

GEOCHEMICAL SURVEY OF MISSOURI

Chemical Variation of Soils in Missouri
Associated with Selected Levels of the
Soil Classification System



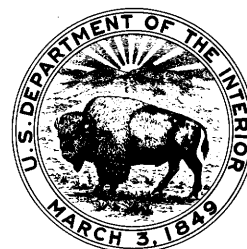
Chemical Variation of Soils in Missouri Associated with Selected Levels of the Soil Classification System

By RONALD R. TIDBALL

GEOCHEMICAL SURVEY OF MISSOURI

GEOLOGICAL SURVEY PROFESSIONAL PAPER 954-B

*An examination of sampling efficiency in
geochemical studies of standard soil
classes in Missouri*



UNITED STATES DEPARTMENT OF THE INTERIOR

THOMAS S. KLEPPE, *Secretary*

GEOLOGICAL SURVEY

V. E. McKelvey, *Director*

Library of Congress Cataloging in Publication Data

Tidball, Ronald R.

Chemical variation of soils in Missouri associated with levels of the soil classification system.
(Geochemical survey of Missouri) (Geological Survey Professional Paper 954-B)

Bibliography: p.

Supt. of Docs. no.: I 19.16:954-B

1. Soils—Missouri—Analysis. 2. Soils—Missouri—Classification. 3. Geochemistry—Missouri.

I. Title: Chemical variation between and within selected levels of the soil classification system... II. Series: Geochemical survey of Missouri. III. Series: United States Geological Survey Professional Paper 954-B.

S599.M52T52 631.4'4'09778 75-619380

For sale by the Superintendent of Documents, U.S. Government Printing Office

Washington, D.C. 20402

Stock Number 024-001-02797-6

CONTENTS

	Page		Page
Abstract	B1	Results	B7
Introduction	1	Geochemical components of variability	7
Acknowledgments	2	Reliability of Suborder means computed from first-stage	
Review of literature	2	data	8
Materials sampled and methods	3	Second-stage sample designs	9
Soil maps	3	Geochemical summaries	11
Analytical procedures	3	Discussion	13
Target and sampled populations	4	Conclusions	14
Statistical model	6	References cited	14

ILLUSTRATION

	Page
FIGURE 1. Map showing soil sampling localities in Missouri	B5

TABLES

	Page
TABLE 1. Analytical methods and lower limits of determination used in analyses for elements in Missouri soils	B4
2. Taxa of Missouri soils sampled in this study	4
3. Nested analysis-of-variance design used for first-stage sampling program of Missouri soils	7
4. Logarithmic variance components for each element apportioned among the several levels of the analysis-of-variance design ..	8
5. Comparison of the observed variance of the mean within Suborders with the maximum acceptable variance of the mean	9
6. Comparison of average compositions of Missouri soil Suborder samples	12

GEOCHEMICAL SURVEY OF MISSOURI

CHEMICAL VARIATION OF SOILS IN MISSOURI ASSOCIATED WITH SELECTED LEVELS OF THE SOIL CLASSIFICATION SYSTEM

By RONALD R. TIDBALL

ABSTRACT

Natural variation in the total elemental composition of the plow zone (approximately 6 in. (15 cm) in depth) of soils in Missouri has been examined using a five-level hierarchical analysis-of-variance design. This study is an initial investigation of the chemical variation in the soils and can serve as a basis for planning future intensive sampling. The design includes three levels (Suborders, Subgroups, and Series) of the Soil Taxonomy developed by the National Cooperative Soil Survey and permits an objective evaluation of the geochemical properties of the taxonomic system. Components of geochemical variance for 31 elements were estimated for differences between soil Suborders, between soil Subgroups, and between soil Series. Significant differences in composition between Suborders were found for 23 elements, including the trace metals arsenic, barium, chromium, copper, gallium, lanthanum, lead, scandium, selenium, strontium, vanadium, ytterbium, and zinc. Such differences, however, account for no more than 45 and commonly less than 35 percent of the natural variation (excluding analytical effects) of any element. Maps based on the distribution of Suborders, therefore, would not be particularly useful in interpretation of geochemical characteristics. Variation between soil Series is, for the most part, much larger than variation within a given Series. Only cobalt, nickel, lead, and zirconium were found to have variation within Series larger than that between Series. Thus, maps based on the distribution of Series might serve as fairly good geochemical maps for most elements.

INTRODUCTION

This report is one in a series describing a variety of field studies undertaken as a part of a general survey of the geochemical environment in Missouri. The survey was undertaken in cooperation with medical scientists of the Environmental Health Surveillance Unit of the University of Missouri and all results have been published in a series of progress reports (U.S. Geological Survey, 1972a-1973), which were compiled at 6-month intervals during the 4 years the work was under way. The work was performed partially in support of the Health Surveillance Units studies on the possible role of environmental geochemistry in health and disease. A preliminary description of the survey and its implications for epidem-

ologic research was published by Connor and others (1972), and a general statement of goals and methods was given by Miesch (1976). Preliminary results of the present study were reported by Tidball (1971) and this paper is a final report.

The role of trace elements in health and disease has been of interest to medical researchers for a long time, but an increasing public awareness of trace-element hazards in the environment has focused attention on the need for data about the expected concentrations and ranges of many trace elements in ordinary environmental materials under natural conditions. Any attempt to describe the geochemistry of soils over a broad area must take into account the possible variations in trace-element concentrations at local scales. Developing such a description without generating an inordinately large analytical load (as might easily happen by sampling a very large landscape unit with a very small sampling interval) is a problem in sampling efficiency.

The principles underlying sampling efficiency in trace-element investigations that were described by Miesch (1976) were in part formulated and tested in much of the Missouri work. Briefly, sampling efficiency is approached through the concept of sampling in more than one stage. That is, an initial or first stage of sampling, which is essentially exploratory, is undertaken to investigate the peculiarities of the particular kind of variation faced by the investigator. Later stages of sampling, if required, are then undertaken using a design specifically directed to delineation of the variation discovered during the first stage. Such a design should produce a stable or reproducible map of that variation. Discovery of natural variation lacking a regional component may make secondary sampling stages unnecessary. Likewise, failure to undertake the first stage subjects the investigator to the risk of either oversampling or undersampling when attempting to des-

cribe or map natural geochemical variation. Maps based on too few samples may be erroneous—that is, not reproducible; maps based on too many samples are financially wasteful—that is, inefficient.

In order to minimize the sampling effort, similar soils should, where possible, be classed together in some minimum number of larger groups rather than attempting to describe each of a large number of individual soils. The hierarchical taxonomy of the Soil Classification System (U.S. Soil Conservation Service, Soil Survey Staff, 1960; 1967, 1968, in press) provides one method of grouping similar soils. The six levels of the hierarchy in descending sequence are as follows: Order, Suborder, Great Group, Subgroup, Family, and Series. This classification seems especially useful for the present purpose for several reasons: (1) if soil composition is uniform within taxonomic groups at any one level of the hierarchy, then estimates derived for a sample of one group (taxon) at that level may be extrapolated to the other groups (taxa) at the same level, thereby reducing the total number of samples required for description; (2) the geographic distribution of many soil taxons in the classification system is at least partially known; (3) some genetic inferences about the distribution of elements may be possible from the definition of the taxon being described; and (4) because the soil classification system provides a standard language for description of soils, communication with other interested researchers is facilitated.

This report examines some properties of the geochemical variation in the classified soils of Missouri. At the time the study was initiated, about 275 (H. E. Grogger, oral commun., 1969) soil Series were recognized in Missouri, even though part of the State has yet to be mapped by modern standards. Because of this diversity of soil types, a first-stage, or exploratory, sampling program was undertaken to estimate some of the properties of the chemical variability in the soils of the State, in hopes of designing an efficient secondary sampling scheme for describing that variation.

The soils at any taxonomic level are grouped by the distinctive attributes they have in common; accordingly the greatest number of general laws can then be deduced about the taxon. Each of the several attributes defining the taxon is ideally the cause, or at least the mark (referred to as a "marker" property), of numerous other properties. This commonality of attributes is called a natural taxonomy (Smith, 1963; Cline, 1963). The attributes which define a taxon are predominantly morphological, and only to a minor extent are chemical attributes included. We can, however, hypothesize that a taxon which is defined by certain fundamental marker attributes has a distinctive chemical composition and that the concentration of elements in the soils of the taxon is controlled by mechanisms related to those for morphologic properties.

The objective of the study on which this report is based was to test the above hypothesis by answering two questions, as follows:

1. Because the number of taxa increase downward in the hierarchical system, what level of the system yields the most information on geochemical variability for the least number of taxa? A level containing relatively few, but geochemically distinct, taxa constitutes an efficient soil classification system for definition of strata in stratified sampling designs.
2. Having defined the most efficient level in the system, how many samples must be collected from each taxon at that level to reliably estimate the mean concentration of elements in each?

ACKNOWLEDGMENTS

The assistance of several staff personnel of the U.S. Soil Conservation Service throughout the State of Missouri in the selection of sampling locations is gratefully acknowledged. In particular, Mr. H. E. Grogger, State Soil Scientist (now retired), compiled the list of soil Series, current to May 1969, at the beginning of this study. The following U.S. Geological Survey chemists made the analyses for this study: Floyd Brown, Joseph Budinsky, G. T. Burrow, J. P. Cahill, I. C. Frost, Johnnie Gardner, Claude Huffman, Jr., Wayne Mountjoy, V. E. Shaw, D. G. Shipley, J. A. Thomas, and E. P. Welsch made the wet chemical analyses and atomic absorption analyses, M. W. Solt, and J. S. Wahlberg did the X-ray fluorescence analyses, and Harriet Neiman, the spectrographic analyses. Special acknowledgment is given to J. G. Boerngen, who gave much assistance in data processing. All statistical computations were performed on the U.S. Geological Survey's IBM 360/65 computer in Washington, D. C., using programs written or adapted by R. N. Eicher, G. I. Selner, Robert Terrazas, and George Van Trump.

REVIEW OF LITERATURE

Chemical variation in the soil may be described as a function of either geography (and thus increases as distances between sampling sites increase) or taxonomy (and thus increases as the degrees of differentiation between soil attributes increase). A review of the literature on the lateral variability of soil was given by Beckett and Webster (1971). The lateral variability of the soil is a function of the relative geographical homogeneity of the soil and will, in theory, define the size of the sampling interval needed for an accurate representation of the soil characteristics: that is, a homogeneous soil can be adequately represented by rather widely separated samples, whereas a heterogeneous soil requires a much shorter sampling interval. In practice, the soil of a region is likely to have several geographic components of variance, each associated with increasing distance between sampling localities, and so

both short and long sampling intervals would be necessary to describe all of the variance. Even a very small area may have a large lateral variation. For example, McKenzie (1955) intensively sampled a "uniform" soil within a 3- by 8-foot (0.9- by 2.4-m) area and concluded that significant differences in copper, gallium, magnesium, manganese, molybdenum, and vanadium occurred between samples separated by a lateral distance of only 8.5 inches (21.5 cm). The development of chemical maps for the major fertilizer nutrient elements over areas the size of a field or a farm was discussed by Vazhenin and others (1961). They found that concentrations of nutrient elements could be related not to taxonomic soil groups, but, rather, to previous history of land use.

Thus, some workers have used designs that have examined variation as a function of distance, whereas others have approached the problem by using taxonomic or other differentiation factors rather than distance. Oertel (1959) discussed a method for estimation of the trace-element status of soils over an area of several thousand square miles. The objective was to express the trace-element status of an area by giving the range of expected concentrations for each soil group in the area. Yet, as Oertel (1959, p. 59) pointed out, "results must be expected" in which very different soil groups, taxonomically speaking, are chemically the same; and also, any one group may differ within itself from one geographic area to another. In other words, variance between taxonomic groups may be less than variance within groups. This is the circumstance that Petersen and Calvin (1965, p. 56) cautioned against: namely, that differences between groups cannot be estimated with any degree of reliability in the presence of excessive within-group variance. Any subtle differences between groups could be defined, therefore, only with a large number of samples; in fact, the number may be so large as to be totally impractical to collect and analyze. Variation between and within several mapping units (based on soil Series) in Ohio was estimated using a nested analysis-of-variance design similar to the one used in the study reported on here (Wilding and others, 1965).

