Structured Under#Specification of Life Cycle Impact Assessment Data for Building Assemblies

Structured under-specification of life cycle impact assessment data for building assemblies

Paolo Tecchio[1], Jeremy Gregory[2], Randa Ghattas[3] Randolph Kirchain[4]

[1]Materials Systems Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

[2]Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

[3]Campus Construction, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

[4]Materials Processing Center, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

Address correspondence to:

Paolo Tecchio,

Department of Applied Science and Technology, Politecnico di Torino, 24 Corso Duca degli Abruzzi, Torino, (IT).

Email: paolo.tecchio@gmail.com

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jiec.12746.

Summary

The existence of uncertainties and variations in data represents a remaining challenge for life cycle assessment. Moreover, a full analysis may be complex, time consuming and implemented mainly when a product design is already defined. Structured under-specification, a method developed to streamline life cycle assessment, is here proposed to support the residential building design process, by quantifying environmental impact when specific information on the system under analysis cannot be available. By means of structured classifications of materials and building assemblies, it is possible to use surrogate data during the life cycle inventory phase, and thus to obtain environmental impact and associated uncertainty. The bill of materials of a building assembly can be specified using minimal detail during the design process. The low-fidelity characterization of a building assembly and the uncertainty associated with these low levels of fidelity are systematically quantified through structured under-specification using a structured classification of materials. The analyst is able to use this classification to quantify uncertainty in results at each level of specificity. Concerning building assemblies, an average decrease of uncertainty of 25% is observed at each additional level of specificity within the data structure. This approach was used to compare different exterior wall options during the early design process. Almost 50% of the comparisons can be statistically differentiated at even the lowest level of specificity. This data structure is the foundation of a streamlined approach that can be applied not only when a complete bill of materials is available, but also when fewer details are known.

[running head: METHODS, TOOLS, AND SOFTWARE]

[running foot: *Tecchio et. al.,* Structured under-specification LCA data for buildings]

Key words: buildings, industrial ecology, streamlined life cycle assessment, structured under-specification, uncertainty analysis

Conflict of interest statement: The authors have no conflict to declare.

<heading level 1>Introduction

Life Cycle Assessment (LCA) is a scientific and structured methodology, based on international standards (ISO 14040 2010; ISO 14044 2010), that represents the reference for environmental impact assessment. The aim of LCA is to quantify the environmental consequences of products and services from “cradle to grave” (Finnveden et al. 2009; Curran 2013). The increased awareness of environmental sustainability in the construction sector has resulted in a significant number of publications focused on LCA applied to buildings and building products (for reviews of these publications see Ghattas et al. 2016; Cabeza et al. 2014; Ortiz et al. 2009; Khasreen et al. 2009; Buyle et al. 2013), and in the growth of various assessment frameworks and tools for rating building sustainability (Schwartz and Raslan 2013; Bayer et al. 2010; Haapio and Viitaniemi 2008; Anand and Amor 2017).

Unfortunately, conventional LCAs are complex and time-consuming (Hochschorner and Finnveden 2003; Schulz et al. 2012), and designers are discouraged by both this complexity (Zabalza Bribián et al. 2009) and the information lag within the process (Malin 2005). As such, despite the large academic literature on buildings LCA, the use of LCA tools by building design professionals is still uncommon (Olinzock et al. 2015). When it is used, LCA is typically applied by experienced practitioners in a resource-intensive effort at the end of the design process, excluding *de facto* the possibility of obtaining environmental results to drive the decisions during the early design phases (Malmqvist et al. 2011). By contrast, if LCA can be applied to explore and innovate at the early design stage when few parameters are defined and a broad selection of options is still available, there is more potential to impact decisions and drive toward lower impact designs (Bragança et al. 2014). Recent surveys on the building design process show that few details of the building design are finalized in the early phases of the process (Ghattas et al. 2015; Olinzock et al. 2015).

