych na nich MSDO. W podkatalogu \\wyniki\statystyki\Mapy\ znajdują się wszystkie mapy, które umieszczono w rozdziale V, wraz z danymi źródłowymi, na podstawie których zostały one wygenerowane.

W katalogu \\wyniki\maskowanie\ umieszczono dwa pliki XLS. W pierwszym z nich – Maskowanie\_56-80.xls – podano za pomocą kodowania binarnego 1/0, które dane MSDO zostały użyte, a które zamaskowano w trakcie analizy i modelowania struktury przestrzennej (patrz dodatek X.5). Materiały te są uporządkowane w tabelach rocznych znajdujących się w osobnych arkuszach skoroszytu. W drugim pliku – Zamaskowane.xls – zawarto podsumowujące procedurę maskowania statystyki i wykresy.

W folderze \\wyniki\normalizacja\ zapisano 25 plików TXT, po jednym dla każdego roku analizowanego wielolecia. Zostały one wygenerowane przez program ISATIS (Bleinès i in. 2007) i zawierają wszystkie parametry procedury normalizacji poszczególnych serii źródłowych miesięcznych i rocznych MSDO (patrz dodatek X.4), a wśród nich współczynniki wielomianów Hermite.

Najważniejsze rezultaty niniejszej rozprawy znajdują się w folderze \\wyniki\struktura\. Materiały umieszczone w podkatalogu \\wyniki\struktura\semiwariogramy empiryczne\ stanowią uporządkowane latami pliki tekstowe (TXT), zawierające dla każdej miesięcznej i rocznej serii danych MSDO wyliczone wartości semiwariancji empirycznych. Każdy z tych plików, wygenerowanych z programu ISATIS, zawiera zarówno semiwariancje danych surowych, normalizowanych, jak i kodowanych. W podkatalogu \\wyniki\struktura\modele semiwariogramów\ jest najwięcej, bo aż 5258 plików. Przede wszystkim zostały tam zapisane w 4 plikach XLS zbiorcze zestawienia parametrów modeli semiwariancji danych surowych, znormalizowanych i kodowanych MSDO oraz obliczone za pomocą programu Vmodel wartości modeli danych znormalizowanych co 500 m (patrz rozdz. VI.6). W folderze tym jest także 25 podkatalogów, po jednym dla każdego roku wielolecia (na przykład: \\wyniki\struktura\modele semiwariogramów\1956\). Zawierają one pliki tekstowe (TXT) z parametrami modeli struktury przestrzennej MSDO, wyeksportowane z programu ISATIS. Dane te, zestawione w wspomnianych wyżej tabelach zbiorczych, były podstawą do opisanych w rozdziałach VI i VII niniejszej rozprawy analiz. W każdym z owych folderów rocznych jest także katalog \ryciny z plikami WMF – po jednym dla każdego wykonanego modelu danych surowych, znormalizowanych i kodowanych. Jak wspomniano w załączniku XII.2.2 są to wykresy z semiwariancjami empirycznymi i dopasowanymi do nich funkcjami modeli.

Wyniki opisanych w rozdziale VIII analiz zmienności czasoprzestrzennej rocznych MSDO zamieszczono w folderze \\wyniki\roczne MSDO\. Ich część, która była treścią podrozdziałów od VIII.1 do VIII.3 zapisano w folderze \\wyniki\roczne MSDO\Prawdopodobieństwo\. Jest tam tabela, na podstawie której wygenerowano rycinę 112, zawierająca zgeneralizowane w pentadach i przedziałach co 5 mm wysokości rocznych MSDO z całego wielolecia (plik Roczne\_MSDO\_pentady\_klasy.xls). W pliku KDE\_10na10nor\_prawdopodobienstwo.xls zamieszczono estymowane, metodą interpolacji, gęstości skupień prawdopodobieństwa wystąpienia danego dnia na obszarze Polski maksymalnych rocznych sum dobowych opadów o określonej wysokości (dane ryc. 114). Obie ryciny w formacie SRF wraz ze źródłowymi siatkami interpolacyjnymi (pliki GRD) zostały zapisane w folderze \\wyniki\roczne MSDO\Prawdopodobieństwo\Ryciny\. Oprócz tego w pliku Statystyki\_prawdopodobienstwa\_rocznych\_MSDO.xls znajdują się przetworzone dane prawdopodobieństwa rocznych MSDO, które przedstawiono w tabeli 25 i na rycinach 115-117. W folderze \\wyniki\roczne MSDO\zmienność przestrzenna\ umieszczono wyniki obliczeń omówionych w rozdziale VIII.4. Są to ryciny 118, 119 i 120 w plikach SRF oraz źródłowe siatki interpolacyjne (GRD). Rezultaty modelowania struktury przestrzennej terminów rocznych MSDO zostały zestawione w folderze \\wyniki\roczne MSDO\struktura przestrzenna terminów\. Analogicznie jak w wypadku analizy struktury wysokości opadów, oprócz zbiorczej tabeli (plik Zbiorczy\_modele\_terminow.xls), zamieszczono źródłowe pliki tekstowe z semiwariancjami empirycznymi i ich modelami oraz pliki WMF z wykresami (w folderze \\wyniki\roczne MSDO\struktura przestrzenna terminow\ryciny\). W trakcie obliczania semiwariancji empirycznych terminów rocznych MSDO również wykonywano maskowanie danych odstających. Zapis tej procedury jest zawarty w dwóch plikach XLS dla lat 1956-1970 i 1971-1980.

W folderze \\wyniki\symulacje\_MSDO\ i jego podkatalogach umieszczono siatki interpolacyjne **GRD** z warunkowymi symulacjami pól MSDO oraz przykłady ich wizualizacji. Szczegółowe wyjaśnienia zawarte są w dodatku X.2 i rozdziale VI.3. Pliki **GRD** uporządkowane są według kolejnych miesięcy roku kalendarzowego i „całych” lat. Ich nazwy jednoznacznie identyfikują serię danych MSDO, z której zostały wygenerowane. Należy pamiętać, że w celu umożliwienia bezpośredniego porównywania zmienności przestrzennej MSDO pomiędzy różnymi okresami, symulacji dokonywano w przestrzeni danych znormalizowanych. Obrazy symulowane odzwierciedlają wiernie dane pomiarowe w ich lokalizacjach oraz z pewnym przybliżeniem rozkład statystyczny wartości i strukturę reprezentowaną przez model semiwariancji. Nie są jednak najlepszą prognozą wartości punktowych w miejscach, w których nie wykonywano pomiarów. W głównym folderze \\wyniki\symulacje\_MSDO\ umieszczono przykłady wizualizacji symulowanych pól MSDO w programach Surfer i Global Mapper (pliki **SRF** i **GMW** – patrz. załącznik XII.2.2 i XII.2.3.3). Znajdują się tam również warstwy wektorowe z granicami Polski, głównymi rzekami i miastami w formacie **SHP** ArcGIS/ArcInfo (kwartet plików o tej samej nazwie i rozszerzeniach **SHP/SHX/DBF/PRJ**). Pliki te można użyć do tworzenia wizualizacji zarówno w Surferze, jak i Global Mapperze.

### 2.3.3. Wersja elektroniczna kolorowych rycin i tabel

Elektroniczne kopie oryginalnych kolorowych rycin i tabel zostały zapisane w folderze \\ryciny\_kolorowe. Są to pliki **JPG** uporządkowane w podfolderach według poszczególnych rozdziałów rozprawy. Ich nazwy są zgodne ze stosowaną ciągłą numeracją rycin i tabel, np. ryc\_25.jpg lub tab\_03.jpg.