The first-stage sampling program used in the present study is based on a five-level nested analysis-of-variance design which incorporates both taxonomic and geographic components of variability. A detailed discussion of the mechanics and interpretation of the nested analysis of variance in geological applications is given by Olson and Potter (1954), Potter and Olson (1954), and Krumbein and Slack (1956).

MATERIALS SAMPLED AND METHODS

SOIL MAPS

Since 1904, soil survey maps have been published for 67 of the 114 Missouri counties (U.S. Soil Conservation Service, 1968); of these county maps, eight are based on

the 1938 soil classification system (U.S. Department of Agriculture, 1938; revised by Riecken and Smith, 1949; Thorp and Smith, 1949). These eight counties are: Boone (Krusekopf and Scrivner, 1962), Daviess (Grogger, 1964), Holt (Shrader and others, 1953), Jasper (Shrader and others, 1954), Linn (Shrader and others, 1945), Livingston (Rourke and others, 1956), Moniteau (Scrivner and Frieze, 1964), and St. Charles (Shrader and Krusekopf, 1956).

Six county soil surveys—for Worth (Brown, 1968), Pemiscot (Brown, 1971), Dent (Gilbert, 1971), Caldwell (Jeffrey, 1974), Barton (Hughes, 1974), and Lafayette (Jeffrey, 1975) Counties—have been published since the adoption in 1965 of the comprehensive system of soil classification (U.S. Soil Conservation Service, Soil Survey Staff, 1960, 1967, 1968, in press). Unpublished detailed soil surveys have either been completed or are in progress in 31 other counties and 3 ranger districts of the Mark Twain National Forest (J. H. Lee, Missouri State Soil Scientist, oral commun., 1975). Soil-association maps for counties not otherwise covered were published by the University of Missouri Agricultural Experiment Station, as follows: Blackwater and Lamine Townships, Cooper County (Scrivner and Baker, 1961); Mississippi delta area of southeast Missouri (Krusekopf, 1966); Howard County (Scrivner, 1961); and Taney County (Scrivner and Frieze, 1953). Soil-association maps are in preparation by the experiment station staff for 19 counties and for the Missouri River flood plain from Kansas City to Jefferson City (C. L. Scrivner, oral commun., 1970). Statewide soil-association maps present the broad picture (Missouri University Agricultural Experiment Station, 1931; Scrivner and others, 1966; U.S. Soil Conservation Service, in U.S. Geological Survey, 1970, p. 86).

ANALYTICAL PROCEDURES

The analytical procedures used in this study were described by Huffman and Dinnin (1976), Bartel and Millard (1976), Neiman (1976), and Wahlberg (1976). Table 1 gives the procedures and the limits of determination for all of the elements commonly detected in the soil samples. A standard semiquantitative emission spectrographic procedure was used to scan for about 60 elements, but the concentration ranges for more than half of these elements are typically below the limit of determination for this method. Therefore, supplementary methods (listed in table 1) were used for some of the elements that were of particular interest and where the analytical methods were not highly expensive. The spectrographic data for elements with concentrations above the lower limits of determination are reported in six geometric classes per order of magnitude, which have approximate midpoints of 0.00010, 0.00015, 0.00020, 0.00030, 0.00050, 0.00070, 0.0010, . . . , in each order of magnitude, up to 10.0 percent. Logarithmic transformation of data reported in such classes, as suggested by Bennett and Franklin (1954, p. 91), is essential.

TABLE 1.—Analytical methods and lower limits of determination used in analyses for elements in Missouri soils

Element	Analytical Method	Lower limit of determination (parts per million)
Al	X-ray fluorescence	10,000
As	Dithiocarbamate-colorimetric	.2
B	Spectrographic	20
Ba	do	1.5
C (total)	Induction furnace	500
Ca	X-ray fluorescence	1,000
Co	Spectrographic	3
Cr	do	1
Cu	do	1
F	Selective ion electrode	10
Fe	X-ray fluorescence	1,000
Ga	Spectrographic	5
Hg	Instrumental atomic absorption	.01
K	X-ray fluorescence	1,000
La	Spectrographic	30
Mg	Atomic absorption	300
Mn	Spectrographic	1
Na	Atomic absorption	100
Ni	Spectrographic	5
P	X-ray fluorescence	300
Pb	Spectrographic	10
Sc	Spectrographic	5
Se	X-ray fluorescence	.1
Si	do	10,000
Sr	Spectrographic	5
Ti	do	2
V	do	7
Y	do	10
Yb	do	1
Zn	Atomic absorption	10
Zr	Spectrographic	10

TARGET AND SAMPLED POPULATIONS

The soils of the State of Missouri cannot be regarded as a population in a statistical sense until the individuals are defined; the aggregate of these individuals may then be regarded as a population (Krumbein, 1960). The target population—or population of interest—is, therefore, defined as all of the potential samples of the uppermost horizon (0–6 in. (0–15 cm) in depth) of soil profiles throughout the State that had been mapped and classified by soil Series at the time of collection. A soil profile is defined as a vertical cross section of the sequence of soil horizons that extends downward from the ground surface to a variable depth, typically several feet, where the original properties of the parent material are more or less evident. The fact that the target population is specified as being “classified” implies that the soil individuals are defined in terms of taxonomic units. Although the coverage of Missouri soils by soil mapping is still incomplete, it is sufficiently extensive to consider the defined population as a reasonable representation of all soils in the State.

The total variation in chemical composition of the soil population may be subdivided into variance components associated with each taxonomic level. Apportioning the variance in this manner permits a judgment of the rela-

tive efficiency of any taxonomically dependent design that might be used in a second-stage sampling program. The present study includes only three levels of the taxonomic system—Suborder, Subgroup, and Series—because of inadequate representation of some levels among Missouri soils. The source-of-variation designations for the more ideal complete design and for the abridged design used in this study are compared as follows:

<i>Complete design, source of variation</i>	<i>Abridged design, source of variation</i>
Between Orders; within State	
Between Suborders; within Orders	Between Orders and Suborders; within State
Between Great Groups; within Suborders	
Between Subgroups; within Great Groups	Between Great Groups and Subgroups; within Suborders
Between Families; within Subgroups	
Between Series; within Families	Between Families and Series; within Subgroups
Between Localities; within Series	Between Localities; within Series
Between Samples; within Localities	Between Samples; within Localities
Between Analyses; within Samples	Between Analyses; within Samples

In the abridged design some levels are combined within the next lower level. The first three levels evaluate taxonomic variation, the next two levels evaluate local geographic variation, and the last level evaluates analytical variation.

The sampled population is composed of the taxa of Missouri soils as shown in table 2.

TABLE 2.—Taxa of Missouri soils sampled in this study [Leaders (...) indicate missing Series]

Suborder	Subgroup	Series
Aqualf	Aeric ochraqualf	McCune, Dundee, Calwoods.
	Typic ochraqualf	Lightening, McGirk, Moniteau.
Udalf	Typic fragiudalf	Hobson, Loring, Viraton.
	Aquollic hapludalf	Liberal, Lineville, Mystic.
Fluvent	Aquic udifluent	Collins, Dockery, Waubonsie.
	Typic udifluent	Caruthersville, Elsay, Nodaway.
Ochrept	Fluventic dystrochrept	Beulah, Gladden, ...
	Dystric fluventic eutrochrept	Nebo, Ray, ...
Aquept	Aeric fluventic haplaquept	Commerce, Falaya, Westerville.
	Vertic haplaquept	Alligator, Jacob, Sharkey.
Alboll	Argiaquic argialboll	Arbela, Humestone, ...
	Typic argialboll	Edina, ...
Aquoll	Fluventic haplaquoll	Levasy, Myrick, Radley.
	Vertic haplaquoll	Booker, Bosworth, McCroskie.
Udoll	Aquic argiudoll	Lagonda, Grundy, Nevin.
	Aquic fluventic hapludoll	Cooter, Gilliam, Waldron.
Udalt	Ochreptic fragiudult	Nixa, ...
	Typic paleudult	Baxter, Clarksville, Coulstone.

At the top level of the first-stage design, all the Suborders (nine) that had full representation at the lower hierarchical levels were selected systematically. Random selections were made within each succeeding lower level as follows: two Subgroups within each Suborder, three Series within each Subgroup, two localities (areas 100 feet (30 m) square) within each Series, and two samples within each locality. Also, duplicate analyses were performed on 25 randomly selected samples; although variation due to the analytical procedures was estimated independently of the other variance components, it constitutes, in effect, a sixth level in the nested design. The geographic dis-

tributions of the Series are available from both published and unpublished soil survey reports by the U.S. Soil Conservation Service, and the University of Missouri Agriculture Experiment Station. A random selection of sampling localities was achieved by having Soil Survey Staff personnel identify a number of widely scattered occurrences of each Series, then two localities (each representing a separate occurrence) for each Series were selected using a random number table. Randomly selected sampling localities are shown in figure 1. The soils on the major river flood plains have considerable morphologic diversity; accordingly, a disproportionate number of

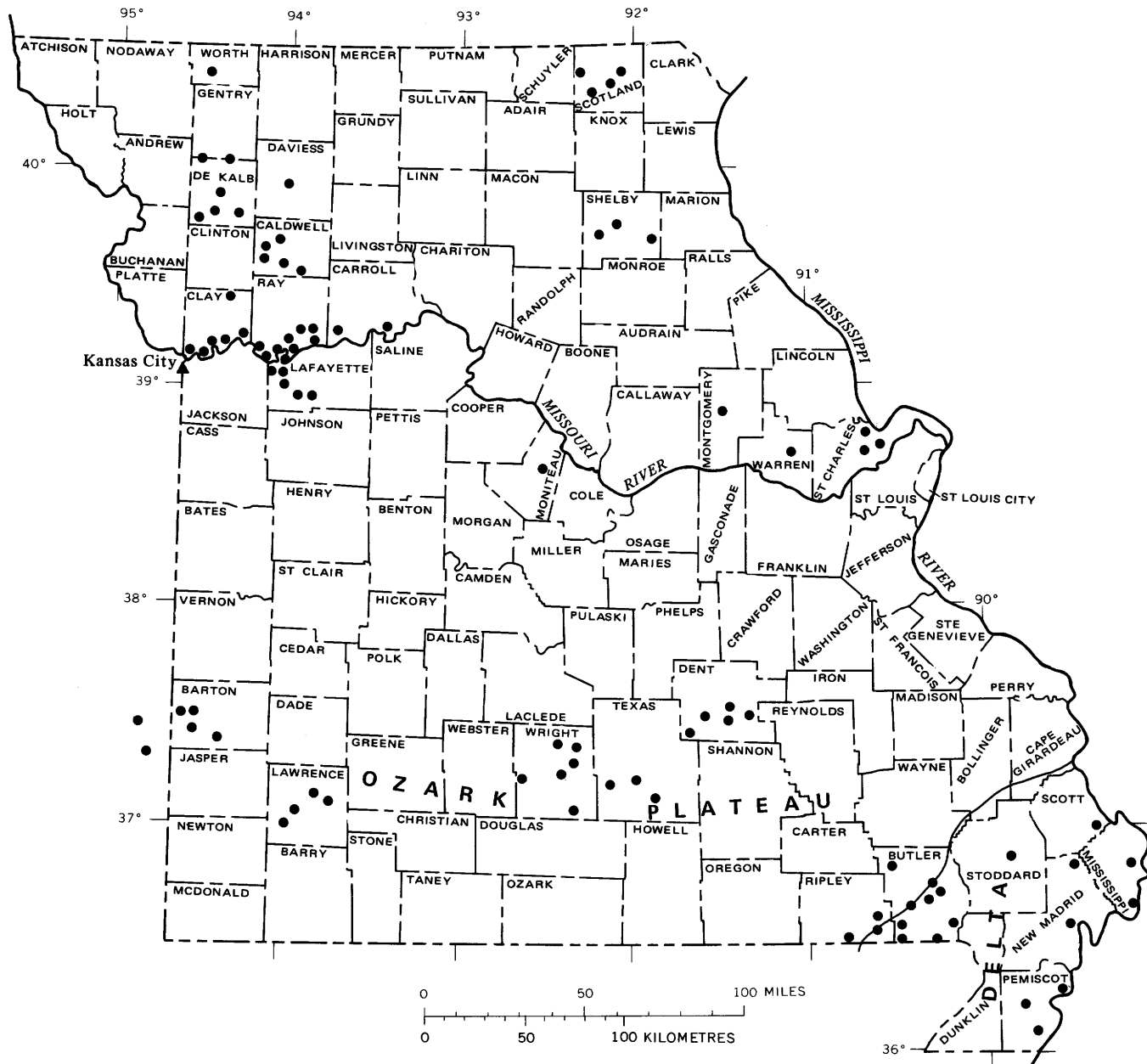


FIGURE 1.—Soil sampling localities (dots) in Missouri.

flood plain soil Series in relation to the total area of alluvium have been recognized and sampled. Sampling localities are, therefore, more concentrated along the Missouri River valley east of Kansas City and in the Mississippi River delta area of southeast Missouri. Areas without representation in the sample were excluded largely because of a lack of modern mapping. Two sampling localities are located in Kansas just across the border from Barton County, Missouri.