In this context, more simplified or streamlined LCA methods that allow for evaluations with limited and uncertain information are needed (De Soete et al. 2014; Hunt et al. 1998). This need is not new to the field of LCA. Streamlined LCA techniques emerged early in the development of LCA to reduce the amount of effort required to conduct a study (Baumann and Tillman 2004; Pesonen and Horn 2012). Several streamlined tools exist, but they often limit the scope of analysis (Schulz et al. 2012; Bala et al. 2010), are qualitative in nature (Hochschorner and Finnveden 2003; Pesonen and Horn 2012), rely on proxy data (Tecchio et al. 2016), or do not address uncertainty analysis (Pelton and Smith 2015; Steubing et al. 2016). If not applied carefully, many streamlining methods, particularly those that limit analytical scope, will give incorrect conclusions (Hunt et al. 1998).

Streamlining has also been an important topic of research for LCA of buildings. This focus has led to the creation of a number of specialized tools aimed to facilitate the analysis of buildings (Malmqvist et al. 2011; Anand and Amor 2017). A commonly cited challenge is the burden of collecting data that characterizes the building life cycle (often referred to as the bill of materials (BOM)). One strategy to address this has been scope reduction either in terms of life cycle activities, detail of components, or impact assessments (Kellenberger and Althaus 2009; Blengini and Di Carlo 2010; Lewandowska et al. 2015; Ghattas et al. 2016) (for a detailed review of scope reduction approaches see Soust-Verdaguer et al. (2016)). In recent years, a number of studies have described the use of building information management (BIM) software to reduce this burden by automatically translating architectural drawings into the BOMs needed for a LCA. This approach has been applied to a broad range of cases including single family homes (Iddon and Firth 2013; Houlihan Wiberg et al. 2014; Hollberg and Ruth 2016; Motuzienė et al. 2016), office buildings (Flager et al. 2012; Basbagill et al. 2013, 2014), the selection of refurbishment strategies for a multi-story office building (Seo et al. 2007), school buildings (Ajayi et al. 2015) and an apartment building (Hollberg and Ruth 2016) (See review by Soust-Verdaguer et al. (2017)). One challenge for this strategy is that digital drawings may not exist at the earliest stages of building design before electronic drawings are created. Hollberg and Ruth (2016) and Basbagill et al. (2013) address this challenge by using specific BIM models that require less detail and can easily permute designs. Wang et al. (2005) and Heeren et al. (2015) avoid the challenges of BIM tools altogether by developing fully analytic models based on limited geometric information.

Each of these approaches reduces the burden of estimating quantities within the BOM. To complete that LCA, these BOM quantities must be multiplied by an inventory or impact factor per unit of material. At the early design stage, the material is also not settled.

Basbagill describes an approach that samples building designs made from many materials (Flager et al. 2012; Basbagill et al. 2013, 2014), but each component in each design is associated with a specific material from existing databases. The work described here complements earlier streamlining approaches by describing a data structure and sampling approach that would allow early stage geometric representations to be matched with an appropriately generalized set of inventory data all the while estimating the associated uncertainty in impact.

Some of the authors have previously proposed an approach to streamlining that attempts to reduce data collection burden and provides an estimate of uncertainty in the result. As described in Olivetti et al. (2013), the structured under-specification approach involves developing an inventory from the combination of a low-fidelity description of the system under investigation and a structured classification of materials (or other activities) designed to match with that low-fidelity description. By accommodating low-fidelity (i.e. high-level) information, this approach can be applied during the early design process when specific information on the system cannot be available. This approach streamlines LCA because the practitioner spends less time collecting the data required to specify details on all aspects of the system being analyzed. This makes it particularly suited for identifying promising alternatives with lower impacts even in early-stage building design where details are not settled. Finally, by leveraging the structured classifications of materials, it is possible to estimate both the environmental impact and associated uncertainty.

We applied the structured under-specification approach to building assemblies as the foundation for a broader method to streamline building LCAs. The core enabler of the under-specification approach is a hierarchically-structured dataset that captures the distribution of possible impacts for a given activity, when that activity is described with a limited amount of information. We describe such a structured dataset for building-relevant materials and assemblies and demonstrate the impacts of calculating environmental indicators (EIs) at several levels of specification. Significant uncertainty is present in all LCA results. Streamlined methods only amplify that uncertainty. As such, a key metric for any streamlining method is its ability to generate results that differentiate alternatives in a statistically-defensible manner (Heijungs and Huijbregts 2004; Heijungs and Kleijn 2001). Herein, such statistically-differentiable results will be referred to as resolvable.