### 2.3.4. Pliki instalacyjne programów komputerowych

W folderze \\programy umieszczono w osobnych podfolderach pliki instalacyjne trzech programów komputerowych. Nazwy podfolderów identyfikują ich zawartość. Oprócz wspomnianych już wyżej Surfera dla Windows i IrfanView znajduje się tam również wersja demonstracyjna programu Global Mapper (<http://www.global-mapper.com/>). Jest to, oprócz innych funkcji, uniwersalna przeglądarka i translator plików z kategorii GIS. Może on również posłużyć do wizualizacji wygenerowanych w niniejszym opracowaniu siatek rastrowych GRD w połączeniu z innymi danymi wektorowymi i rastrowymi. Przykładem może być plik środowiska pracy Global Mappera (ang. *Global Mapper workspace file*) umieszczony w folderze \\wyniki\symulacje\_MSDO\, a noszący nazwę rok-80\_sym01.gmw. Zawiera on odwołania do szeregu plików umieszczonych na płycie DVD i generuje obraz symulacji pola rocznych MSDO z roku 1980, nałożony na cyfrowy model rzeźby i uzupełniony o granicę państwa, główne rzeki i miasta. Poprawne wykonanie zawartych w pliku **GMW** poleceń wymaga modyfikacji ścieżek lokalizacji poszczególnych plików z danymi źródłowymi. Ponieważ jest on zapisany w formacie tekstowym, modyfikacji takiej można dokonać za pomocą każdego edytora tekstu, poczynając od Notatnika Windows. Siatki interpolacyjne w formacie **GRD** nie zawierają informacji o stosowanym w nich układzie współrzędnych. W trakcie ich importu do programu Global Mapper należy zatem wybrać w oknie dialogowym „*Select projection for ....*” w polu „*projection*” opcję „Poland 1992/19”. Zamieszczona na płycie DVD wersja demonstracyjna programu Global Mapper zawiera jednak szereg ograniczeń funkcjonalności, między innymi pozwala jednocześnie załadować maksymalnie tylko 4 warstwy danych i nie pozwala korzystać z plików środowiska pracy (**GMW**).

# XIII

---

## List of figures and tables

### Figures

#### Chapter II

1. Time and space domains in MDPT analysis with reference to major atmospheric occurrences (Hense, Friederichs 2006, modified). Legend: 1 - direct spatial range of MDPT data, 2 - direct temporal range of MDPT data, 3 - indirect spatial range of MDPT data, 4 - indirect temporal range of MDPT data. Domains designated "A", "B", "C" and "D" are effects of a combination of the above-mentioned ranges. Detailed explanations in the text.
2. Relation between the area ( $F$ , km<sup>2</sup>) and circumference ( $U^2$ , (km)<sup>2</sup>) of single *Cumulonimbus* clouds and their clusters (generating convective precipitation) in Germany (Hauf, Theusner 2003). Colour is used to present the number of clouds/clusters examined in the particular area and circumference classes. Detailed explanations in the text.

#### Chapter III

3. **h**-scattergrams of the first six lag distances, 2,500 m each, for MDPT data from July 1977. The diagrams also give the range of the distance, mean lag for all the data pairs in the given interval, number of pairs, and the coefficient of linear correlation between measurement results. Worth noting is a decreasing correlation of MDPTs with a growing distance between the sites.
4. Correlogram and semivariogram of MDPT data from July 1977 calculated for 85 lag intervals 2,500 m each. Marked in the diagram are the zero autocorrelation level and the sample variance amounting to 153.72, denoted in geostatistics as  $C(0)$  - autocovariance for a 0 lag distance.
5. Geometric interpretation of the value of semivariogram  $\gamma(\mathbf{h})$  as the mean of all the squared orthogonal distances  $d_{\alpha}$  from the diagonal (the first bisector) of the **h**-scattergram (Goovaerts 1997).
6. Autocovariance and a semivariogram of coded data can be interpreted as a proportion of points (data pairs) which occur in specified parts of the **h**-scattergram: autocovariance - the horizontally hatched area, semivariogram - the vertically hatched area (Goovaerts 1997). In preparing the figure use was made of the maximum annual daily total of precipitation of 18 July 1970. The diagram shows the

- values for 392 pairs of sites separated by a distance of 1.5 to 4.5 km. The threshold value  $z_k$  is the daily total of 100 mm.
7. Profile of rainfall stations (one dimensional data) working in May 1980, used to illustrate methods of the geostatistical analysis of coded data.
  8. MDPTs registered in May 1980 at the stations situated along the profile (Fig. 7).
  9. Cumulative distribution of the probability (*cdf*) of MDPTs registered at rainfall stations throughout Poland in May 1980. Marked in the diagram are daily totals with probabilities 0.2, 0.4, 0.6, 0.8, 0.9 and 0.95 of being exceeded (the 20th, 40th, ..., 95th percentiles).
  10. Threshold values of MDPTs in May 1980 with probabilities of 0.2, 0.4, 0.6, 0.8, 0.9 and 0.95 (Fig. 9, the 20th, 40th, 60th, 80th, 90th and 95th percentiles) mapped onto the profile data (Figs 7 and 8).
  11. MDPTs registered in May 1980 along the analysed profile (Figs 7 and 8), coded into binary data vectors in accordance with formula [12] depending on how much they exceed threshold values determined from the global curve of the cumulative probability distribution (the 20th, 40th, 60th, 80th, 90th and 95th percentiles, Fig. 9).
  12. Empirical semivariograms and their models for coded values (the 20th, 40th, 60th, 80th, 90th and 95th percentiles, Fig. 9) of MDPTs in Poland in May 1980.
  13. Basic semivariogram models employed to describe the spatial structure of MDPTs and the estimation and simulation of their probability field (A). An example of a complex (nested) semivariogram model consisting of three basic models: nugget, exponential, and spherical (B). In the diagrams basic parameters of the models are indicated: sill variance (C), variance of the components ( $C_x$ , partial sills), range (a), and ranges of the components ( $a_x$ , partial ranges).
  14. Examples of problems with modelling the spatial structure of normalised data. A fuller explanation in the text.
  15. Examples of empirical semivariograms of coded data with outlying values of the first lag distances, and their models. A dashed line marks the variance level of data. A fuller explanation in the text.
  16. Examples of chaotic empirical semivariograms of coded data, and models fitted to them. A dashed line marks the variance level of data. A fuller explanation in the text.
  17. Examples of empirical semivariograms of coded data displaying no clear structure, and their models. A dashed line marks the variance level of data. A fuller explanation in the text.

#### Chapter IV

18. Annual precipitation totals on the territory of Poland over the years 1881-2000. Data for the century 1901-2000 come from the TYN CY 1.1 database ([http://www.cru.uea.ac.uk/~timm/cty/obs/TYN\\_CY\\_1\\_1.html](http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1.html)), while the 1881-1990 figures have been derived from Kożuchowski (1985) and converted using linear regression.
19. Mean monthly precipitation totals in the century 1901-2000 (A) and mean deviations from them (B) in three multi-year periods: 1901-1955, 1956-1980 (marked with hatching) and 1981-2000. The data come from the TYN CY 1.1 database ([http://www.cru.uea.ac.uk/~timm/cty/obs/TYN\\_CY\\_1\\_1.html](http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1.html)).
20. Relations between mean annual air temperatures and annual precipitation totals on the territory of Poland in the century 1901-2000 on the basis of the TYN CY 1.1 database ([http://www.cru.uea.ac.uk/~timm/cty/obs/TYN\\_CY\\_1\\_1.html](http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1.html)). The multi-year period under analysis is marked by a different symbol.