The analysis-of-variance procedures (Anderson and Bancroft, 1952, p. 328) are sufficiently flexible that some missing data can be tolerated in lower levels of the design. Although omissions were not originally anticipated or intended, some vacancies in the sample design had to be accepted when certain soil Series could not be found in the field, and no alternate choices were available. These vacancies are indicated in table 2.

Each soil sample consists of a channel composite of the surface horizon (approximately 6 in. (15 cm) in depth). This horizon was selected on the premise that it exercises the greatest influence on plant growth and is the horizon most intimately involved in any phenomena at or near the ground surface. Most soil properties tend to exhibit their greatest variation in the surface horizon and the variation decreases with depth (Petersen and Calvin, 1965, p. 56).

STATISTICAL MODEL

The statistical model used for this study is the standard model of the nested analysis of variance (Anderson and Bancroft, 1952, p. 327-330; Bennett and Franklin, 1954, p. 402-409), as follows:

$$X_{ijklmn} = \mu + \alpha_i + \beta_{ij} + \gamma_{ijk} + \delta_{ijkl} + \eta_{ijklm} + \epsilon_{ijklmn}, \quad (1)$$

where the subscripts are as follows:

- i refers to a Suborder $1 \leq i \leq 9$,
- j refers to a Subgroup, $1 \leq j \leq 2$,
- k refers to a Series, $1 \leq k \leq 3$,
- l refers to a Locality, $1 \leq l \leq 2$,
- m refers to a Sample, $1 \leq m \leq 2$, and
- n refers to an Analysis, $1 \leq n \leq 2$.

The term X_{ijklmn} is a single analytical value for the m th Sample, within the l th Locality, representing the k th Series, and so forth. Equation 1 shows that X_{ijklmn} is equal to μ , the grand mean for the entire soil population, plus the following deviations: α_i is the difference between the grand mean μ and the mean for the i th Suborder; β_{ij} is the difference between the mean for the j th Subgroup and the mean for the i th Suborder; γ_{ijk} is the difference between the mean for the k th Series and the mean for the j th Subgroup; δ_{ijkl} is the difference between the mean for the l th Locality and the mean for the k th Series; η_{ijklm} is the difference between the concentration of the element in the m th Sample and the mean for the l th Locality; and ϵ_{ijklmn} is the difference between the n th Analysis and the concentration in the m th Sample. The last term reflects the analytical precision.

The analytical values for most minor elements have been transformed to the logarithmic form because the change in scale tends to improve the homogeneity of the variances, thus fulfilling one of the basic requirements for the analysis of variance. In addition the frequency distributions of geochemical data at trace concentrations are typically more nearly symmetrical on a logarithmic scale than on an arithmetic scale. The estimation of variance components in the analysis-of-variance method is not seriously affected by slight departures from normality, but tests for the significance of the differences among means may be more affected (Miesch, 1967a, p. A14).

All of the subscripted terms on the right side of equation 1 have means of zero; the purpose of the first-stage sampling program is to estimate the variance of each term. These components of variance have the property of being additive, giving:

$$\sigma_x^2 = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_\delta^2 + \sigma_\eta^2 + \sigma_\epsilon^2, \quad (2)$$

where σ_x^2 is the total variance within the sampled population of soils; all the terms on the right side of the equation represent the parts of the total variance apportioned to each level.

The nested analysis-of-variance design, excluding the analytical (lowest) level is shown in table 3. The subscripted S^2 terms are estimates of the population variance components, σ^2 . Note that each mean square is a progressive summation of weighted variance components from all lower levels; each component is multiplied by a coefficient that represents the number of samples that have contributed to that particular variance estimate. Fractional coefficients appear for S_α^2 and S_β^2 because of the unequal sizes of the classes in levels 2 and 3; the fractional coefficients are considered "effective" sample sizes at these levels and are slightly less than they would be if the class sizes were equal. The analysis of variance is used both to test for significant differences between group means at each level by using the F -ratio and to estimate the variance components.

The F -test for the equality of two variances (Snedecor and Cochran, 1967, p. 117) is used in the analysis-of-variance design to test the hypothesis that the last variance component represented in each summation of the mean square (see table 3) equals zero. For example, consider the variance ratio of levels four and five,

$$\frac{MS_4}{MS_5} = \frac{S_\eta^2 + 2S_\delta^2}{S_\eta^2} = F, \quad (3)$$

where the hypothesis that σ_δ^2 (variance among localities) equals zero is being tested. As S_δ^2 approaches zero, F approaches one. The judgment that calculated F is or is not significantly greater than one at a given probability is based on comparison with critical values from an F -table for known degrees of freedom. If calculated F exceeds the critical F for a probability level of, say 0.05, then it

TABLE 3.—Nested analysis-of-variance design used for first-stage sampling program of Missouri soils

Level	Source of variation ¹	Degrees of freedom	Mean-square estimates	Variance component
1	Between Suborders.....	8	$MS_1 = S\eta^2 + 2S\delta^2 + 4S\gamma^2 + 10.7S\beta^2 + 20.8S\alpha^2$	$S\alpha^2 = \frac{MS_1 - EMS_1}{20.8} \approx \sigma\alpha^2$
2	Between Subgroups: within Suborders.....	9	$EMS_1 = S\eta^2 + 2S\delta^2 + 4S\gamma^2 + 10.7S\beta^2$ $MS_2 = S\eta^2 + 2S\delta^2 + 4S\gamma^2 + 10.2S\beta^2$	$S\beta^2 = \frac{MS_2 - MS_3}{10.2} \approx \sigma\beta^2$
3	Between Series: within Subgroups.....	29	$MS_3 = S\eta^2 + 2S\delta^2 + 4S\gamma^2$	$S\gamma^2 = \frac{MS_3 - MS_4}{4} \approx \sigma\gamma^2$
4	Between Localities: within Series.....	47	$MS_4 = S\eta^2 + 2S\delta^2$	$S\delta^2 = \frac{MS_4 - MS_5}{2} \approx \sigma\delta^2$
5	Between Samples: within Localities	94	$MS_5 = S\eta^2$	$S\eta^2 \approx \sigma\eta^2$
	Total.....	187		

	Population size	Sample size
Suborders in the State.....	$N_\alpha = 9$	$n_\alpha = 9$
Subgroups per Suborder.....	$N_\beta = 4.6$ (average)	$n_\beta = 2$
Series per Subgroup.....	$N_\gamma = 7$ (average)	$n_\gamma = 2.6$ (average)
Localities and samples.....	$N_\delta = N = \infty$	$n_\delta = n_\eta = 2$

follows that the term $S\delta^2$ is greater than zero for some reason other than chance, probably because $\sigma\delta^2$ is greater than zero. Only one time out of 20 should the calculated F exceed the critical value of F for the 0.05 level because of chance alone. A significant variation is present between the localities; that is, some or all localities within any Series differ from one another.

The calculation for the estimates of variance components is shown in table 3. It is possible that on occasion the estimated mean square at some lower level is greater than at the next higher level. This results in a negative estimate of a variance component. A true variance component cannot be negative, but the probability of occurrence of negative estimates of the components is sometimes quite large (Leone and Nelson, 1966, p. 465). By convention, negative variance-component estimates are set equal to zero (Olson and Potter, 1954, p. 43) on the assumption that the real variance is either zero or very small, or that an unusual sampling event occurred (Krumbein and Slack, 1956, p. 757). The convention has been applied in this study with the understanding that the subject has been a matter of some debate (Nelder, 1954; Thompson, 1962; Thompson and Moore, 1963; Anderson, 1965; Leone and Nelson, 1966, p. 464).

RESULTS

GEOCHEMICAL COMPONENTS OF VARIABILITY

The estimated logarithmic variance components associated with each level of the sample design for a number of elements commonly detected in Missouri soils are shown in table 4. The relative distribution of the total variance among the several levels is more readily seen when each variance component is expressed as a percentage.

The percentage distribution of the variance components shows that the analytical procedure accounts for about 5 percent or less of the total variance for elements such as aluminum, calcium, iron, potassium, sodium, selenium, silicon, strontium, and zinc. In contrast, about 80 percent of the total variance in mercury can be attributed to the analytical procedure. Determination of the distribution of mercury in these soils must clearly be based on a more precise analytical method. On the average, nearly two-thirds or more of the total elemental variance can be attributed to the taxonomic variation and the remaining third is associated about equally with geographic localities (within the Series) and analytical procedures.

Most of the elements that exhibit significant variation between Suborders are also estimated to exhibit *no* variation between Subgroups. This feature has an important bearing on any future sampling designs based on soil taxonomy. For such elements, it would be wasteful to use any second-stage designs that employ Subgroups as a sampling unit. Moreover, because the Subgroup component of variance includes variations between Groups as well, it is clear that inclusion of either Groups or Subgroups in future designs for these particular elements would be inefficient. Taxonomic levels used in further sampling should be Suborders or Series. Because sampling of Suborders describes no more than about 45 percent of the total logarithmic variance for any element (notably aluminum), and describes a much smaller percentage for many of the elements studied, it seems reasonable to require that taxonomically related designs must be based on soil Series. Sampling Series and describing differences between them would result in describing about 60–80 percent of the geochemical variance, but this method clearly could be costly, at least in Missouri, because of the large number of Series present.

The actual number of localities required within each Series and the optimum spacing of localities depend on

TABLE 4.—Logarithmic variance components for each element apportioned among the several levels of the analysis-of-variance design

[Numbers in parentheses express each variance component as a percentage of the total logarithmic variance. Significance of the 0.05 probability level for values in the Suborder, Subgroup, Series, and Locality columns is indicated by an asterisk (*). No test for significance was made for Sample and Analysis values]