This work has three primary contributions. The first is the application of the structured under-specification approach to building materials and assemblies, which required the development of a classification scheme and new technique in which both materials and assemblies are under-specified (as opposed to just materials, which has been done in the past). The second contribution is the use of four impact categories to demonstrate the efficacy of the method; Patanavanich (2011) and Olivetti et al. (2013) only analyzed cumulative energy demand (CED). These contributions provide new insights on the use of the approach and its implications for streamlining building LCA. The third contribution is the proposal and application of metrics to evaluate the performance of under-specification data structures.

<heading level 1>Methodology

We applied the structured under-specification concept to develop a hierarchical data structure by which a wide range of building materials and assemblies could be categorized. Using this categorization system and data compiled from several databases, it was then possible to obtain distributions of environmental impacts for a range of building materials and assemblies (predominantly from the residential building sector) at different levels of specificity, which can be used at different phases of the design process. Developing this data structure involves several steps, which are detailed in this section for both materials and assemblies datasets.

<heading level 2>Scope definition

Our focus is residential building materials and assemblies, primarily those related to the building structure and envelope. As such, the scope of the life cycle datasets is cradle-to-gate. A reference unit of 1 kg of mass for each material and 1 m^2 of surface area for each building assembly was used. Materials used for electrical wiring, water, sanitary and HVAC systems, furniture and appliances are not included in the current analysis.

<heading level 2>Data collection

We assembled LCI datasets of construction materials from four relevant databases: ecoinvent 2.2. (Ecoinvent 2013), PE International Professional database (Thinkstep 2014), USLCI (NREL 2014a), and Athena Sustainable Materials Institute (ASMI 2014). We used different databases, as differences in LCA results can arise from differences in the methodological approaches used for LCI modelling (Hischier and Achachlouei 2014), other than variability due to differences in LCIA modeling (Alvarenga et al. 2016; Benini and Sala 2016; Pizzol et al. 2011).

The LCI datasets used in this analysis include all the lifecycle phases from the extraction of raw materials (or use of secondary raw materials) to the production of the material (e.g. cement) or the product to be used in a building assembly (e.g. plastic connectors or concrete blocks). Transportation to the construction sites, use phase and end of life were not considered.

The database of materials initially contained 580 datasets complied from all sources, but was reduced to 530 by removing repetitious, out of scope processes (e.g., products that are not directly used in constructions), and datasets with inconsistent scope (i.e., gate-to-gate datasets).

Bills of materials (BOMs) and technical details about residential building assemblies have been retrieved from the textbook *Architectural Graphic Standards for Residential Construction* (The American Institute of Architects 2010) and from the tools Athena Impact Estimator for Buildings V4.5 and Building Energy Optimization (BEopt) V2.1 (NREL 2014b). Each assembly contained in the building assembly database is therefore represented by a BOM dataset, which contains information about quantity and type of materials.

<heading level 2>Classification

We used and adapted the MasterFormat® structure, defined by the Construction Specifications Institute (2014) (CSI), as a starting point to create a classification system (taxonomy) for the materials datasets. MasterFormat® is a standard for organizing specifications and other written information for building projects in the USA and Canada. It is structured with nested sub-divisions, providing a structured hierarchy for all the activities of a construction site.

The MasterFormat® data structure is organized into four levels which we will refer to as material levels (ML). Formally, it is a strictly finite, 1-to-n hierarchical tree which means that each entry at ML4 (the terminal level in the tree) is univocally assigned to a preceding level (ML3) and by extension therefore to an ML2 and ML1. We extended this classification system by assigning Individual material datasets to specific ML4 categories. These material datasets, therefore, represent a new fifth level of the hierarchical tree. In the end, the data structure is a five level hierarchical tree, organized as ML1 to ML5 with ML1 being the most general classification and ML5 being the most specific (i.e., individual datasets). The categorization of individual material datasets within the classification system was based on the supporting information provided in the documentation for individual LCI datasets and expert judgment. A subset of the hierarchical structure is described in Table 1, while the full structure is available in the supporting information S1 available on the Journal's website, Table S1-1.