21. Mean monthly air temperature in the century 1901-2000 (A), and mean deviations from them (B) in three multi-year periods: 1901-1955, 1956-1980 (marked with hatching) and 1981-2000. Data come from the TYN CY 1.1 database ([http://www.cru.uea.ac.uk/~timm/cty/obs/TYN\\_CY\\_1\\_1.html](http://www.cru.uea.ac.uk/~timm/cty/obs/TYN_CY_1_1.html)).

## Chapter V

22. Distribution and number of precipitation measuring sites in the multi-year period 1956-1980 with identical location co-ordinates: A - distribution of such sites in 1956 (a) and 1975 (b); symbol (c) denotes the location of major towns; B - changes in their number in the successive years of the multi-year period; a darker hue marks the number of cases where as many as three sites had the same co-ordinates.
23. List of the number of MDPT measurements in Poland and Germany used in the present work: a - monthly data from PIHM/IMGW Precipitation Yearbooks, b - annual data from PIHM/IMGW Precipitation Yearbooks, c - monthly data from the German side of the border zone, d - annual data from the German side of the border zone.
24. Statistics of the nearest neighbour distances and randomness of the spatial distribution of measuring points for which MDPTs were analysed in the successive months (A) and years (B) of the multi-year period 1956-80. Legend: (a) - mean distance to nearest neighbour, (b) - standard deviation of distance to nearest neighbour, (c) - skewness of distance to nearest neighbour, (d) - Clark-Evans index of random spatial distribution.
25. Distribution of measuring points in the months and years with their smallest and greatest numbers.
26. Frequency of the number of data points within a radius of 35 km from a node of the 1×1 km interpolation grid for the months and years of the multi-year period 1956-80 with the smallest and greatest density of the measuring network.
27. Frequency of the distance from a node of the 1×1 km interpolation grid to the nearest data point for the months and years of the multi-year period 1956-80 with the smallest and greatest density of the measuring network.
28. Distance of a node of the interpolation grid (1×1 km) from the nearest measuring site in the month with their smallest (II-56: 1,890) and greatest number (VIII-75: 2,707).
29. Distance of a node of the interpolation grid (1×1 km) from the nearest measuring site in the year with their smallest (1956: 1,768) and greatest number (1975: 2,698).
30. Number of measuring sites within a radius of 35 km from each node of the interpolation grid (1×1 km) in the month with their smallest (II-1956: 1,890) and greatest density (VIII-1975: 2,707).
31. Number of measuring sites within a radius of 35 km from each node of the interpolation grid (1×1 km) in the year with their smallest (1956: 1,768) and greatest density (1975: 2,698).
32. Difference in the number of data points within a radius of 35 km from a node of the 1×1 km interpolation grid between August 1975 (maximum number of rainfall measuring sites in the multi-year period 1956-80) and February 1956 (minimum number of sites) - A. Difference in the distance of a node of the 1×1 km interpolation grid to the nearest data point between August 1975 and February 1956 - B.

33. Comparison of empirical distribution functions of all the daily precipitation totals from the 1956-1980 period (1) with the distributions of monthly (2) and annual (3) MDPTs for three sites located in the German side of the border zone (cf. Appendix XII.1).
34. Examples of four types of statistical distribution characterising MDPT data sets. Points represent empirical values, and lines, best-fitted theoretical functions of probability density: VI.1961 - log-normal, VII.1969 - Weibull, XII.1976 - gamma, and V.1978 - of extreme values (Gumbel I).
35. Empirical distribution functions of MDPTs for the particular months of the multi-year period 1956-80: January to June. The thick red line marks mean distributions for the entire data set.
36. Empirical distribution functions of MDPTs for the particular months of the multi-year period 1956-80: July to December. The thick red line marks mean distributions for the entire data set.
37. Cumulative mean distributions of MDPTs for the particular months of the multi-year period 1956-80.
38. Cumulative distributions of annual MDPTs for the particular years of the multi-year period 1956-80. Colour is used to distinguish the mean distribution (for the entire data set) from extreme years significantly departing from the mean.
39. Median, inter-quartile interval and 90% interval of monthly and annual MDPTs in the multi-year period 1956-80.
40. Mean values (b), standard deviation (a) and skewness (c) of MDPTs in the successive months (A) and years (B) of the multi-year period 1956-80.
41. Hypothetical data profiles over time or space illustrating common relations between local means and local variability (Isaaks, Srivastava 1989): (a) - the mean, represented by a horizontal line, and variability are constant, (b) - trend of the mean, variability remains constant, (c) - the mean is constant while variability increases or decreases, (d) - both local means and local variability increase or decrease.
42. Relation between the mean and the standard deviation of MDPTs in the particular months of the multi-year period 1956-80: A - global statistics, B - differences in rainfall amount in comparison with the nearest neighbour. Marked in the diagrams are the regression curve and its 99% confidence interval. Also given is the formula for the function and the coefficient of determination of the regression.
43. Median, inter-quartile interval and range of monthly and annual mean differences in MDPTs at the nearest sites in the multi-year period 1956-80.
44. Values of the mean (a), standard deviation (b) and skewness (c) of differences in MDPTs at the nearest sites in the successive months (A) and years (B) of the multi-year period 1956-80.
45. Trends in the statistics of differences in relation to the nearest neighbour in monthly MDPTs for March in the multi-year period 1956-1980.
46. Point maps of standardised rainfall amounts for months characterised by the smallest (II-1976) and greatest (VII-1965) mean local variability of MDPTs.
47. Point maps of standardised rainfall amounts for the years characterised by the smallest (1962) and greatest (1965) mean local variability of MDPTs.
48. Recurrence (number of cases per year) of specified values of maximum monthly and annual daily totals in Poland determined on the basis of data from the multi-year period 1956-1980.



49. The dates and amounts of maximum annual daily precipitation recorded in Poland and in the German side of the border zone in the multi-year period 1956-80 (61,940 cases).

## Chapter VI

50. Directionless (isotropic, omnidirectional) semivariogram models of normalised data for monthly sets of MDPTs from January to June. The thick red line marks the mean semivariogram for the 25-year period 1956-1980.
51. Directionless semivariogram models of normalised data for monthly sets of MDPTs from July to December. The thick red line marks the mean semivariogram for the 25-year period 1956-1980.
52. 25-year means of isotropic semivariogram models of normalised data for the particular monthly sets of MDPTs (cf. Figs 50 and 51).
53. Directionless semivariogram models of normalised data for annual MDPT sets. The thick red line marks the mean semivariogram for the 25-year period 1956-80. Colour is used to mark a few extreme cases.
54. Single images of conditional simulations of the MDPT field carried out for normalised data from January 1960 and March 1964.
55. Single images of conditional simulations of the annual MDPT field carried out for normalised data from 1965 and 1970.
56. Examples of empirical semivariograms of normalised data and their models for the five distinguished types of spatial structure of MDPTs. Arrows mark rough ranges of the particular structures.
57. Frequency of the distinguished types of spatial structure of MDPTs for monthly (A) and annual (B) data sets.
58. Seasonal variability of the five basic types of spatial structure of MDPTs in the 25-year period 1956-80.
59. Frequencies of values of the nugget variance ( $C_0$ ) and threshold variance of the successive components ( $C_1$ ,  $C_2$  and  $C_3$ ) of models of the spatial structure of normalised monthly MDPT data from the multi-year period 1956-80. Both axes are scaled in a way allowing a direct comparison of the diagrams.
60. Frequencies of the ranges of the successive components ( $A_1$ ,  $A_2$  and  $A_3$ ) of models of the spatial structure of normalised monthly MDPT data from the multi-year period 1956-80. Only the frequency axis is scaled in a way allowing a direct comparison of the diagrams.
61. Frequencies of the ranges of all the components ( $A_1$ ,  $A_2$  and  $A_3$  jointly) of models of the spatial structure of normalised monthly MDPT data from the multi-year period 1956-80.
62. Frequency of the component variance values of the trend assessed for a lag distance of 212.5 km.
63. Share of the particular components in the total spatial variability of monthly MDPTs in the multi-year period 1956-80.
64. Share of the particular components in the total spatial variability of annual MDPTs in the multi-year period 1956-80.
65. Dendrite of similarity among models of the spatial structure of normalised monthly MDPT data obtained using Ward's method of hierarchical clustering. A: the entire dendrite; B-C: its fragments enlarged to show markings of the particu-