Element	Suborder (σ_{α}^2)	Subgroup (S_{β}^2)	Series (S_{γ}^2)	Locality (S_{δ}^2)	Sample (S_{η}^2)	Analysis (S_{ϵ}^2)
Al.....	0.01794*	0.0	0.01517*	0.00368*	0.00298	0.00081
	(44)	(0)	(37)	(9)	(7)	(3)
As.....	.00522*	.0	.00925*	.00643*	.00148	.00277
	(21)	(0)	(37)	(26)	(5)	(11)
B.....	.0	.00226	.00764*	.0	.00539	.01104
	(0)	(9)	(29)	(0)	(20)	(42)
Ba.....	.01428*	.0	.01588*	.00320*	.00456	.00556
	(33)	(0)	(37)	(7)	(10)	(13)
C.....	.00603	.00615*	.00333	.00937*	.00036	.01750
	(14)	(14)	(8)	(22)	(1)	(41)
Ca.....	.05437*	.0	.10151*	.01558*	.00276	.00500
	(30)	(0)	(57)	(9)	(1)	(3)
Co.....	.00674	.0	.01526*	.01949*	.00289	.00642
	(13)	(0)	(30)	(38)	(6)	(13)
Cr.....	.00951*	.0	.01699*	.00122	.00074	.00790
	(26)	(0)	(47)	(3)	(2)	(22)
Cu.....	.02593*	.00989	.01484*	.00890*	.0	.01135
	(37)	(14)	(20)	(13)	(0)	(16)
F.....	.02812*	.00705	.01697*	.00986*	.00642	.02110
	(31)	(8)	(19)	(11)	(7)	(24)
Fe.....	.00750*	.00368	.00873*	.00619*	.00956	.00061
	(21)	(10)	(24)	(17)	(26)	(2)
Ga.....	.02851*	.0	.02193*	.00329	.01040	.00496
	(41)	(0)	(32)	(5)	(15)	(7)
Hg.....	.00702	.0	.01262*	.0	.0	.08756
	(7)	(0)	(12)	(0)	(0)	(81)
K ¹08995*	.0	.14513*	.03652*	.01298	.00571
	(31)	(0)	(50)	(13)	(5)	(1)
La.....	.00729*	.0	.01331*	.00546*	.00281	.01127
	(18)	(0)	(33)	(14)	(7)	(28)
Mg.....	.04942*	.0	.04852*	.01142*	.00983	.01633
	(36)	(0)	(36)	(9)	(7)	(12)
Mn.....	.00632	.0	.02847*	.03084*	.00397	.01401
	(8)	(0)	(34)	(37)	(5)	(16)
Na ¹02939*	.0	.06481*	.01051*	.00051	.00182
	(27)	(0)	(61)	(10)	(-1)	(2)
Ni.....	.00951	.00701*	.00650	.01966*	.01137	.01086
	(15)	(11)	(10)	(30)	(17)	(17)
P.....	.01435*	.0	.00828*	.01085*	.0	.02284
	(25)	(0)	(15)	(19)	(0)	(41)
Pb.....	.00193*	.0	.00259	.00331*	.00403	.00404
	(12)	(0)	(17)	(21)	(25)	(25)
Sc.....	.02466*	.00534	.01649*	.00354*	.01062	.00228
	(39)	(8)	(26)	(8)	(17)	(4)
Se.....	.07265*	.00377	.04240*	.03246*	.07610	.02470
	(29)	(1)	(17)	(13)	(30)	(10)
Si ¹	6.109*	.81605	4.814*	1.757*	4.5970	78500
	(41)	(6)	(33)	(12)	(3)	(5)
Sr.....	.03145*	.0	.03374*	.01302*	.00596	.00462
	(35)	(0)	(38)	(15)	(7)	(5)
Ti.....	.00210	.0	.01237*	.00124	.00612	.00547
	(8)	(0)	(45)	(5)	(22)	(20)
V.....	.02399*	.00037	.01506*	.01113*	.00566	.00598
	(39)	(< 1)	(24)	(18)	(9)	(10)
Y.....	.00202	.0	.00519*	.00496*	.00228	.00403
	(11)	(0)	(28)	(27)	(12)	(22)
Yb.....	.00249*	.0	.00435*	.00374*	.00186	.00248
	(17)	(0)	(29)	(25)	(12)	(17)
Zn.....	.02374*	.00374	.01304*	.01329*	.00520	.00042
	(40)	(6)	(22)	(22)	(9)	(< 1)
Zr.....	.00273	.00798	.01968*	.01084*	.00811	.00935
	(5)	(14)	(34)	(18)	(14)	(15)

¹Variance components are arithmetic rather than logarithmic.

the total variance within the Series as well as on the relative magnitudes of the variance between localities and between samples within localities. If the variance between localities is small enough to be ignored—as is the case for about one-third of the elements—then such elements could be described by sampling within any one randomly chosen locality in each Series. If the variance between localities is relatively large—as in the case of cobalt, manganese, nickel, or yttrium—then samples must be obtained from a number of localities within each Series.

RELIABILITY OF SUBORDER MEANS COMPUTED FROM FIRST-STAGE DATA

The previous section discussed in a general way the understanding to be gained about elemental variation from first-stage sampling. The components of variance

can also be used to assess the reliability of any mean values that might be computed from the first-stage results, as well as to estimate the additional sampling necessary to insure a given reliability in such means.

The following discussion will focus on the Suborders as the soil mapping unit of principal interest, rather than on lower taxonomic levels, for two reasons. First, the mapping units of Suborders cover a much larger area; hence, large areas can probably be described with a minimum sampling effort. Secondly, more reliable summary statistics for several elements may be calculated from the first-stage data because of the greater quantity of data per Suborder.

The reliability of the estimate of a Suborder mean depends on the ratio of the variance between Suborders (σ_{α}^2 from equation 2) to the total variance within Suborders ($\sigma_{\beta}^2 + \sigma_{\gamma}^2 + \sigma_{\delta}^2 + \sigma_{\eta}^2 + \sigma_{\epsilon}^2$). This is the variance ratio, v , used by Miesch (1976):

$$v = \frac{Nv}{Dv} = \frac{S_{\alpha}^2}{S_{\beta}^2 + S_{\gamma}^2 + S_{\delta}^2 + S_{\eta}^2 + S_{\epsilon}^2}, \quad (4)$$

where the S^2 terms are sample estimates of the respective population parameters, σ^2 . The ratio, v , is a fundamental property of the soil and is dependent both on the inherent variation of the soil and on the analytical procedure used to estimate that variation. The larger v is, the smaller is the number, n_r , of randomly selected specimens needed to make a reliable estimate of the mean of each Suborder. The terms v and n_r are related to the F -statistic, as follows:

$$F = 1 + n_r v, \quad (5)$$

where n_r is adjusted until F takes a value equal to or greater than the critical F -value for 1 and $2n_r - 2$ degrees of freedom at a specified probability level. Graphs from which n_r may be read as a function of v at given probability levels are shown by Miesch (1976).

After determining the value of n_r , the error variance of the Suborder means, E_r , may be calculated, as follows:

$$E_r = \frac{S_{\beta}^2 + S_{\gamma}^2 + S_{\delta}^2 + S_{\eta}^2 + S_{\epsilon}^2}{n_r}. \quad (6)$$

This equation, which is based on a random sampling design, has the boundary condition that n_r is the minimum number of samples required for distinguishing between two Suborder means, and, accordingly, E_r becomes a maximum acceptable error variance.

The observed error variance of the mean obtained from a nested design, E_s (s subscript denotes stratified), is calculated according to a rearranged version of the general equation of Cochran (1963, p. 268) as follows:

$$E_s = (1 - f_{\beta}) \frac{S_{\beta}^2}{n_{\beta}} + (1 - f_{\beta} f_{\gamma}) \frac{S_{\gamma}^2}{n_{\beta} n_{\gamma}} + (1 - f_{\beta} f_{\gamma} f_{\delta}) \frac{S_{\delta}^2}{n_{\beta} n_{\gamma} n_{\delta}} + (1 - f_{\beta} f_{\gamma} f_{\delta} f_{\eta}) \frac{S_{\eta}^2}{n_{\beta} n_{\gamma} n_{\delta} n_{\eta}} + (1 - f_{\beta} f_{\gamma} f_{\delta} f_{\eta} f_{\epsilon}) \frac{S_{\epsilon}^2}{n_{\beta} n_{\gamma} n_{\delta} n_{\eta} n_{\epsilon}} \quad (7)$$

where the quantities in parentheses are finite population correction factors, and $f_{\beta} = n_{\beta}/N_{\beta}$, the sampling fraction at the Subgroup level, $f_{\gamma} = n_{\gamma}/N_{\gamma}$, the sampling fraction at the Series level, and so forth. A correction factor is applied when more than 10 percent of the population is sampled (Cochran, 1963, p. 24). The values of f_{δ} , f_{η} , and f_{ϵ} in the present study are very small; therefore, these correction factors may be ignored and the final equation is as follows:

$$E_s = (1-f_{\beta}) \frac{S_{\beta}^2}{n_{\beta}} + (1-f_{\beta}f_{\gamma}) \frac{S_{\gamma}^2}{n_{\beta}n_{\gamma}} + \frac{S_{\delta}^2}{n_{\beta}n_{\gamma}n_{\delta}} + \frac{S_{\eta}^2}{n_{\beta}n_{\gamma}n_{\delta}n_{\eta}} + \frac{S_{\epsilon}^2}{n_{\beta}n_{\gamma}n_{\delta}n_{\eta}n_{\epsilon}} \quad (8)$$

The error variance, E_r , represents a maximum acceptable value or "ceiling" under which the observed error variance, E_s , must fall, according to the inequality, as follows:

$$E_s \leq E_r \quad (9)$$

No further sampling is necessary if E_s is equal to or less than E_r . Otherwise, additional second-stage sampling is required; that is, the n 's in equation 8 must be increased so that the estimate of E_s is decreased as required by expression 9. The several parameters discussed above, v , n_r , E_r , and E_s , are shown in table 5 for each element.

Assuming that the purpose of estimating means is to describe differences among Suborders, the first-stage data

TABLE 5.—Comparison of the observed variance of the mean within Suborders with the maximum acceptable variance of the mean

[v , observed variance ratio according to equation 4; n_r , required minimum number of randomly selected soil samples within each Suborder according to equation 5 for a 95 percent probability level; E_r , maximum acceptable value of the variance of the Suborder means according to equation 6; E_s , observed variance of the mean within Suborders based on the first-stage data according to equation 8. Asterisk indicates that the inequality of expression 9 is fulfilled. Leaders (...) indicate that no calculation could be made]

Element	v	n_r	E_r	E_s
Al.....	.79	6	.00377	.00296*
As.....	.26	13	.00153	.00225
B.....	000240
Ba.....	.49	8	.00365	.00322*
C.....	.16	20	.00184	.00361
Ca.....	.44	9	.01387	.01812
Co.....	.15	21	.00210	.00463
Cr.....	.35	10	.00268	.00308
Cu.....	.58	7	.00643	.00632*
F.....	.46	8	.00768	.00649*
Fe.....	.26	13	.00221	.00352
Ga.....	.70	6	.00676	.00447*
Hg.....	.07	40	.00244	.00414
K ¹45	8	.02479	.02763
La.....	.22	15	.00219	.00308
Mg.....	.57	7	.01230	.00979*
Mn.....	.08	36	.00209	.00808
Na ¹38	10	.00777	.01153
Ni.....	.17	19	.00292	.00573
P.....	.34	10	.00420	.00293*
Pb.....	.14	22	.00064	.00103
Sc.....	.64	7	.00547	.00507*
Se.....	.40	9	.01994	.01528*
Si ¹71	6	1.439	1.2168*
Sr.....	.55	7	.00819	.00709*
Ti.....	.08	36	.00066	.00254
V.....	.63	7	.00546	.00402*
Y.....	.12	38	.00043	.00152
Yb.....	.20	16	.00078	.00121
Zn.....	.67	6	.00595	.00470*
Zr.....	.05	55	.00102	.00709

¹Based on arithmetic data rather than logarithmic.

are adequate to estimate Suborder means for 13 elements (as indicated by the asterisks in table 5). The value of n_r required to make adequate estimates for the remaining elements could be as high as 55 (for zirconium) per Suborder if all of the elements are to be described. However, such an element must necessarily be very important to warrant the excessive expense of sampling over and above the more modest requirements of the other elements. Therefore, in the interest of economy, it may be appropriate to limit n_r to a more moderate value that would not so excessively oversample for most elements, and which would require sacrificing only a few possibly unimportant elements for which the sampling would be inadequate. The total sample loads for several values of n_r in a random design including all of the 13 Suborders recognized in Missouri are as follows:

n_r	Number of elements described	Total number of samples
10	18	130
15	20	195
22	25	286
37	26	481
41	29	533
55	30	715

It is evident that a point of diminishing return occurs beyond n_r equal to about 22, for to add just one more element requires increasing n_r to 37—a considerable increase in cost for little additional information.

Thus, as much as 44 percent of the variation of 25 elements in the soils of Missouri could be described by collecting 22 random samples from each of the 13 Suborders present in the State.

SECOND-STAGE SAMPLE DESIGNS

The development of an efficient second-stage sample design depends on a balance between field costs and the amount of information to be obtained. The principle cost factors to consider are (1) the analytical and sampling costs, which increase with the number of samples, (2) the travel time and distance between localities, and (3) the difficulty of finding or identifying the mapping unit, especially if the area has never been mapped. These last two factors have major bearing on sampling costs.

As a general principle, the larger the number of localities that have to be visited, the greater the cost for field work. The field costs within a nested design are less for an increase in the number of samples, n_{η} , than for an increase in the number of localities, n_{δ} , because having once arrived at a site the cost of obtaining additional samples is small. Similarly, the cost of increasing n_{δ} is less than that of increasing the number of Series, n_{γ} , because less time and travel are required for sampling a few small areas intensively than for collecting the same number of samples from more widely separated localities. Finally, on a per-unit basis, the costs of most types of

chemical analysis are less than the field costs for the same samples. Thus, although a simple random-sample design always requires the smallest total number of samples, the increased amount of travel required results in a total cost greater than that for most nested designs.