Some additional modifications were necessary to adapt the MasterFormat® structure to accommodate all of the collected LCI datasets. First of all, we added an additional division in this structure that includes basic materials (e.g., gravel, sand, water, etc.) that may not be used *as is* in a building, but are clearly related to the sector (about 170 individual datasets were classified in this group). Additionally, the subcategories of the MasterFormat® category 05 00 00-Metals were expanded because in its current form the structure does not differentiate stainless and low-alloyed steel, primary and secondary aluminum, nor other non-ferrous metals. Adaptions were considered necessary by the authors in order to create a link between common practice in the building sector and information contained in life cycle inventories. Finally, we note that for almost all cases it was only possible to populate one ML3 category within each ML2. In that form, any analysis of ML2 would have an identical statistical dispersion, that this is the observed spread or variation in values often measured in terms of standard deviation or variance, as the single ML3 within it. We felt that would be misrepresentative and as such have omitted the ML2 level from reporting.

Table 1 gives an example of the hierarchical categorization scheme for thermal insulation. For example, the dataset “Rock wool, packed, at plant” (ML5) is univocally classified as “thermal and moisture protection” (low-level of specificity, ML1), “thermal insulation” (medium-level of specificity, ML3) and “insulation blanket” (high-level of specificity, ML4). These materials datasets act as the basis for the assemblies datasets described next. The full list of materials is available in the supporting information S1 on the Web, Table S1-1.

Table 1. Extract of MasterFormat® adapted to classify LCA datasets in nested groups (material level, ML). The extract is focused on the ML3 category Thermal Insulation. ML5 lists individual datasets

ML1	ML3	ML4	ML5
(070000) THERMAL AND MOISTURE PROTECTION	(072100) THERMAL INSULATION	(072106) Blanket insulation	Fiberglass Batt R11-50
			Glass wool mat, at_plant
			Mineral wool Batt R11-50
			Glass wool
			Rock wool
			Rock wool, at_plant
			Rock wool, packed, at_plant
			Wood wool, u=20%, at_plant
		(072126) Blown insulation	Cellulose fibre, inclusive blowing in, at_plant
			Blown Cellulose
		(072113) Board insulation	Cork slab, at_plant
			Foam glass, at_plant
			Foam glass, at_regional_storage
			Foam glass, at_regional_storage
			Polystyrene foam slab, 100% recycled, at_plant
			Polystyrene foam slab, 45% recycled, at_plant
			Polystyrene foam slab, at_plant
			Polystyrene, extruded CO2 blown, at_plant
			Polystyrene, extruded, at plant
			Polystyrene, extruded, HFC-134a blown, at_plant
			Polystyrene, extruded, HFC-152a blown, at_plant
			Urea formaldehyde foam slab, hard, at_plant
			Raw cork, at_forest_road
			Wood, cork oak, under bark, u=70%, at_forest_road
			Wood wool boards, cement bonded, at_plant
			Polyisocyanurate (high-density foam)
			Expanded Polystyrene
			Extruded Polystyrene
			Polyiso Foam Board (unfaced)
		(072119) Foamed-in-place insulation	Urea formaldehyde foam, in situ foaming, at_plant
			Urea formaldehyde resin, at_plant
			Polyurethane, flexible foam, at_plant
			Polyurethane, rigid foam, at_plant
		(072123) Loose-fill insulation	Fiberglass low filled Cavity Fill R15-38
			Fiberglass low filled Open Blow R13-60
		Pipe insulation	Tube insulation, elastomer, at_plant

ML1	ML3	ML4	ML5
			Synthetic rubber, at_plant

Another database was developed to classify a series of building assemblies, divided into eight main categories: exterior walls, interior walls, foundations, roofs, floors, windows, doors, and exterior finishes. Figure 1 provides an example of some of the building assemblies considered in this study, highlighting a specific insulated concrete form (ICF) wall. We analyzed and characterized almost three hundred building assemblies typically used in the U.S. residential sector (the full list is available in the supporting information S1 on the Web, Table S1-2). Each assembly is characterized by a BOM. Each entry within that BOM contains a material type (i.e., a reference to an entry from the material data structure) and a quantity referred to the assembly area of 1 m^2. For these cases, quantities consists of the thickness of a given layer or, in specific cases (e.g., paint), the amount in terms of kg/m^3 or kg/m^2. Table 2 provides an example of structured under-specification of materials (ML3-5) for a specific insulated concrete form wall. It is possible to note the presence of basic materials (e.g. polypropylene used as proxy data for plastic connectors).