lar cases. The rough ranges of the fragments are indicated with red arrows in Figure A. Fragments D-F are presented in Figure 66.

66. Enlarged fragments of the dendrite of similarity among models of the spatial structure of normalised monthly MDPT data obtained using Ward's method of hierarchical clustering. The entire dendrite and rough ranges of its fragments are presented in Figure 65A. Parts B and C of the dendrite can also be found in Figure 65.
67. Dendrite of similarity among models of the spatial structure of normalised annual MDPT data obtained using Ward's method of hierarchical clustering.
68. Criterion of the division of dendrites of similarity among models of the spatial structure of normalised MDPT data (Figs 65A and 76). Legend: a - agglomeration curve of annual data; b - level of division of annual data dendrite into 3 classes with collapsing distance equal to 5; c - agglomeration curve of monthly data; d - level of division of monthly data dendrite into 7 classes with collapsing distance equal to 24; e - level of division of monthly data dendrite into 4 classes with linkage distance equal to 40.
69. Mean semivariance values (thick lines) for the distinguished 7 classes of models of the spatial structure of normalised monthly MDPT data. Hatching is used to mark the range of one standard deviation, and thin lines, models of selected months most typical of each class (cf. maps in Figs 71 to 75).
70. Mean semivariance values (thick lines) for the distinguished 4 types of models of the spatial structure of normalised monthly MDPT data. Hatching is used to mark the range of one standard deviation, and thin lines, models of selected months most typical of each class (cf. maps in Figs 71 to 75).
71. Single images of conditional simulations of the MDPT field carried out for normalised data from January 1980 (an example of class 1 out of 4) and August 1958 (examples of class 2 out of 4 and 2 out of 7).
72. Single images of conditional simulations of the MDPT field carried out for normalised data from January 1974 (an example of class 3 out of 4) and August 1977 (examples of class 4 out of 4 and 5 out of 7).
73. Single images of conditional simulations of the MDPT field carried out for normalised data from March 1978 (an example of class 1 out of 7) and August 1976 (an example of class 3 out of 7).
74. Single images of conditional simulations of the MDPT field carried out for normalised data from January 1973 (an example of class 4 out of 7) and October 1970 (an example of class 6 out of 7).
75. Single image of a conditional simulation of the MDPT field carried out for normalised data from February 1978 (an example of class 7 out of 7).
76. Mean semivariance values (thick lines) for the distinguished 3 types of models of the spatial structure of normalised annual MDPT data. Hatching is used to mark the range of one standard deviation, and thin lines, models of selected years most typical of each class (cf. maps in Figs 77 and 78).
77. Single images of conditional simulations of the MDPT field carried out for normalised annual data from 1964 (an example of class 1 out of 3) and 1975 (an example of class 2 out of 3).
78. Single image of a conditional simulation of the MDPT field carried out for normalised annual data from 1964 (an example of class 3 out of 3).
79. Seasonal variability of mean variance values of the particular components of models of the spatial structure of monthly normalised MDPT data. The 95% confidence interval for the mean is marked.

80. Seasonal variability of mean ranges of the particular components of models of the spatial structure of monthly normalised MDPT data. The 95% confidence interval for the mean is marked. The range of  $A_3$  in January exceeded 168 km, which greatly affected diagram readability; therefore it has been omitted.
81. Relations between  $C_0$  and  $C_1$  means in the particular months of the year in the models of the spatial structure of normalised MDPT data and their standard deviations. Dashed lines mark the 95% interval of regression confidence.
82. Seasonal variability of the share of the nugget variance ( $d$ ) set against the climatological correction factor (CF) of precipitation ( $a$ ), and the correction factors for the years 1996 (b) and 1997 (c) obtained during the BALTEX experiment. Data for curves a, b and c were obtained from Ungersböck et al. (2001) and Rudolf and Rubel (2005). Detailed explanations in the text.
83. Temporal trend in the share of the nugget variance in models of the spatial structure of normalised monthly MDPT data for the multi-year period 1956-80. The abbreviation 'nr. M' in the formula denotes the consecutive month number in the multi-year period.
84. Temporal trend in the variance values of the first component in models of the spatial structure of normalised monthly MDPT data for the multi-year period 1956-1980. The abbreviation 'nr. M' in the formula denotes the consecutive month number in the multi-year period.
85. Temporal trend in the range values of the first component in models of the spatial structure of normalised monthly MDPT data for the multi-year period 1956-1980. The abbreviation 'nr. M' in the formula denotes the consecutive month number in the multi-year period.
86. Seasonal variability of the mean values of residuals from the nugget variance trend in models of the spatial structure monthly normalised MDPT data. The 95% confidence interval for the mean is marked.
87. Relation between absolute nugget variance values of monthly MDPT sets calculated from the normalised data models and those coming from the original models of 'raw' measurement data. Detailed explanations in the text.
88. Variability of absolute variance values of the particular components in models of the spatial structure of monthly MDPTs. The absence of a component was treated as zero.
89. Variability of absolute variance values of the particular components in models of the spatial structure of annual MDPTs. The absence of a component was treated as zero.
90. Seasonal variability of mean values of the root of the variance of the particular components in models of the spatial structure of 'raw' monthly MDPT data. The 95% confidence interval for the mean is marked.
91. Dependence of the nugget variance ( $C_0$ ) of monthly MDPTs on the standard deviation of precipitation total (SD) and the root of the mean squares of differences in precipitation in relation to the nearest neighbour (RMS). The colour of data points denotes the extent of their deviation from the regression surface (blue: < 1 SD of residuals, green: 1 - 2 SD, yellow: 2 - 3 SD, red: > 3 SD).