Equation 8 shows how the distribution of variance components determines the levels at which second-stage sampling should be carried out. As a first principle, a larger number of samples should be collected at those levels with large variance components, and a fewer number at those levels with small variance components. However, a second principle evident from equation 8 is that an increase in n_β is more effective in reducing E_s than an increase of any of the n 's for levels within Subgroups. Similarly, an increase in n_γ is more effective than an increase of any of the n 's for levels within Series. The procedure for adjusting n 's to achieve a reduction in the estimate of the variance of the mean according to equation 8 is just the opposite of the adjustment for minimizing field expenses. Therefore, the most efficient sample design requires a balance between information gained and cost. Some examples will illustrate how different distributions might be handled.

If S_β^2 is large compared to the remaining components within Suborders, then the Subgroup mapping units are considered to be homogeneous within themselves and relatively distinct from each other. Equation 8 shows that an increase of n_β will be the most effective means of reducing E_s . Thus a nested design is appropriate, such as $(n_\beta, n_\gamma, n_\delta, n_\eta, n_\epsilon)^1$ where n_β has a large value, and $n_\gamma, n_\delta, n_\eta,$ and n_ϵ , are each equal to a small value, respectively. None of the elements in this study have such a distribution.

If S_γ^2 is much larger than the components within Series, then the Series mapping units are homogeneous within themselves and are distinct from one another. In this case, it would be inefficient to stratify the sampling by either Suborders or Subgroups unless the effect of the large variance at the Series level could be overcome. The elements boron, scandium, sodium, and titanium are examples of this case.

The variance-component distribution of boron well illustrates how a second-stage sampling design is developed for a specific element because boron is a somewhat unusual case. The Suborder variance component is very small and the components at both the Series and analytical levels are large.

Because the variance component at the Suborder level for boron is so small, it would be impractical to map the

distribution of boron at the Suborder level. Rather, a map of Subgroups could be developed with provision for the large variance at both the Series level and analytical level. The first step is to calculate v using a modification of equation 4 as follows:

$$v = \frac{N_\nu}{D_\nu} = \frac{S_\alpha^2 + S_\beta^2}{S_\gamma^2 + S_\delta^2 + S_\eta^2 + S_\epsilon^2} = \frac{0+.00226}{.00764+0+.00539+.01104} = .094 \quad (10)$$

A value of $n_r=32$ is obtained from Miesch (1976, fig. 2) corresponding to v . Then E_r is calculated according to a modification of equation 6:

$$E_r = \frac{S_\gamma^2 S_\delta^2 + S_\eta^2 + S_\epsilon^2}{n_r} = \frac{.02407}{32} = .00075 \quad (11)$$

Substituting the appropriate n 's in a revised version of equation 8,

$$E_s = (1-f_\gamma) \frac{S_\gamma^2}{n_\gamma} + \frac{S_\delta^2}{n_\gamma n_\delta} + \frac{S_\eta^2}{n_\gamma n_\delta n_\eta} + \frac{S_\epsilon^2}{n_\gamma n_\delta n_\eta n_\epsilon} \quad (12)$$

minimizes E_s to fulfill the inequality of 9. Thus, several nested sampling designs for mapping boron at the Subgroup level are as follows:

$n_\gamma n_\delta n_\eta n_\epsilon$	E_s	Total number of analyses per Subgroup	Total number of localities per Subgroup
5 2 3 5	0.00069	150	10
5 3 3 3	.00064	135	15
5 2 3 3	.00074	90	10
6 2 2 2	.00064	48	12

In each case E_s is less than the limit of E_r equal to 0.00075, but some designs are more efficient than others. For example, the design 5, 3, 3, 3 results in an estimate of E_s equal to 0.00064, based on a total of 135 analyses and travel to 15 (n_γ times n_δ) localities within each Subgroup. The design 6, 2, 2, 2 gives the same estimate of E_s but specifies only 48 analyses and travel to 12 localities. Based on an average of five Subgroups per Suborder, the latter sample design indicates that the total number of samples needed to map the varying concentrations of boron between Subgroups is 1,560 (5 Subgroups times 13 Suborders times 24 samples) collected from 780 localities and 3,120 analyses. We observed earlier that a sample design directed at the Subgroup level would be wasteful because of the very small variance at that level. Keeping this caution in mind, what is the consequence of sampling boron at the Subgroup level? We find that at this level we can expect to describe only about 9 percent of the total variation in boron (table 4).

By a somewhat greater investment in sampling and analyses, about 38 percent of the variation in boron (table 4) could be described by mapping concentrations at the Series level. Using a similar procedure with appropriate modifications of equations 10, 11, and 12 the value, v , for boron at the Series level equals 0.603, n_r equals 7, and E_r equals 0.00234. An efficient nested sampling design of

¹The notation describes a nested sampling design in which the value of n_β is the number of Subgroups randomly selected within each Suborder, n_γ is the number of Series randomly selected within each Subgroup, n_δ is the number of localities randomly selected within each Series, n_η is the number of samples randomly selected within each locality, and n_ϵ is the number of independent analyses of each sample.

$n_\delta, n_\eta, n_\epsilon$ equal to 5, 2, 1 estimates E_s as 0.00218. Collecting two samples at each of five localities within each of 275 Series within the State gives a total of 2,750 samples and analyses.

Several nested designs for second-stage sampling, which differ depending on the element, have been considered. It is impractical to have a separate sampling design for each element and it becomes necessary to select a single design which would serve best for the greatest number of elements. As noted previously, the point of diminishing returns occurs where n_γ equals 22; this sampling intensity will provide data adequate to describe 25 elements at the Suborder level. The elements carbon, cobalt, and lead require values of n_γ of 20, 21, and 22, respectively. Nested designs for second-stage sampling within Suborders for these elements are as follows (where the limiting values of E_γ for carbon, cobalt, and lead are 0.00184, 0.00210, and 0.00064, respectively):

Element	n_β	n_γ	n_δ	n_η	n_ϵ	E_s	Total number of analyses per Subgroup	Total number of localities per Subgroup
C.....	5	1	3	1	2	0.00180	30	15
	3	3	4	1	1	.00174	36	36
Co.....	2	4	6	1	1	.00203	48	48
	3	3	4	1	1	.00202	36	36
Pb.....	2	4	4	1	1	.00060	32	32
	3	3	4	1	1	.00052	36	36

The sample design 3, 3, 4, 1, 1 is not necessarily the most efficient for any one of the three elements taken separately, but it is a design that would comfortably accommodate all of them without excessively oversampling any one. This design would also serve adequately for the remainder of the 25 elements, although the degree of oversampling will increase as the values of n_γ decrease for these elements. The total number of samples necessary to describe Suborders is 468 (36 times 13 Suborders).

A sampling design for the case in which S_γ^2 and S_δ^2 are both large compared with all other components is illustrated by manganese. It is assumed that we wish to describe Suborders. Using equations 4, 6, and 8, the necessary calculations give values as follows: v equals 0.08, n_γ equals 37, E_γ equals 0.00209. Several sampling designs are as follows:

n_β	n_γ	n_δ	n_η	n_ϵ	E_s	Total number of analyses per Subgroup	Total number of localities per Subgroup
1	7	10	1	1	0.00388	70	70
1	7	20	1	1	.00372	140	140
1	7	30	1	1	.00342	210	210
2	7	4	1	1	.00202	56	56
3	4	3	1	2	.00199	72	36

The variance component, S_β^2 , is equal to zero or is very small, and accordingly, in the first 3 designs, n_β is set equal to one as a first approximation and n_γ is set equal to the maximum possible value of seven, which is the population size (see table 3). Although none of these first 3 designs reduces E_s sufficiently, they do illustrate the effect of increasing n_δ : namely, there is a large increase in the total number of analyses and localities but little

reduction in E_s . Alternatively, an increase in n_β from one to two has a dramatic effect on the overall efficiency by reducing E_s to a value less than E_γ using 75 percent fewer analyses. The efficiency is further improved by increasing n_β to three and reducing travel requirements to 36 localities. Using the latter sample design, the distribution of manganese among Suborders could be described with 936 (72×13) analyses of samples collected throughout the State from 468 (36×13) localities. However, only 8 percent of the total variation in manganese (table 4) would be accounted for. If a higher minimum percentage of variance is specified, say 40 percent, then more intensive sampling sufficient to the 275 Series is required. This could be done using a 5, 2, 1 nested design by collecting 2,750 samples from 1,375 localities and analyzing each sample once.

GEOCHEMICAL SUMMARIES

Summary statistics based on the first-stage results are shown in table 6. The means are calculated using the maximum likelihood techniques of Cohen (1959, 1961) (compare with Miesch, 1967b) for those elements that are censored by the lower limit of determination. All of the means, except those for potassium, sodium, and silicon, are geometric. The variability is measured by the geometric deviation, which is the antilogarithm of the standard deviation of the logarithms. The statistics for potassium, sodium, and silicon are expressed as the arithmetic mean and standard deviation.

The elements are arranged in table 6 according to increasing value of the variance mean ratio, v_m (Miesch, 1976), as follows:

$$v_m = \frac{S_\alpha^2}{E_s}$$

where E_s is computed for the data on hand. The value of v_m serves as an index of the relative stability² of the observed differences between the Suborder means. For small values of v_m , the observed differences between the Suborder means are unstable. The estimate of the grand mean based on all of the data is a more reliable estimate for any given Suborder than is any individual estimate. As v_m increases, the observed differences among the individual means tend to be ever more stable, and the superiority of the grand mean as an estimate of any Suborder mean diminishes. A value of v_m equal to 1.0 is approximately equivalent to the F -test criterion (equation 5) at the 80-percent probability level.

The analysis of variance shows that, at least between the two extreme Suborder means, 23 of the 31 elements studied, including all of the elements in table 6 with v_m equal to or greater than 1.8, exhibit significant differences

²A stable estimate is one that tends to be reproducible even if new samples are collected and analyzed.

TABLE 6.—Comparison of average compositions of Missouri soil Suborder samples

[Values are in parts per million, except as noted in percent. Relative stability of estimates is given according to ascending order of v_m . The first number is the geometric mean, except where arithmetic mean noted, arranged in ascending order across the row. The second number is the geometric deviation, except where standard deviation noted. Suborder name appears below. The underscore designates groups of Suborders which are indistinguishable according to Duncan's multiple range test (Duncan, 1955). Asterisk (*) indicates elements that are significant according to *F*-test in analysis of variance (Snedecor and Cochran, 1967, p. 265)]