We classified the individual assembly datasets and their environmental impacts into five hierarchical levels of specificity (AL, assembly level), AL1 to AL5 with AL1 being the first and most general classification, namely the eight main categories described before.

AL2 data were developed by using the main assembly typologies available in the literature and in tools like Athena Impact Estimator for Buildings 4.5 (ASMI 2014) and

Building Energy Optimization (BEopt) V2.1 (NREL 2014b). As an example of AL2 categorization, Figure 1 lists five wall typologies that make up the AL1 exterior wall group: a) concrete masonry units (18 individual assemblies), b) ICFs (12), c) precast concrete walls (5), d) structural insulated panels (5), and e) wood stud walls (12). Each AL2 category comprises a series of specific assemblies at AL3. Each AL3 assembly is defined by details on materials used (type) and layer thicknesses (quantity). Information about the most common used assemblies were retrieved from literature and design-support tools (The American Institute of Architects 2010; ASMI 2014; NREL 2014b). AL4 and AL5 refine the description of the assembly by using high-fidelity categories of materials. As in the previous case, the data structure created is a finite hierarchical tree, such that each entry at AL5 is univocally assigned to preceding levels (AL1-4). As such, Table 2 also provide an example of structured under-specification of an assembly (AL3-5).

The full list of assemblies described at AL1-5 levels is available in the supporting information S1 on the Web, Table S1-2 and related spreadsheets.

Table 2. Bill of materials of an insulated concrete form exterior wall (1 m^2), specified at AL5 (with ML5 datasets), and under-specified at AL4 and AL3 (with materials categories ML3 and ML4).

Assembly defined at AL5	**Assembly defined at AL4**	**Assembly defined at AL3**	**Mass [kg/m^2]**
ML5: Stucco, at plant	ML4: Cement-plastering	ML3: Plaster-gypsum	18,29
ML5: Polystyrene foam slab, at plant	ML4: Board insulation	ML3: Thermal-insulation	1,52
ML5: Concrete, normal, at plant	ML4: Cast-in-place 2400 kg/m^3	ML3: Cast-in-place	241,81
ML5: Polystyrene foam slab, at plant	ML4: Board insulation	ML3: Thermal-insulation	1,52
ML5: Nails	ML4: Metal-fastener	ML3: Metal-fastener	1,99E-08

ML5: Gypsum fibre board, at plant	ML4: Gypsum-board	ML3: Plaster-gypsum	14,29
ML5: Polypropylene, granulate, at plant	ML4: Polypropylene	ML3: Plastic	0,52
ML5: Reinforcing steel, at plant	ML4: Bar	ML3: Concrete-reinforcing	1,10
ML5: Joint compound	ML4: Supports-gypsum-board	ML3: Plaster-Gypsum	1,10
ML5: Paper tape	ML4: Supports-gypsum-board	ML3: Plaster-Gypsum	0,01
ML5: Sawn timber, softwood, raw, kiln dried, u=20%, at plant	ML4: Soft-dried wood	ML3: Rough-carpentry	7,20
ML5: Alkyd paint, white, 60% in H2O, at plant	ML4: Paint	ML3: Interior	0,08

Note: ML = material level; AL = assembly level

AL1

- Exterior walls (52)
  - **AL2**
  - Concrete Masonry Units (18)
  - Insulated Concrete Forms (12)
    - **AL3**
    - 2" insulation, 4" concrete, 2" insulation
    - 3" insulation, 4" concrete, 3" insulation
    - 4" insulation, 4" concrete, 4" insulation
    - 2" insulation, 6" concrete, 2" insulation
    - 3" insulation, 6" concrete, 3" insulation
    - 4" insulation, 6" concrete, 4" insulation
    - …
  - Precast Concrete walls (5)
  - Structural insulated panels (5)
  - Wood stud walls (12)
- Interior walls (4)
  - **AL2**
  - Insulated walls (2)
  - Non-insulated walls (2)
- Foundations (49)
  - **AL2**
  - Concrete Masonry Units (24)
  - Cast in place (8)
  - Insulated Concrete Forms (9)
  - Slab on grade (8)
- Doors (7)
  - **AL2**
  - Glazed doors (3)
  - Unglazed doors (4)
- Windows (48)
  - **AL2**
  - Aluminum frame (8)
  - Fiberglass frame (8)
  - PVC frame (8)
  - Wood frame (8)
  - Wood+Aluminum frame (8)
  - Wood+vinyl frame (8)
- Roofs and ceilings (96)
  - **AL2**
  - Wood joists ceilings (12)
  - Wood joists roofs (36)
  - I-joists ceilings (12)
  - I-joists roofs (36)
- Floors (12)
  - **AL2**
  - Wood joists floors (4)
  - Wood truss (8)
- Exterior finishes (29)
  - **AL2**
  - Wall finishes (18)
  - Roof finishes (11)