## Chapter VII

92. Example of standardised empirical semivariograms (A) and their models (B) for a sequence of threshold values (the 1st, 5th, 10th, 20th, ..., 90th, 95th, 99th percentiles) of MDPTs from July 1963. To facilitate comparison, the same colours are

- used, and in the successive diagrams the last semivariogram of the preceding diagram is repeated.
93. Difference between the observed and the expected frequencies of use of the elementary models (spherical, exponential and Gaussian) to model the first component of the spatial structure of coded data in relation to the season (A) and relative rainfall amount (B).
  94. Seasonal variability of mean variance values of the particular components in models of the spatial structure of coded monthly MDPT data. The 95% confidence interval for the mean is marked.
  95. Variability of mean variance values of the particular components in models of the spatial structure of coded monthly MDPT data depending on the relative rainfall amount. The 95% confidence interval for the mean is marked.
  96. Variability of mean range values of the first component in models of the spatial structure of coded monthly MDPT data depending on the relative rainfall amount. The 95% confidence interval for the mean is marked.
  97. Variability of mean variance values of the particular components in models of the spatial structure of coded annual MDPT data depending on the relative rainfall amount. The 95% confidence interval for the mean is marked.
  98. Variability of mean range values of the first component of models of the spatial structure of coded annual MDPT data depending on the relative rainfall amount. The 95% confidence interval for the mean is marked.
  99. Statistical distribution (the minimum, median, third quartile, maximum) of absolute MDPT values for the particular thresholds for monthly (A) and annual (B) data. To improve diagram readability, the right axis (for the 60th-99th percentiles) is scaled in reverse order.
  100. Variability of the nugget variance depending on the absolute rainfall amount in models of the spatial structure of coded monthly MDPT data. Legend: a - data, b - polynomial regression curve with 95% confidence interval marked, c - distribution of data smoothed with spline functions (Loess, B-spline of order 3, 12% of smoothing).
  101. Variability of the variance of the first component depending on the absolute rainfall amount in models of the spatial structure of coded monthly MDPT data. Legend: a - data, b - polynomial regression curve with 95% confidence interval marked, c - distribution of data smoothed with spline functions (Loess, B-spline of order 3, 10% of smoothing).
  102. Variability of the variance of the second component depending on the absolute rainfall amount in models of the spatial structure of coded monthly MDPT data. Legend: a - data, b - polynomial regression curve with 95% confidence interval marked, c - distribution of data smoothed with spline functions (Loess, B-spline of order 3, 10% of smoothing). Eliminated from calculations were four most outlying results. One of them, visible in the diagram, is marked with a symbol different from those used for the other data.
  103. Variability of the range of the first component depending on the absolute rainfall amount in models of the spatial structure of coded monthly MDPT data. Legend: a - data, b - polynomial regression curve with 95% confidence interval marked, c - distribution of data smoothed with spline functions (Loess, B-spline of order 3, 8% of smoothing). Eliminated from calculations were 97 most outlying results. Some of them, visible in the diagram, are marked with symbols different from those used for the other data.

104. Variability of the nugget variance depending on the absolute rainfall amount in models of the spatial structure of coded MDPT data for the particular months, illustrated with the help of polynomial functions. For January and July, the 95% confidence interval for the functions is marked.
105. Variability of the variance of the first component depending on the absolute rainfall amount in models of the spatial structure of coded MDPT data for the particular months, illustrated with the help of polynomial functions. For January and July, the 95% confidence interval for the functions is marked.
106. Variability of the variance of the second component depending on the absolute rainfall amount in models of the spatial structure of coded MDPT data for the particular months, illustrated with the help of polynomial functions. For January and July, the 95% confidence interval for the functions is marked.
107. Variability of the range of the first component depending on the absolute rainfall amount in models of the spatial structure of coded MDPT data for the particular months, illustrated with the help of polynomial functions. For January and July, the 95% confidence interval for the functions is marked.
108. Variability of the nugget variance depending on the absolute rainfall amount in models of the spatial structure of coded annual MDPT data. Legend: a - data, b - polynomial regression curve with 95% confidence interval marked, c - distribution of data smoothed with spline functions (Loess, B-spline of order 3, 20% of smoothing). Eliminated from calculations was the most outlying result. It is marked in the diagram with a symbol different from those used for the other data.
109. Variability of the variance of the first component depending on the absolute rainfall amount in models of the spatial structure of coded annual MDPT data. Legend: a - data, b - polynomial regression curve with 95% confidence interval marked, c - distribution of data smoothed with spline functions (Loess, B-spline of order 3, 20% of smoothing). Eliminated from calculations were four most outlying results. Two of them, visible in the diagram, are marked with symbols different from those used for the other data.
110. Variability of the variance of the second component depending on the absolute rainfall amount in models of the spatial structure of coded annual MDPT data. Legend: a - data, b - polynomial regression curve with 95% confidence interval marked, c - distribution of data smoothed with spline functions (Loess, B-spline of order 3, 20% of smoothing). Eliminated from calculations were three most outlying results. None of them is visible in the diagram.
111. Variability of the range of the first component depending on the absolute rainfall amount in models of the spatial structure of coded annual MDPT data. Legend: a - data, b - polynomial regression curve with 95% confidence interval marked, c - distribution of data smoothed with spline functions (Loess, B-spline of order 3, 20% of smoothing). Eliminated from calculations were five most outlying results. Four of them, visible in the diagram, are marked with symbols different from those used for the other data.

## Chapter VIII

112. Dates of occurrence and MDPTs in Poland in the multi-year period 1956-1980 generalised in pentads and 5-mm rainfall intervals (cf. Fig. 49). Legend: A - number of cases; B - probability.

113. One-dimensional example of kernel density interpolation for 5 points and a bandwidth pattern conforming to the normal distribution. Figures A and B present the effect of the interpolation at various bandwidths. They were plotted using the example given in CrimeStat III (2004).
114. Probability of occurrence of a specified value of annual MDPTs on a given day in Poland, estimated using the kernel density interpolation method - an isarithmic map (top) and a block diagram (bottom). Note: To enhance the visualisation of low probabilities on the map, the class intervals are not equal.
115. Probability of occurrence of a specified value of annual MDPTs and lower on a given day in Poland, and a division into seasons.
116. Temporal profiles of the probability of occurrence of a specified value of an annual MDPT.
117. Profiles of the probability of occurrence (A) and share of a specified value of an annual MDPT (B) for selected days of the year.
118. Mean (A) and the difference between the mean and the median (B) of the date of occurrence of an annual MDPT in the multi-year period 1956-80. Hatching is used to mark the range of tentative regions whose data were used to test the significance of the differences.
119. First (A) and third (B) quartile of the date of occurrence of an annual MDPT in the multi-year period 1956-80.
120. Standard deviation of the date of occurrence of an annual MDPT in the multi-year period 1956-80.
121. Distribution of the dates of occurrence of an annual MDPT in the five tentatively distinguished regions (cf. Fig. 118). Marked in the diagrams are also mean dates and their 95% confidence intervals. The frequency scale is the same in all the diagrams.
122. Directionless (isotropic, omnidirectional) semivariogram models of the dates of occurrence of annual MDPTs. The thick red line marks the mean semivariogram for the 25-year period 1956-1980. Colour is used to mark a few extreme cases.

## **X. Addenda**

123. Typical pattern of analysis of climatological data. Detailed explanations in the text.
124. Pattern of analysis of climatological data accommodating all the available measurement results, and next their spatial and then a temporal synthesis. Detailed explanations in the text.
125. Pattern of a full geostatistical analysis. The elements covered in the present dissertation occur against a grey background. A fuller explanation in the text.
126. Comparison of quantiles (0.001, 0.005, 0.01, 0.02, 0.03, . . . , 0.97, 0.98, 0.99, 0.995, 0.999) of the distribution of empirical annual MDPT data from 1976 with (a) the results of its estimation (expected IK mean: E-type mean) and (b) its simulation (the first three realisations) in a grid of  $1 \times 1$  km.
127. Comparison of an empirical semivariogram (a) and its model (b) of measurement data of annual 1976 MDPTs with empirical semivariograms (c) of its estimation (expected IK mean: E-type mean) and (d) its simulation (the first three realisations). Semivariograms (c) and (d) were calculated for 2,700 data taken at random from a  $1 \times 1$  km grid (319,114 nodes).