Element	Suborder means and deviations									Grand mean and deviation	v_m
B.....	31 $\frac{\%}{\pm}$ 1.62 Ochrept	32 $\frac{\%}{\pm}$ 1.48 Udult	35 $\frac{\%}{\pm}$ 1.55 Udalf	37 $\frac{\%}{\pm}$ 1.37 Fluvent	40 $\frac{\%}{\pm}$ 1.39 Aquept	41 $\frac{\%}{\pm}$ 1.33 Aquoll	41 $\frac{\%}{\pm}$ 1.33 Udoll	43 $\frac{\%}{\pm}$ 1.34 Aqualf	44 $\frac{\%}{\pm}$ 1.26 Alboll	38 $\frac{\%}{\pm}$ 1.45	0
Zr.....	220 $\frac{\%}{\pm}$ 1.86 Aquoll	250 $\frac{\%}{\pm}$ 1.72 Udoll	270 $\frac{\%}{\pm}$ 1.84 Ochrept	290 $\frac{\%}{\pm}$ 1.53 Udult	290 $\frac{\%}{\pm}$ 1.62 Aquept	300 $\frac{\%}{\pm}$ 1.35 Alboll	360 $\frac{\%}{\pm}$ 1.73 Fluvent	400 $\frac{\%}{\pm}$ 1.61 Udalf	480 $\frac{\%}{\pm}$ 1.49 Aqualf	310 $\frac{\%}{\pm}$ 1.75	0.39
Mn.....	520 $\frac{\%}{\pm}$ 1.78 Fluvent	610 $\frac{\%}{\pm}$ 2.49 Udult	640 $\frac{\%}{\pm}$ 1.62 Udoll	670 $\frac{\%}{\pm}$ 1.77 Alboll	680 $\frac{\%}{\pm}$ 1.38 Aquoll	700 $\frac{\%}{\pm}$ 1.77 Aquept	830 $\frac{\%}{\pm}$ 1.42 Ochrept	880 $\frac{\%}{\pm}$ 2.26 Udalf	970 $\frac{\%}{\pm}$ 1.59 Aqualf	710 $\frac{\%}{\pm}$ 1.95	.78
Ti.....	0.31 $\frac{\%}{\pm}$ 1.51 Fluvent	0.33 $\frac{\%}{\pm}$ 1.84 Ochrept	0.33 $\frac{\%}{\pm}$ 1.45 Udult	0.40 $\frac{\%}{\pm}$ 1.36 Udalf	0.40 $\frac{\%}{\pm}$ 1.37 Udoll	0.41 $\frac{\%}{\pm}$ 1.35 Aquoll	0.43 $\frac{\%}{\pm}$ 1.31 Aquept	0.48 $\frac{\%}{\pm}$ 1.16 Alboll	0.49 $\frac{\%}{\pm}$ 1.32 Aqualf	0.39 $\frac{\%}{\pm}$ 1.46	.83
Y.....	18 $\frac{\%}{\pm}$ 1.61 Ochrept	20 $\frac{\%}{\pm}$ 1.60 Udult	24 $\frac{\%}{\pm}$ 1.23 Fluvent	25 $\frac{\%}{\pm}$ 1.40 Udalf	26 $\frac{\%}{\pm}$ 1.22 Alboll	27 $\frac{\%}{\pm}$ 1.21 Aquoll	27 $\frac{\%}{\pm}$ 1.20 Aqualf	27 $\frac{\%}{\pm}$ 1.29 Udoll	28 $\frac{\%}{\pm}$ 1.18 Aquept	25 $\frac{\%}{\pm}$ 1.37	1.33
Co.....	6.9 $\frac{\%}{\pm}$ 1.98 Udult	7.6 $\frac{\%}{\pm}$ 1.46 Ochrept	8.0 $\frac{\%}{\pm}$ 1.44 Alboll	8.8 $\frac{\%}{\pm}$ 1.35 Fluvent	8.8 $\frac{\%}{\pm}$ 1.32 Aquept	10 $\frac{\%}{\pm}$ 1.54 Udoll	11 $\frac{\%}{\pm}$ 1.54 Aqualf	12 $\frac{\%}{\pm}$ 1.86 Udalf	12 $\frac{\%}{\pm}$ 1.32 Aquoll	9.6 $\frac{\%}{\pm}$ 1.68	1.46
Ni.....	10 $\frac{\%}{\pm}$ 1.88 Udult	14 $\frac{\%}{\pm}$ 1.64 Ochrept	15 $\frac{\%}{\pm}$ 2.10 Udalf	15 $\frac{\%}{\pm}$ 1.84 Fluvent	17 $\frac{\%}{\pm}$ 1.59 Aqualf	18 $\frac{\%}{\pm}$ 1.42 Alboll	22 $\frac{\%}{\pm}$ 1.34 Aquept	23 $\frac{\%}{\pm}$ 1.61 Udoll	29 $\frac{\%}{\pm}$ 1.44 Aquoll	18 $\frac{\%}{\pm}$ 1.80	1.66
C.....	.87 $\frac{\%}{\pm}$ 1.81 Fluvent	.95 $\frac{\%}{\pm}$ 1.53 Udalf	1.3 $\frac{\%}{\pm}$ 1.63 Udult	1.3 $\frac{\%}{\pm}$ 1.80 Aquept	1.3 $\frac{\%}{\pm}$ 1.45 Alboll	1.3 $\frac{\%}{\pm}$ 1.41 Aquoll	1.6 $\frac{\%}{\pm}$ 1.23 Aqualf	1.7 $\frac{\%}{\pm}$ 1.42 Udoll	1.9 $\frac{\%}{\pm}$ 1.31 Udult	1.3 $\frac{\%}{\pm}$ 1.61	1.67
Hg.....	.039 $\frac{\%}{\pm}$ 2.14 Fluvent	.040 $\frac{\%}{\pm}$ 1.60 Ochrept	.042 $\frac{\%}{\pm}$ 1.84 Udult	.048 $\frac{\%}{\pm}$ 1.29 Alboll	.050 $\frac{\%}{\pm}$ 2.09 Aqualf	.050 $\frac{\%}{\pm}$ 1.32 Udoll	.058 $\frac{\%}{\pm}$ 1.53 Aquoll	.067 $\frac{\%}{\pm}$ 2.25 Udalf	.077 $\frac{\%}{\pm}$ 2.28 Aquept	052 $\frac{\%}{\pm}$ 2.02	1.70
Pb*.....	15 $\frac{\%}{\pm}$ 1.45 Udult	17 $\frac{\%}{\pm}$ 1.20 Fluvent	18 $\frac{\%}{\pm}$ 1.27 Ochrept	20 $\frac{\%}{\pm}$ 1.24 Aquept	20 $\frac{\%}{\pm}$ 1.24 Udoll	20 $\frac{\%}{\pm}$ 1.32 Udalf	21 $\frac{\%}{\pm}$ 1.40 Alboll	22 $\frac{\%}{\pm}$ 1.30 Aquoll	23 $\frac{\%}{\pm}$ 1.31 Aqualf	19 $\frac{\%}{\pm}$ 1.34	1.87
Yb*.....	2.0 $\frac{\%}{\pm}$ 1.51 Ochrept	2.1 $\frac{\%}{\pm}$ 1.38 Udult	2.5 $\frac{\%}{\pm}$ 1.34 Fluvent	2.6 $\frac{\%}{\pm}$ 1.27 Udalf	2.8 $\frac{\%}{\pm}$ 1.30 Udoll	2.8 $\frac{\%}{\pm}$ 1.17 Aquoll	2.9 $\frac{\%}{\pm}$ 1.15 Aquept	2.9 $\frac{\%}{\pm}$ 1.20 Aqualf	3.0 $\frac{\%}{\pm}$ 1.00 Alboll	2.6 $\frac{\%}{\pm}$ 1.32	2.06
Fe* 1.....	1.6 $\frac{\%}{\pm}$ 1.54 Ochrept	1.9 $\frac{\%}{\pm}$ 1.30 Fluvent	2.1 $\frac{\%}{\pm}$ 1.53 Udult	2.1 $\frac{\%}{\pm}$ 1.28 Aqualf	2.2 $\frac{\%}{\pm}$ 1.86 Udalf	2.3 $\frac{\%}{\pm}$ 1.17 Alboll	2.5 $\frac{\%}{\pm}$ 1.38 Aquept	3.0 $\frac{\%}{\pm}$ 1.32 Udoll	3.6 $\frac{\%}{\pm}$ 1.37 Aquoll	2.3 $\frac{\%}{\pm}$ 1.55	2.13
As*.....	4.9 $\frac{\%}{\pm}$ 1.54 Ochrept	6.2 $\frac{\%}{\pm}$ 1.22 Aquept	6.5 $\frac{\%}{\pm}$ 1.21 Fluvent	6.9 $\frac{\%}{\pm}$ 1.38 Udalf	7.2 $\frac{\%}{\pm}$ 1.53 Udult	7.7 $\frac{\%}{\pm}$ 1.25 Alboll	7.7 $\frac{\%}{\pm}$ 1.47 Aqualf	9.0 $\frac{\%}{\pm}$ 1.32 Udoll	9.5 $\frac{\%}{\pm}$ 1.26 Aquoll	7.2 $\frac{\%}{\pm}$ 1.44	2.32
La*.....	34 $\frac{\%}{\pm}$ 1.65 Udult	44 $\frac{\%}{\pm}$ 1.46 Fluvent	45 $\frac{\%}{\pm}$ 1.76 Ochrept	46 $\frac{\%}{\pm}$ 1.44 Udalf	56 $\frac{\%}{\pm}$ 1.18 Alboll	57 $\frac{\%}{\pm}$ 1.24 Aqualf	58 $\frac{\%}{\pm}$ 1.28 Udoll	58 $\frac{\%}{\pm}$ 1.19 Aquoll	60 $\frac{\%}{\pm}$ 1.24 Aquept	51 $\frac{\%}{\pm}$ 1.59	2.37
Na* 1 2.....	.19 \pm .094 Udult	.45 \pm .22 Udalf	.47 \pm .30 Ochrept	.54 \pm .20 Aquoll	.63 \pm .23 Aquept	.68 \pm .20 Udalf	.73 \pm .41 Alboll	.76 \pm .89 Udoll	.78 \pm .12 Aqualf	.59 \pm .33	2.55
Ca* 1.....	.12 $\frac{\%}{\pm}$ 1.78 Udult	.23 $\frac{\%}{\pm}$ 2.20 Udalf	.27 $\frac{\%}{\pm}$ 2.08 Ochrept	.34 $\frac{\%}{\pm}$ 1.41 Aqualf	.39 $\frac{\%}{\pm}$ 3.20 Fluvent	.45 $\frac{\%}{\pm}$ 1.17 Udalf	.49 $\frac{\%}{\pm}$ 1.97 Aquept	.72 $\frac{\%}{\pm}$ 1.52 Udoll	.77 $\frac{\%}{\pm}$ 2.05 Aquoll	.46 $\frac{\%}{\pm}$ 2.65	3.00
Cr*.....	38 $\frac{\%}{\pm}$ 1.60 Udult	41 $\frac{\%}{\pm}$ 1.93 Ochrept	46 $\frac{\%}{\pm}$ 1.59 Fluvent	61 $\frac{\%}{\pm}$ 1.35 Udalf	61 $\frac{\%}{\pm}$ 1.24 Aqualf	66 $\frac{\%}{\pm}$ 1.14 Alboll	67 $\frac{\%}{\pm}$ 1.29 Udoll	71 $\frac{\%}{\pm}$ 1.27 Aquept	83 $\frac{\%}{\pm}$ 1.29 Aquoll	59 $\frac{\%}{\pm}$ 1.55	3.09
K* 1 2.....	.89 \pm .48 Udult	1.2 \pm .48 Ochrept	1.2 \pm .33 Udalf	1.5 \pm .55 Fluvent	1.5 \pm .30 Aqualf	1.6 \pm .27 Aquept	1.6 \pm 1.60 Alboll	1.8 \pm .29 Udoll	1.9 \pm .46 Aquoll	1.5 \pm .54	3.26
Cu*.....	9.4 $\frac{\%}{\pm}$ 1.61 Udult	11 $\frac{\%}{\pm}$ 1.87 Ochrept	16 $\frac{\%}{\pm}$ 1.44 Fluvent	16 $\frac{\%}{\pm}$ 1.65 Udalf	17 $\frac{\%}{\pm}$ 1.36 Aqualf	20 $\frac{\%}{\pm}$ 1.23 Alboll	26 $\frac{\%}{\pm}$ 1.53 Aquept	28 $\frac{\%}{\pm}$ 1.54 Udoll	38 $\frac{\%}{\pm}$ 1.69 Aquoll	19 $\frac{\%}{\pm}$ 1.84	4.10
F*.....	100 $\frac{\%}{\pm}$ 2.47 Udult	130 $\frac{\%}{\pm}$ 2.36 Ochrept	200 $\frac{\%}{\pm}$ 1.66 Udalf	230 $\frac{\%}{\pm}$ 1.73 Fluvent	250 $\frac{\%}{\pm}$ 1.58 Aqualf	260 $\frac{\%}{\pm}$ 1.24 Alboll	300 $\frac{\%}{\pm}$ 1.66 Aquept	400 $\frac{\%}{\pm}$ 1.44 Udoll	450 $\frac{\%}{\pm}$ 1.65 Aquoll	260 $\frac{\%}{\pm}$ 2.50	4.33
Ba*.....	430 $\frac{\%}{\pm}$ 1.53 Udult	630 $\frac{\%}{\pm}$ 1.50 Udalf	720 $\frac{\%}{\pm}$ 1.49 Ochrept	720 $\frac{\%}{\pm}$ 1.65 Fluvent	870 $\frac{\%}{\pm}$ 1.35 Aqualf	970 $\frac{\%}{\pm}$ 1.36 Aquoll	980 $\frac{\%}{\pm}$ 1.33 Aquept	1100 $\frac{\%}{\pm}$ 1.32 Alboll	1100 $\frac{\%}{\pm}$ 1.22 Udoll	810 $\frac{\%}{\pm}$ 1.62	4.43
Sr*.....	52 $\frac{\%}{\pm}$ 1.69 Udult	100 $\frac{\%}{\pm}$ 1.78 Udalf	100 $\frac{\%}{\pm}$ 1.84 Ochrept	120 $\frac{\%}{\pm}$ 2.26 Fluvent	150 $\frac{\%}{\pm}$ 1.33 Aqualf	170 $\frac{\%}{\pm}$ 1.48 Aquept	180 $\frac{\%}{\pm}$ 1.15 Alboll	200 $\frac{\%}{\pm}$ 1.28 Udoll	220 $\frac{\%}{\pm}$ 1.28 Aquoll	140 $\frac{\%}{\pm}$ 1.99	4.44
Se*.....	.11 $\frac{\%}{\pm}$ 2.96 Ochrept	.19 $\frac{\%}{\pm}$ 2.36 Fluvent	.20 $\frac{\%}{\pm}$ 2.33 Udult	.35 $\frac{\%}{\pm}$ 1.92 Udalf	.40 $\frac{\%}{\pm}$ 1.53 Aqualf	.40 $\frac{\%}{\pm}$ 2.05 Aquept	.49 $\frac{\%}{\pm}$ 1.47 Udoll	.51 $\frac{\%}{\pm}$ 1.71 Aquoll	.55 $\frac{\%}{\pm}$ 1.18 Alboll	.33 $\frac{\%}{\pm}$ 3.13	4.75
Sc*.....	4.1 $\frac{\%}{\pm}$ 1.75 Udult	4.7 $\frac{\%}{\pm}$ 2.27 Ochrept	6.6 $\frac{\%}{\pm}$ 1.53 Fluvent	7.6 $\frac{\%}{\pm}$ 1.48 Udalf	8.6 $\frac{\%}{\pm}$ 1.24 Aqualf	11 $\frac{\%}{\pm}$ 1.28 Alboll	11 $\frac{\%}{\pm}$ 1.36 Aquept	12 $\frac{\%}{\pm}$ 1.32 Udoll	13 $\frac{\%}{\pm}$ 1.25 Aquoll	8.6 $\frac{\%}{\pm}$ 1.78	4.86
P* 1.....	.019 $\frac{\%}{\pm}$ 1.63 Udult	.029 $\frac{\%}{\pm}$ 1.60 Udalf	.031 $\frac{\%}{\pm}$ 1.62 Ochrept	.032 $\frac{\%}{\pm}$ 1.38 Aqualf	.032 $\frac{\%}{\pm}$ 1.24 Alboll	.036 $\frac{\%}{\pm}$ 1.35 Aquept	.043 $\frac{\%}{\pm}$ 1.58 Fluvent	.046 $\frac{\%}{\pm}$ 1.45 Udoll	.047 $\frac{\%}{\pm}$ 1.25 Aquoll	.035 $\frac{\%}{\pm}$ 1.69	4.90