Figure 1. Extract of the hierarchical structure developed for building assemblies (AL1 and AL2). The highlighted example at AL3 is for an ICF wall. Values in parenthesis represent the number of assemblies within a given category.

Note: AL = assembly level

[running head: METHODS, TOOLS, AND SOFTWARE]

[running foot: *Tecchio et. al.,* Structured under-specification LCA data for buildings]

<heading level 2>Probabilistic impact assessment and aggregation

We used existing datasets to calculate life cycle impacts using the life cycle impact assessment (LCIA) method Tool for the Reduction and Assessment of Chemical and other environmental Impacts (TRACI) version 2.1, with a reference unit of 1 kg of mass for each material and 1 m^2 for each assembly. Conversions were made where necessary (e.g., unit of volume to unit of mass) by using information available from the primary source of data, database supporting information or from RSMeans (2013). TRACI is a midpoint-oriented LCIA method that includes a range of EIs (Bare 2012; Bare et al. 2006) and is implemented in LCA tools used for this research. Therefore, results were obtained directly from these LCA tools, namely Simapro, GaBi and Athena Impact Estimator for Buildings. However, due to different impact assessment methods in some LCA tools and databases, some impact categories were not evaluated (e.g., Athena Impact Estimator for Buildings V4.5 did not include carcinogenic and non-carcinogenic effects, ecotoxicity and fossil fuel depletion). The categories included in our analysis are acidification (AP), eutrophication (EP), global warming (GW), and smog creation (SM). We did not include normalization, grouping, or weighting.

The procedure for calculating and aggregating probabilistic impact assessment results for the structured hierarchy is schematically represented in the supporting information S2 available on the Journal's website, section S1.

For ML5 entries, uncertainty is present for a number of reasons, including measurement error, (in)completeness, and natural variation as well as the technical, temporal, and geographic correlation of the data with its intended use here (Frischknecht et al. 2004; Heijungs and Suh 2002; Clavreul et al. 2013, 2012). Although these factors would vary for each LCA dataset, to test the data structure, we have applied a uniform estimate of uncertainty at ML5. ML5 uncertainty of a given environmental impact, was evaluated using a log-normal distribution with the mean based on the nominal value presented in the data set and the squared geometric standard deviation (perturbation term) equal to 1.27. This value happens to correspond to the geometric standard deviation that would result from selecting midrange values for all indicators within the pedigree matrix first described by Weidema and Wesnaes (1996) and more recently updated by Muller et al. (2016). The authors chose this point of reference as it suggested that such a level of uncertainty was expected in at least some forms of LCA studies. ML5 entries were modeled as following a lognormal distribution, as it is frequently observed in real life populations (Koch 1966), is representative of many LCA input parameters (Huijbregts et al. 2003), and the majority of parameters for real life populations are always positive (Weidema et al. 2013). Using a uniform estimate of the uncertainty, instead, was a simplified approach adopted in Olivetti et al. (2013), in which any existing estimates of uncertainty for individual LCI datasets were therefore overwritten.

Because of the hierarchical nature of the data structure, by defining the impact and the associated uncertainty of each specific material (at ML5), distributions of impacts at other levels of specificity (from ML1 to ML4) are implicitly assigned, as they derive from the aggregation of all of the distributions of results of those materials included in a given set of data (e.g., for ML3 “Thermal insulation” this consists of all the ML5 results for the datasets included in that set of data). The distribution of impacts for levels ML1 to ML4 were estimated through a two stage Monte Carlo sampling process. First, a specific material (ML5) that is a member of given category was randomly selected. Then an impact value was randomly sampled from the distribution associated with that ML5 material.