128. Variability of the field of annual 1976 MDPTs, estimated (expected IK mean: E-type mean) and simulated in a  $1 \times 1$  km grid (the first three realisations).
129. Schematic diagram of the complete procedure of analysis and modelling of the spatial structure, and of the estimation and simulation of the probability field of maximum daily precipitation for a single set of monthly or annual data. Optional operations are marked with a dashed line. The steps followed in the present work are shown against a grey background.
130. Schematic diagram of the surface of models of simple (A) and nested (B) semivariances with an isotropic autocorrelation of data. Legend: a - raster map in flat projection, b, c, d - images in perspective: raster map, isarithmic map, and three-dimensional wireframe model, respectively.
131. Schematic diagram of the surface of nested models of anisotropic semivariances with a geometric (A) and geometric-zonal (B) type of anisotropy. Legend: a - raster map in flat projection, b, c, d - images in perspective projection: raster map, isoline map, and three-dimensional grid model, respectively.
132. Surfaces (maps) of semivariograms for the years 1956-1967. The range of values presented in each of the maps is different. Each pixel has  $4 \times 4$  km. Detailed explanations in the text.
133. Surfaces (maps) of semivariograms for the years 1968-1980. The range of values presented in each of the maps is different. Each pixel has  $4 \times 4$  km. Detailed explanations in the text.
134. Surfaces (maps) of semivariance for the particular months of 1977. The range of values presented in each of the maps is different. Each pixel has  $4 \times 4$  km. Detailed explanations in the text.
135. Location of two control points of Gaussian anamorphosis using Hermitian polynomials (Bleines et al. 2007). A fuller explanation in the text.
136. Schematic diagram of determining a 'practical' and an 'absolute definition interval' of the model of Gaussian anamorphosis (Bleines et al. 2007). A detailed description in the text.
137. Example of a normalisation (Gaussian anamorphosis) of MDPT data from July 1976. Legend: a - cumulative empirical distribution function (*cdf*), b - distribution curve of normalised data plotted using 30 Hermitian polynomials, c - unrealistic extrapolation of distribution outside the range of empirical data, d - distribution limits of empirical data (3.1 - 166.1 mm), e - interval limits of normalised data (-3.43 - 3.43).
138. Histograms and descriptive statistics of MDPTs from July 1976 : A - raw measurement data in mm, B - data normalised using the methodology described in the text.
139. Spatial structure of MDPT data from July 1976. Legend: a, b - semivariogram and variance of the entire set of raw data (2,685 cases), c, d - semivariogram and variance of selected raw data (2,672 cases), e, f - semivariogram and variance of the entire data set after normalisation, g, h - semivariogram and variance of selected normalised data.
140. Directionless (isotropic, omnidirectional) empirical semivariogram of raw measurement MDPT data from June 1976. It was calculated for 15 lag distances, each 2.5 km wide. With each semivariance value is given the number of data pairs from which it was calculated.

141. Schematic diagram of the procedure of interactive cleaning of a semivariogram based on the cloud diagram of the semivariogram (A) and a location map with symbols proportional to measurement results (B). Detailed explanations in the text.
142. Corrected directionless empirical semivariogram of raw measurement MDPT data from June 1976 - compare with Figure 140. With each semivariance value is given the number of data pairs from which it was calculated; comparison with Figure 140 offers information about the number of eliminated (masked) pairs.
143. Number and proportion of data masked in the course of analysis of the spatial structure against the total number of data in the particular monthly MDPT sets.
144. Number and proportion of data masked in the course of analysis of the spatial structure against the total number of data in the particular annual MDPT sets.
145. Distribution of rainfall stations at which those MDPTs had been measured which were masked in the course of analysis of the spatial structure: A - monthly sets, B - annual sets. For simplicity, marked on map A are measuring points from only five selected years: 1956, 1960, 1965, 1970 and 1975.
146. Distribution of rainfall stations at which MDPTs masked in the course of analysis of the spatial structure of monthly sets had been measured more than 5 times. In the remaining part of Poland, not shown on the map, there were only four such locations: Szczecin - Warszewo (6 times), Warsaw - the National Museum (8), Wyszaków on the Bug (7), and Klusy on the south-eastern margin of the Great Mazurian Lakeland (7).
147. Probability curves of finding a data point within a given radius (distance) from another point for months with the smallest (February 1956) and greatest (August 1975) number of data (details in chapter III). Marked for both data sets is a distance with 95% of probability that the point is situated farther.

## Tables

### Chapter III

1. Approximate proportion of coded semivariograms with anomalous values for first lag distances depending on the month and threshold value (percentile, P). Legend: R - annual MDPTs,  $\bar{x}$  - mean value. Colour is used to mark value intervals: 0-5, 6-10, and then every 10, from 11 to 80.
2. Approximate proportion of coded semivariograms with chaotic fluctuations of values of for the entire distance range analysed, depending on the month and threshold value (percentile, P). Legend: R - annual MDPTs,  $\bar{x}$  - mean value. Colour is used to mark value intervals: 0-5, 6-10, and then every 10, from 11 to 80.
3. Approximate proportion of coded semivariograms with no distinct spatial structure depending on the month and threshold value (percentile, P). Legend: R - annual MDPTs,  $\bar{x}$  - mean value. Colour is used to mark value intervals: 0-5, 6-10, and then every 10, from 11 to 80.

### Chapter V

4. Local statistics of the number of data points within a radius of 35 km from a node of the  $1 \times 1$  km interpolation grid, and distances from the node to the nearest data point for months and years with the smallest and the greatest density of the measuring network.



5. Descriptive statistics of annual MDPTs in the years 1956-1980.
6. Descriptive statistics of MDPTs for the particular months and the entire period 1956-1980.
7. General result of the Mann-Kendall test for the presence of a linear trend in the statistics of monthly MDPTs.
8. General result of the Mann-Kendall test for the presence of a linear trend in the statistics of differences in the rainfall amount of monthly MDPTs in the nearest sites.
9. List of annual MDPTs exceeding 200 mm registered in Poland in the years 1956-80.

## Chapter VI

10. Types of functions (models) used to fit the particular components of composite models of the spatial structure of monthly normalised MDPT data.
11. Summary of the results of testing the significance of seasonal parameter variability in models of the spatial structure of normalised monthly MDPT data using one-way analysis of variance. Differences significant at the  $p < 0.05$  level are marked in bold. Legend: SK - sum of squares, df - number of degrees of freedom,  $\acute{S}K$  - mean squares.
12. General result of the Mann-Kendall test for the presence of a linear trend in the parameters of the spatial structure of coded monthly MDPTs.
13. Descriptive statistics of absolute values of the square root of the variance (in mm of precipitation total) of the particular components of models of the spatial structure of MDPTs.
14. Summary of the results of testing the significance of seasonal parameter variability in models of the spatial structure of monthly MDPT data using one-way analysis of variance. Legend: SK - sum of squares, df - number of degrees of freedom,  $\acute{S}K$  - mean squares.