TABLE 6.—Comparison of average compositions of Missouri soil Suborder samples—Continued

Element	Suborder means and deviations										Grand mean and deviation	v_m
	Aquoll	Udoll	Aquept	Alboll	Udalf	Aqualf	Fluvent	Ochrept	Udult			
Si* ^{1 2}	33 ± 3.92	34 ± 2.29	36 ± 2.98	36 ± 1.33	38 ± 3.12	38 ± 1.56	39 ± 2.52	40 ± 3.62	42 ± 2.64		37 ± 3.83	5.02
Mg*	.12 ± 2.10	.16 ± 1.84	.22 ± 1.50	.29 ± 1.49	.35 ± 2.71**	.37 ± 1.41	.46 ± 1.90	.58 ± 1.70	.63 ± 1.89		.33 ± 1.31	5.05
Zn*	24 ± 1.51	30 ± 1.65	41 ± 1.79	49 ± 1.45	51 ± 1.47	56 ± 1.15	64 ± 1.36	69 ± 1.46	90 ± 1.46		51 ± 1.75	5.05
V*	45 ± 1.55	49 ± 1.90	53 ± 1.72	61 ± 1.51	75 ± 1.31	89 ± 1.27	93 ± 1.50	110 ± 1.44	140 ± 1.56		76 ± 1.78	5.97
Al* ¹	2.2 ± 1.56	3.0 ± 1.56	3.8 ± 1.54	3.8 ± 1.37	4.3 ± 1.21	5.0 ± 1.08	5.3 ± 1.32	5.9 ± 1.20	6.4 ± 1.39		4.3 ± 1.59	6.06
Ga*	5.6 ± 1.76	8.4 ± 1.90	9.0 ± 1.67	10 ± 1.54	13 ± 1.25	15 ± 1.09	15 ± 1.49	17 ± 1.24	21 ± 1.54		12 ± 1.83	6.33

¹Values given in percent.

²Arithmetic mean and standard deviation given.

at the 0.05 probability level. The first-stage sample data are judged, by the criterion of expression 9, to be adequate for describing the differences between Suborder means for 13 elements. (See table 5.) These are the same 13 elements that appear in table 6 with v_m equal to or greater than 4.0. The considerable overlap between the groups for most of the elements suggests that the only individual Suborders that can be demonstrated to be significantly different are those at the opposite ends of the range. For elements such as boron and zirconium none of the means are significantly different, and, accordingly, the Duncan test shows that the Suborders are essentially the same with respect to these elements.

The grand mean concentrations of most trace elements in soils, as reported in table 6, are about equal to or slightly higher than the means reported by Shacklette and others (1971) for U.S. soils located east of the 97th meridian. In Missouri soils the mean concentrations of major elements such as aluminum, calcium, iron, magnesium, phosphorus, potassium, sodium, and titanium are 1.25 to 2 times higher than in the soils of the eastern United States. This may be due in part to the fact that Missouri is covered to a large extent by loess and glacial material, which tend to be relatively more enriched in a number of elements than are the more highly weathered parent materials of the southeastern United States. The grand means in table 6 are nearly equal to corresponding means based on 1,140 soil samples collected extensively throughout Missouri in a later study (U.S. Geol. Survey, 1972e, p. 41-55; 1972f, p. 19-57; 1973, p. 16-25).

DISCUSSION

The soil classification scheme is a useful tool for reducing a great complexity of information about soils down to several more manageable levels of generalization (Smith, 1963, p. 7). The classification scheme emphasizes

relationships between the soils and their environmental settings. The taxa are defined, especially at the higher levels, so as to group soils of similar genesis (Smith, 1963, p. 8); however, the criteria used for definition are not the factors of soil formation (climate, topography, parent material, and organisms, each operating through time) (Jenny, 1941) that prevail, but rather the soil characteristics that result from any combination or intensity of these formational factors. Thus, chemical composition of the soil is a characteristic that could be reflected in the taxa, although currently it is not. Moreover, translation of knowledge of the geographic distribution of soils onto maps is a practical necessity for both the sampler and those who will interpret the sampling results for environmental studies. The geographic distribution of soils serves as a bridge between our knowledge about the chemical composition of soils and the physical environment of the Missouri landscape. Unfortunately, coverage of the State by detailed mapping according to modern concepts is incomplete.

Although we would hope that a relatively large part of the total variance would occur at the higher taxonomic levels so that the soils could be more easily sampled and described, the results show that the largest part of the variation occurs at the lower levels. This is perhaps not so surprising if we consider first that the variance is related to distance between sampling localities, in that samples close together tend to be more similar than those farther apart. The soil Series as a mapping unit has a minimum area of five acres; thus the distance between Series mapping units is more than that between localities within a Series but less than the distance between more regional mapping units such as Suborders. A second consideration is that, inasmuch as the taxa at the higher levels of the classification system are defined by their most general properties and the taxa at successively lower levels

are defined by ever more detailed properties, quite a diversity of Series may be permitted within a Suborder. The variance component, S_{λ}^2 , is characteristically large, and sampling only a few of the Series at random within each Suborder (as is implied by concentrating the sampling at the Suborder level) may fail to describe the Suborder adequately. Thus, while a sample of the higher taxa improves the level of generalization, it may describe only a small percentage of the total mappable variance.

The alternative is to sample at the Series level, which means that every Series would have to be sampled. The results show that up to about 85 percent of the variance for some elements can be described by sampling all of the soil Series. However, Smith (1963, p. 7) observed that the broad view of soils (for example, on a Statewide scale) cannot be grasped at the Series level because there are too many categories. If Series were used to map the distribution of elements in the soils of a large region, the very number of detailed mapping units could overwhelm any reasonable effort to define general principles; but a more immediate drawback is that the sampling effort required is quite impractical.

What is the consequence of describing only a small part of the total variance? It provides only a limited ability to understand or to predict what mechanisms or factors are controlling the distribution of elements in the Suborders. In other words, the mechanisms that control the "marker" properties of the Suborders are apparently only distantly related to the mechanisms that control the distribution of elements. The only remedy to this situation is to search for alternative classifications based on a small number of chemically distinct classes.

Twenty-three elements show significant differences between Suborder means (table 6). With the exception of silicon, the highest concentrations of these elements are generally found in the Suborders of the Mollisols (Alboll, Aquoll, and Udoll). The Mollisols have developed in the temperate climatic conditions of the central Great Plains, have thick, nearly black surface horizons, and are high in exchangeable bases, such as calcium, magnesium, potassium, and sodium. The Aquolls in particular have the highest concentrations of all Suborders. These are the seasonally wet members of the Mollisols that typically occur on alluvium in the Missouri River valley (U.S. Geological Survey, 1970, p. 86).

Conversely, the Udults (a Suborder of the Ultisols), with the Ochrepts (a Suborder of the Inceptisols), and the Fluvents (a Suborder of the Entisols), generally contain the lowest concentrations of these elements. The Udults are older soils than the Aquolls, with more intensely weathered profiles, and probably developed under a warmer, more humid climate. The Udults of Missouri are a northern extent of the Ultisols, which cover most of the southeastern United States. The Ultisols are depleted in bases and organic matter, have a thick accumulation of clay

with a minimum of weatherable minerals, and are developed on sedimentary parent materials of Pleistocene or Holocene age. Most of the soils having low elemental concentrations tend to be more base-rich—implying less leaching—than the Udults; although in the case of the Ochrepts, considerable variation in base saturation is permitted by the definition. The amount of organic matter in the Ochrepts and Fluvents also is higher than in the Udults, but as a whole these soils contain considerably less organic matter than the Mollisols.

CONCLUSIONS

Examination using analysis-of-variance techniques of the geochemical properties of the soils classified by the Soil Taxonomy of the National Cooperative Soil Survey suggests the following.

1. Although significant differences are found between the Suborders for many of the elements examined, these differences account for only a limited percentage of the total variation in elemental abundances among the soils of the State. The differences for 13 of the elements are adequately described by the results of the first-stage sampling; the differences for 9 additional elements could be described by implementation of the recommended second-stage sampling. Such multistage sampling is necessary because the list of taxonomic criteria by which the Suborders are defined does not include chemical composition.
2. Geochemical differences between soil Groups and Subgroups within the same Suborder are negligibly small for most elements studied.
3. A large percentage of the total variation in the soils of the State could be described by sampling each soil Series, but the number of Series in the State and the chemical variation within each Series would result in a very large sampling and analytical load.
4. Geographic variation within soil Series is, on the average, about half of that observed between Series.
5. The estimates of the grand mean for each of the elements do not differ significantly from estimates based on a much more extensive sample of Missouri soils.
6. The concentrations of numerous elements are two or more times greater in Suborders of the Mollisols than in Suborders of the Ultisols.
7. An alternate classification scheme based on a small number of chemically distinct classes would be helpful in minimizing the second-stage sampling required to produce geochemical maps.