## Chapter VII

15. General results of multifactor analysis of variance (MANOVA) of the dependence of standardised parameters of coded data models on the season (month, M-c), relative rainfall amount (percentile, P), and their interaction (M-c\*P).
16. Results of multifactor analysis of variance (MANOVA) of the dependence of the particular parameters of coded data models ( $A_1$ ,  $C_0$ ,  $C_1$  and  $C_2$ ) on the season (month, M-c), relative rainfall amount (percentile, P), and their interaction (M-c\*P). Partial variances ( $C_0$ ,  $C_1$  and  $C_2$ ) were standardised prior to calculations.
17. Summary of the results of testing the significance of parameter variability in coded data models of the spatial structure of annual MDPTs depending on the relative rainfall amount using one-way analysis of variance. Differences significant at the  $p < 0.05$  level are marked in bold. Legend: SK - sum of squares, df - number of degrees of freedom,  $\acute{S}K$  - mean squares.
18. Summary of the results of testing the significance of parameter variability in coded data models of the spatial structure of monthly MDPTs depending on the absolute rainfall amount using one-way analysis of variance. Differences significant at  $p < 0.05$  level are marked in bold. Legend: SK - sum of squares, df - number of degrees of freedom,  $\acute{S}K$  - mean squares.

19. Summary of the results of testing the significance of parameter variability in coded data models of the spatial structure of annual MDPTs depending on the absolute rainfall amount using one-way analysis of variance. Differences significant at  $p < 0.05$  level are marked in bold. Legend: SK - sum of squares, df - number of degrees of freedom,  $\bar{SK}$  - mean squares.
20. Summary of the results of modelling relations between the parameters of models of the spatial structure of coded monthly MDPT data and the absolute rainfall amount. Legend: n - number of data considered, Degree - number of model parameters, R - coefficient of determination of model, R-max - maximum possible coefficient of determination irrespective of model type, SEE - standard estimation error, F - F-statistic of model, p - significance level of model, F-LF - F-statistic of lack of fit, p-LF - significance level of lack of fit. Detailed explanations in the text.
21. Summary of the results of modelling relations between the parameters of models of the spatial structure of coded annual MDPT data and the absolute rainfall amount. Legend: n - number of data considered, Degree - number of model parameters, R - coefficient of determination of model, R-max - maximum possible coefficient of determination irrespective of model type, SEE - standard estimation error, F - F-statistic of model, p - significance level of model, F-LF - F-statistic of lack of fit, p-LF - significance level of lack of fit. Detailed explanations in the text.
22. Summary of the results of modelling relations between the parameters of models of the spatial structure of coded MDPT data for the particular months and the absolute rainfall amount. Legend: n - number of data considered, Degree - degree of polynomial, R - coefficient of determination of model, R-max - maximum possible coefficient of determination irrespective of model type, SEE - standard estimation error, F - F-statistic of model, p - significance level of model, F-LF - F-statistic of lack of fit, p-LF - significance level of lack of fit. Detailed explanations in the text.
23. General result of the Mann-Kendall test for the presence of a linear trend in the parameters of the spatial structure of coded monthly MDPTs.
24. General result of the Mann-Kendall test for the presence of a linear trend in the parameters of the spatial structure of coded annual MDPTs.

## Chapter VIII

25. Characteristics of the seasonal probability of occurrence of annual MDPTs. Legend:  $\Sigma P$  (%) - probability sum in season in per cent, minPd (‰) - minimum daily probability in season in per mill, maxPd (‰) - maximum daily probability in season in per mill, minPd (‰) - mean daily probability in season in per mill, gradPd (‰/d) - mean gradient of probability change in season in per mill per day.
26. Statistics of the dates of annual MDPTs in the five regions distinguished. Also presented are the results of Watson's  $U^2$ -tests of the agreement of the observed distributions with von Mises' distribution (corresponding to the normal distribution for cyclic data) and a uniform distribution.
27. Upper part of the table: a comparison matrix of the mean dates of annual MDPT occurrence in the regions using the Watson-Williams F-test. The lower half of the matrix contains values of the F-statistic, and the upper one, an estimated probability of the null hypothesis that the means are equal.  $p$  values given as 0 mean that they were lower than the numerical accuracy of the computer ( $1 \cdot 10^{-16}$ ). Lower part of the table: a comparison matrix of the distributions of annual MDPT

occurrence in the regions using Watson's  $U^2$ -test. The lower half of the matrix contains values of the  $U$  statistic, and the upper one, an estimated probability of the null hypothesis that the distributions are identical.

#### **X. Addenda**

28. Characteristics and applications of geostatistical estimations and simulations.
29. Comparison of the statistics of the original set of measurement data of annual MDPTs from 1976 and the results of its estimation (expected IK mean: E-type mean) and simulation (three first realisations) in a grid of  $1 \times 1$  km.
30. Descriptive statistics of MDPT data from July 1976: VII-76 - original measurement data, VII-76n - normalised data, VII-76bt - data after back conversion.

# Analysis of the spatial and spatial-temporal structure of maximum daily precipitation in Poland in the years 1956-1980

---

## Summary

Among forecasts made on the basis of global and regional climatic models is one of a high probability of an increase in the frequency and intensity of extreme precipitation events. This corresponds with many published analyses of archival measurement series indicating the appearance of such tendencies in the latter half of the 20th century. Learning the regularities underlying the recurrence and spatial extent of extreme precipitation is obviously of great importance, both economic and social. A full insight into those problems can only be gained by combining ground-based point measurements with areal remote-sensing data. Regrettably, it is still hard to formulate generalisations because of the relatively short period for which remote-sensing data are available and difficulties involved in making precise assessments of precipitation intensity on their basis. That is why, irrespective of the new measuring techniques being introduced, it is necessary to perform an analysis and reinterpretation of archival data, making use of the possibilities created by the development of GIS.

In the case of discontinuous meteorological elements - such as precipitation - there can be changes in both, their temporal and spatial characteristics because a change in the point-registered frequency of precipitation in the various intensity or duration classes can be due to a real change in its frequency of occurrence or a change in its spatial range. Hence, the observed increase in the extreme 'amounts' of rain can only be associated with its greater spatial extent, without any changes in its recurrence. Analyses of changes in climate characteristics, which have concentrated almost exclusively on their temporal aspect so far, should also consider a potential evolution of their spatial structure. It is therefore necessary to employ archival multi-year data to identify the basic properties of this structure that could provide a reference point against which possible trends of change might be tested.

The main goal of the dissertation was to analyse regularities underlying spatial and temporal variations in monthly and annual maximum daily precipitation totals (MDPTs) observed in Poland over the years 1956-1980. The basic source material employed was maximum daily precipitation totals recorded at weather and rainfall stations in the particular months of the multi-year period 1956-1980 and published in the Atmospheric Precipitation yearbooks. They were augmented with the date of occurrence of a maximum annual daily precipitation total. Apart from the data from Poland, a similar pluviometric material was ob-

tained for all the measuring sites existing in the years 1956-1980 in the former GDR in a belt of up to 30 km from the Polish border. In total, the database employed contained 747,486 maximum daily precipitation figures in the particular months (2,492 on average) and 61,940 dates/totals of maximum annual daily precipitation (2,478 on average).

These data are specific because apart from being spatially discontinuous, which is typical of precipitation, they are also non-synchronic. A given monthly MDPT figure could have been recorded at any locality on any day of the month. Still, the point-measured MDPTs are not results of independent, totally random phenomena. Each precipitation event takes some time and has a specified spatial range. Hence there is a non-zero probability that it will be recorded in a greater number of localities, and that in at least some of them it will be classed as a periodic MDPT. The present work rests on the hypothesis that the probability of MDPT occurrence displays spatial autocorrelation, and hence spatial continuity.