REFERENCES CITED

- Anderson, R. L., 1965, Negative variance estimates—Answer to query 9: *Technometrics*, v. 7, no. 1, p. 75-76.
- Anderson, R. L., and Bancroft, T. A., 1952, *Statistical theory in research*: New York, McGraw-Hill, 399 p.
- Bartel, A. J., and Millard, H. T., Jr., 1976, Analysis of rocks and soils for iodine by radio-isotope dilution-neutron activation, *in* Miesch,

- A. T., Geochemical survey of Missouri—Methods of sampling, laboratory analysis, and statistical reduction of data: U.S. Geol. Survey Prof. Paper 954-A, p. A15-A17.
- Beckett, P. H. T., and Webster, R., 1971, Soil variability—A review: *Soils and Fertilizers*, v. 34, no. 1, p. 1-15.
- Bennett, C. A., and Franklin, N. L., 1954, *Statistical analysis in chemistry and the chemical industry*: New York, John Wiley, 724 p.
- Brown, B. L., 1968, Soil survey of Worth County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], 50 p.
- 1971, Soil survey of Pemiscot County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], 99 p.
- Cline, M. G., 1963, Logic of the new system of soil classification: *Soil Science*, v. 96, no. 1, p. 17-22.
- Cochran, W. G., 1963, *Sampling techniques* [2d ed.]: New York, John Wiley, 413 p.
- Cohen, A. C., Jr., 1959, Simplified estimators for the normal distribution when samples are singly censored or truncated: *Technometrics*, v. 1, no. 3, p. 217-237.
- 1961, Tables for maximum likelihood estimates—Singly truncated and singly censored samples: *Technometrics*, v. 3, no. 4, p. 535-541.
- Connor, J. J., Feder, G. L., Erdman, J. A., and Tidball, R. R., 1972, Environmental geochemistry in Missouri—A multidisciplinary study: *Internat. Geol. Cong., 24th, Montreal 1972, Symposium 1—Earth sciences and the quality of life*, p. 7-14.
- Duncan, D. B., 1955, Multiple range and multiple F tests: *Biometrics*, v. 11, no. 1, p. 1-42.
- Gilbert, F. L., 1971, Soil survey of Dent County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], 141 p.
- Grogger, H. E., 1964, Soil survey of Daviess County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], ser. 1959, no. 34, 39 p.
- Huffman, Claude, Jr., and Dinnin, J. I., 1976, Analysis of rocks and soils by atomic absorption and other methods, in Miesch, A. T., *Geochemical survey of Missouri—Methods of sampling, laboratory analysis, and statistical reduction of data*: U.S. Geol. Survey Prof. Paper 954-A, p. A12-A14.
- Hughes, H. E., 1974, Soil survey of Barton County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], 76 p.
- Jeffrey, A. H., 1974, Soil survey of Caldwell County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], 56 p.
- 1975, Soil survey of Lafayette County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], 67 p.
- Jenny, Hans, 1941, *Factors of soil formation, a system of quantitative pedology*: New York, McGraw-Hill, 281 p.
- Krumbein, W. C., 1960, The "geological population" as a framework for analyzing numerical data in geology: *Liverpool and Manchester Geol. Jour.*, v. 2, pt. 3, p. 341-368.
- Krumbein, W. C., and Slack, H. A., 1956, Statistical analysis of low-level radioactivity of Pennsylvanian black fissile shale in Illinois: *Geol. Soc. America Bull.*, v. 67, p. 739-761.
- Krusekopf, H. H., 1966, Delta soils of southeast Missouri: *Missouri Univ. Agr. Expt. Sta. Bull.* 854, 38 p.
- Krusekopf, H. H., and Scrivner, C. L., 1962, Soil survey of Boone County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], ser. 1951, no. 12, 39 p.
- Leone, F. C., and Nelson, L. S., 1966, Sampling distributions of variance components—[Pt.] 1, Empirical studies of balanced nested designs: *Technometrics*, v. 8, no. 3, p. 457-468.
- McKenzie, R. M., 1955, Sampling variations in the concentrations of elements in soils: *Australian Jour. Agr. Research*, v. 6, p. 699-706.
- Miesch, A. T., 1967a, Theory of error in geochemical data: U.S. Geol. Survey Prof. Paper 574-A, 17 p.
- 1967b, Methods of computation for estimating geochemical abundance: U.S. Geol. Survey Prof. Paper 574-B, 15 p.
- 1976, Geochemical survey of Missouri—Methods of sampling, laboratory analysis, and statistical reduction of data, *with sections on Laboratory methods*, by Barnett, P. R., Bartel, A. J., Dinnin, J. I., Feder, G. L., Harms, T. F., Huffman, Claude, Jr., Janzer, V. J., Millard, H. T., Jr., Neiman, H. G., Skougstad, M. W., and Wahlberg, J. S.: U.S. Geol. Survey Prof. Paper 954-A, 39 p.
- Missouri University Agricultural Experiment Station, 1931, Reconnaissance survey soil map of Missouri: Columbia, University of Missouri, 1 sheet.
- Neiman, H. G., 1976, Analysis of rocks, soils, and plant ashes by emission spectroscopy, in Miesch, A. T., *Geochemical survey of Missouri—Methods of sampling, laboratory analysis, and statistical reduction of data*: U.S. Geol. Survey Prof. Paper 954-A, p. A14-A15.
- Nelder, J. A., 1954, The interpretation of negative components of variance: *Biometrika*, v. 41, p. 544-548.
- Oertel, A. C., 1959, Estimation of the trace-element status of large areas of soil: *Australian Jour. Agr. Research*, v. 10, p. 58-70.
- Olson, J. S., and Potter, P. E., 1954, Statistical methods, [Pt. 1] of Variance components of cross-bedding direction in some basal Pennsylvanian sandstones of the eastern interior basin: *Jour. Geology*, v. 62, p. 26-49.
- Petersen, R. G., and Calvin, L. D., 1965, Sampling, in Black, C. A., Evans, D. D., Ensminger, L. E., White, J. S., and Clark, F. E., eds., *Methods of soil analysis—Pt. 1, Physical and mineralogical properties including statistics of measurement and sampling*: *Am. Soc. Agronomy, Agronomy Mon.* 9, p. 54-72.
- Potter, P. E., and Olson, J. S., 1954, Geological application, [Pt. 2] of Variance components of cross-bedding direction in some basal Pennsylvanian sandstones of the eastern interior basin: *Jour. Geology*, v. 62, p. 50-73.
- Riecken, F. F., and Smith, G. D., 1949, Lower categories of soil classification—Family, series, type, and phase: *Soil Science*, v. 67, p. 107-115.
- Rourke, J. D., Shrader, W. D., and Atkinson, C. H., 1956, Soil survey of Livingston County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], ser. 1950, no. 1, 54 p.
- Scrivner, C. L., 1961, Soils of Howard County, Missouri: *Missouri Univ. Agr. Expt. Sta. Bull.* 749, 20 p.
- Scrivner, C. L., and Baker, J. C., 1961, Soils of Blackwater and Lamine Townships, Cooper County, Missouri: *Missouri Univ. Agr. Expt. Sta. Bull.* 772, 20 p.
- Scrivner, C. L., Baker, J. C., and Miller, B. J., 1966, Soils of Missouri—A guide to their identification and interpretation: *Missouri Univ. Extension Div. Circ.* 823, 48 p.
- Scrivner, C. L., and Frieze, J. A., 1953, Soils of Taney County, Missouri: *Missouri Univ. Agr. Expt. Sta. Prog. Rept.* 24, 6 p.
- 1964, Soil survey of Moniteau County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], ser. 1953, no. 11, 42 p.
- Shacklette, H. T., Hamilton, J. C., Boerngen, J. G., and Bowles, J. M., 1971, Elemental composition of surficial materials in the conterminous United States: U.S. Geol. Survey Prof. Paper 574-D, 71 p.
- Shrader, W. D., and Krusekopf, H. H., 1956, Soil survey of St. Charles County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], ser. 1939, no. 28, 49 p.
- Shrader, W. D., Springer, M. E., and Hall, N. S., 1953, Soil survey of Holt County, Missouri: U.S. Soil Conserv. Service Soil Survey [Rept.], ser. 1939, no. 21, 61 p.
- Shrader, W. D., Springer, M. E., Hall, N. S., DeYoung, William, and Jordan, H. V., 1945, Soil survey (reconnaissance) of Linn County, Missouri: U.S. Bur. Plant Industry Soils, and Agr. Eng. [Soil Conserv. Service] Soil Survey [Rept.], ser. 1938, no. 7, 28 p.
- Shrader, W. D., Springer, M. E., Hamby, Robert, Pettijohn, W. J., and Miller, J. T., 1954, Soil survey of Jasper County, Missouri:

- U.S. Soil Conserv. Service Soil Survey [Rept.], ser. 1942, no. 5, 67 p.
- Smith, G. D., 1963, Objectives and basic assumptions of the new soil classification system; *Soil Science*, v. 96, no. 1, p. 6-16.
- Snedecor, G. W., and Cochran, W. G., 1967, *Statistical methods* [6th ed.]: Ames, Iowa State Univ. Press, 593 p.
- Thompson, W. A., Jr., 1962, The problem of negative estimates of variance components: *Annals of Math. Statistics*, v. 33, p. 273-289.
- Thompson, W. A., Jr., and Moore, J. R., 1963, Non-negative estimates of variance components: *Technometrics*, v. 5, no. 4, p. 441-449.
- Thorp, James, and Smith, G. D., 1949, Higher categories of soil classification—Order, suborder, and great soil groups: *Soil Science*, v. 67, p. 117-126.
- Tidball, R. R., 1971, Geochemical variation in Missouri soils, in Hemphill, D. D., ed., Trace substances in environmental health—IV [4th Ann. Conf. Proc.]: Columbia, Missouri Univ., p. 15-25.
- U.S. Department of Agriculture, 1938, Soils and men: U.S. Dept. Agriculture, Yearbook of Agriculture 1938, 1232 p.
- U.S. Geological Survey, 1970, National Atlas of the United States of America: Washington, D. C., 417 p.
- 1972a, Geochemical survey of Missouri—Plans and progress for first six-month period (July-December, 1969): U.S. Geol. Survey Open-file Rept., 49 p.
- 1972b, Geochemical survey of Missouri—Plans and progress for second six-month period (January-June, 1970): U.S. Geol. Survey Open-file Rept., 60 p.
- 1972c, Geochemical survey of Missouri—Plans and progress for third six-month period (July-December, 1970): U.S. Geol. Survey Open-file Rept., 33 p.
- 1972d, Geochemical survey of Missouri—Plans and progress for fourth six-month period (January-June, 1971): U.S. Geol. Survey Open-file Rept., 63 p.
- 1972e, Geochemical survey of Missouri—Plans and progress for fifth six-month period (July-December, 1971): U.S. Geol. Survey Open-file Rept., 145 p.
- 1972f, Geochemical survey of Missouri—Plans and progress for sixth six-month period (January-June, 1972): U.S. Geol. Survey Open-file Rept., 86 p.
- 1973, Geochemical survey of Missouri—Plans and progress for seventh six-month period (July-December, 1972): U.S. Geol. Survey Open-file Rept., 59 p.
- U.S. Soil Conservation Service, 1968, List of published soil surveys: Washington, D.C., U.S. Govt. Printing Office, 47 p.
- U.S. Soil Conservation Service, Soil Survey Staff, 1960, Soil classification, a comprehensive system (7th Approximation): Washington, U.S. Govt. Printing Office, 265 p.
- 1967, Supplement to soil classification system (7th approximation): Washington, U.S. Govt. Printing Office, 207 p.
- 1968, Supplement to soil classification system (7th approximation), histosols: Washington, U. S. Govt. Printing Office, 22 p.
- Soil taxonomy of the National Cooperative Soil Survey: Washington, U.S. Govt. Printing Office, (in press).
- Vazhenin, I. G., Muzychkin, Ye. T., Prokhorova, Z. A., and Aleshina, T. N., 1961, Method of compiling large-scale soil agricultural chemical maps for fertilizer application: *Soviet Soil Science* 1961, no. 4, p. 363-374.
- Wahlberg, J. S., 1976, Analysis of rocks and soils by X-ray fluorescence, in Miesch, A. T., Geochemical survey of Missouri—Methods of sampling, laboratory analysis, and statistical reduction of data: U.S. Geol. Survey Prof. Paper 954-A. p. A11-A12.
- Wilding, L. P., Jones, R. B., and Schafer, G. M., 1965, Variation of soil morphological properties within Miami, Celina, and Crosby mapping units in west-central Ohio: *Soil Sci. Soc. America Proc.*, v. 29, p. 711-717.