The main aim of the dissertation was accomplished via several detailed goals:

- determination of the chief statistical features of the analysed sets of monthly and annual MDPTs, including their seasonal and multi-year variability,
- identification and typology of the spatial structure of monthly and annual MDPTs,
- determination of the character and probable origin of events generating MDPTs,
- quantitative assessment of the contribution of the particular events to the overall MDPT figures,
- examination of potential seasonal variations in the spatial structure of MDPTs,
- verification of the hypothesis about differences in the spatial structure of MDPTs in different classes of rainfall amount,
- checking the spatial structure of MDPTs for significant tendencies in the entire period under analysis,
- determination of seasonal differences in the probability of occurrence and the level of annual MDPTs, and
- examination of regional variations in the date of occurrence of annual MDPTs.

The analysis of the spatial structure of MDPTs was carried out in two approaches. In one, use was made of 325 models of spatial structure, one for each of the analysed sets of monthly and annual MDPTs. The models were built on the basis of empirical isotropic semivariograms of normalised data with a range of 212.5 km (85 lag intervals 2.5 km in width).

In the other approach, the analysis of variations in the structure of monthly and annual MDPTs depending on precipitation levels was performed using indicator semivariograms. For each data set, 13 threshold values were determined based on the empirical distribution function: the 1st, 5th, 10th, 20th, ..., 90th, 95th, and 99th percentile (4,225 models in total). Calculations were made for 15 intervals, 2.5 km each.

To assess the probability of the date and level of annual MDPTs, use was made of the method of kernel density estimation. The spatial variability of the dates of occurrence of annual MDPTs was also analysed with the help of the local statistics of their dates (the number of the day in the Julian year).

The MDPT field in Poland is usually the sum of the outcomes of three types of processes operating at various spatial scales: local (<10-20 km), regional (50-150 km), and supra-regional (>200 km). The spatial scales are probably connected with a convective/orographic, a frontal and a 'climatological' genesis of high precipitation (planetary waves?). Their contributions are highly variable. Generally predominant, however, are high daily precipitation totals with a spatial extent of 50 to 150 km connected with mesoscale phenomena and the migration of atmospheric fronts (35-38%). The spatial extent of areas of high local-scale

precipitation usually varies at random, especially in the warm season. At supra-local scales, structures of repetitive size predominate.

The parameters characterising the spatial structure of MDPTs display a marked seasonal cyclicity and variability depending on the precipitation level. The seasonal variability of the proportion of the random component (nugget variance), for instance, is very close to the average multi-year pattern of factors correcting the errors of standard precipitation measurements. Their primary sources are wind-generated turbulence, solid-state precipitation, as well as evaporation- and wetting-related losses. The absolute  $C_0$  value was also found to depend on variations, both global and local, of the MDPT field. Also the relative share of the variance of the first component displayed a highly significant seasonal variability with a maximum in June and May and a minimum in October. Such a yearly cycle probably reflects the changeable proportion of convective precipitation, which attains a maximum in months with the greatest thermal contrasts. The characteristics of the spatial structure of MDPTs depend on a relative rainfall amount connected with its location along the empirical distribution function to a greater extent than on the absolute total.

The multi-year period under analysis was too short to make a reliable assessment of long-term tendencies in the evolution of the climate. Still, the changes found to occur in the elements of the spatial structure, largely connected with local-scale events, show that the hypothesis advanced at the beginning - as to a potential effect of changes in the spatial structure of extreme precipitation events on their point-registered frequency - is a promising research perspective and should continue to be tested.

On the basis of a kernel density estimation, 14 periods were distinguished that varied in the level and gradient of the probability of occurrence of annual MDPTs in various rainfall-amount classes. They last from 12 to 58 days. While an annual MDPT can potentially occur every day on the territory of Poland, the probability of its occurrence between 16 December and 19 March is extremely low. On the other hand, the probability that it will appear between 21 April and 5 November is as high as 95.6%. The period with the greatest chance of occurrence of an annual MDPT lasts from 28 June to 22 August. The probability curve reaches its annual maximum, 1.03%, on 23 July. Its lowest daily values occur at the turn of the year and range from 0.005 to 0.008%. Thus, the difference spans three orders of magnitude. From mid-December to the end of March the predominant annual MDPTs are lower than 20 mm. From mid-June to the last decade of August the most frequent are annual MDPTs of about 40 mm and more. In the remaining part of the year the predominant class of precipitation is that in the interval from 20 to 40 mm. The maximum probability of occurrence of annual MDPTs from the individual rainfall-amount classes falls on the short period between 20 and 26 July. The mean dates of MDPT occurrence in Poland range from 27 June (the 178th day of the year) to 11 August (the 223rd day), with the average date for the entire country being 22 July (the 203rd day). This result differs by a mere day from that read from the probability curve. Variations in the dates of occurrence of annual MDPTs, despite the random character of the highest rainfall, turned out to differ markedly regionally. They seem to be a resultant of factors acting at a variety of scales, one of the major ones being the general oceanic-continental gradient between Western and Eastern Europe. It is modified directly by the Baltic and terrain topography. It was also found that the dates of occurrence of annual MDPTs displayed marked spatial autocorrelation which can be modelled in a standard way. The average ranges of the components of models of the spatial structure of the MDPT dates are similar to those found for the MDPT figures (15.8, 102.8 and 180 km). However, in the case of individual data sets, there are no apparent relations between them. This also holds for the number and type of components making up a model. Probably those dif-

ferences are an effect of a greatly reduced role of the long-term component. The total spatial variability of the date of occurrence of annual MDPTs is largely (in about 69%) shaped by short-range precipitation (from single convective cells) and random factors affecting their internal structure and movement.

While the conducted analyses have thrown much light on the issue of the spatial variability of the highest daily precipitation totals, there is still much to be done. Many of the newly found regularities need a deeper genetic interpretation, which is hard, if not impossible, to conduct without additional data and a wider context of their analysis.

Special studies should be conducted of individual cases typical of the spatial structure classes distinguished. To perform them, it is necessary to have a complete set of daily precipitation totals from the whole of Poland embracing a period of several days before, during, and after the episode analysed, and most importantly, synoptic charts showing the precipitation-generating conditions.

Another potentially fruitful direction, and one not involving a labour-consuming compilation of new databases, is the study of relations between the occurrence of the distinguished spatial structure types of monthly and annual MDPTs and the frequency of various air masses appearing over Poland defined on the basis of calendars of circulation types.

Attention should also be paid to potential regional differences in the spatial structure of maximum daily precipitation. The regularities established in the present work are likely to obtain over a much more extensive area than our country, perhaps even over the entire humid temperate zone. Regional dissimilarities, therefore, may involve not qualitative characteristics, but primarily more or less distinct statistical differences in the frequency of appearance and relative significance of individual components. A study of such differences is currently perfectly possible and scheduled for the next stage of the research.

A promising line of inquiry is also using annual MDPTs to perform co-simulations of both, their level and time of occurrence. This is a logical consequence of the spatial autocorrelation present in both, the set of precipitation totals and the dates of their appearance. As a result, it will be possible to make reliable, fully probabilistic, spatial-temporal assessments of the risk of heavy precipitation at the scale of regions, or even medium-sized river catchments. So far such assessments have been made exclusively on the basis of point data whose spatial representativeness was hard to test. Scale relations allowing a 'conversion' of point probabilities into areal ones have been unknown.

Empirical semivariograms of MDPT data are complex in nature - they consist of two to five elementary models. This is usually interpreted, as has been mentioned in the text, as a result of the operation and summing up of the effects of events occurring at various spatial scales. The method of factorial kriging offers a promising way of differentiation among, and quantitative assessment of, the chief components forming the MDPT field.

*Translated by Maria Kawińska*