Monte Carlo simulations were also used for estimating the distribution of impacts in the assemblies datasets. For each assembly, a Monte Carlo simulation with 1,000 samples was used to obtain a distribution of results for each level of specificity and for each impact category. First of all, each assembly can be seen as a set of components (materials and products, as specified in <heading level 2>Data collection). For each assembly, we calculated a distribution of results using the highest level of specificity, here referred to as AL5. At AL5, each component is represented by a unique ML5 material. As noted earlier, the impact of the materials at ML5 are uncertain and assumed to follow a log-normal distribution. Subsequently, we used the ML4 materials categories to calculate under-specified results for the assembly level AL4. The distribution of AL4 impacts was estimated through a two stage Monte Carlo sampling process; first, selecting a specific material (ML5) in that category, then an impact value from the distribution associated with that ML5 material. In this way the uncertainty in results is increased by the variation of possible materials for each component. The same applies for AL3 with materials coming from ML3 categories.

We then observed that using material category ML1 as the basis for building assembly levels AL1 and AL2 led to an unworkable degree of uncertainty. Therefore, the distribution of results of a given type of building assembly at AL2 (e.g., ICF walls) is represented by all the results obtained at AL3 for individual building assemblies of that type (e.g., ICF walls as specified in supporting information S1 on the Web, Table S1-2). The same applies at AL1, where all the AL3 results for a given class (e.g., exterior walls) building assemblies are combined.

The procedure is also schematically represented in the supporting information S2 on the Web, section S2.

<heading level 2>Evaluation

We evaluate the proposed data structures using two related sets of analyses. First we evaluate the four EIs for a set of two case examples (a material, rock wool, and an assembly, ICF walls) and then for the data structures as a whole and compute how the spread in the estimate of those metrics declines with increasing specificity. These results are compared among the ML1 categories for the data structure described here and against a novel analysis of the data structure utilized by Olivetti et al. (2013). An effective data taxonomy should most frequently reduce the dispersion in the estimate as more information is provided. Second, we test how this taxonomy allows the environmental performance of the building materials and assemblies to be statistically resolved (i.e., provide results with sufficient fidelity that when compared the confidence in their difference is deemed statistically significant). In the following formulas, $x_{i,EI,ALj}$ represents the i^{th} simulated result obtained for a given EI and for a defined level of specificity (ALj, j = 1:5) and $X_{EI,ALj}$ represents a vector of results obtained by Monte Carlo simulations for a given EI [x_1, x_2, x_3,…, x_{1000}] and a given level of specificity.

<heading level 3>Statistical dispersion of results

The hierarchical data structures being tested here serve a role much like a statistical tree model where the metric of interest (environmental impact) is a continuous response. The performance of such models, also referred to as regression trees, is measured in terms of the homogeneity of the categories defined with the tree (Loh 2002; Loh et al. 2008). When categories are more homogenous, the characteristics of the category better represent the characteristics of the members of the category. This is directly analogous to our own use of the taxonomy where the performance of the taxonomy ultimately derives from how well the LCA characteristics of the categories represent the LCA characteristics of category members. The most common metric of homogeneity is the total sum of squared (TSS) errors from the category mean, here computed for the assembly database as TSS $_{EI,ALj}$ (for a given EI specificity (ALj)):

$$TSS_{EI,ALj} = \sum_{i=1}^{n} \left(x_{i,EI,ALj} - mean(X_{EI,ALj})\right)^2 \qquad (1)$$

TSS is a common metric for evaluating alternative categorization schemes for a specific taxonomy. Its magnitude, however, is a function of the number and scale of observations. As such, TSS is not particularly useful for comparing the performance of different taxonomies. For that purpose, we compute average metrics of dispersion scaled to the expected value of the sample.

Specifically, we compute the coefficient of variation (CV), for each EI and for a defined level of specificity (ALj). For a given set of data, the CV $_{EI,ALj}$ can be defined as the ratio between the standard deviation and the mean (Upton and Cook 2008), as in the following formula:

$$CV_{EI,ALj} = \frac{standard\ deviation(X_{EI,ALj})}{mean(X_{EI,ALj})} \quad (2)$$

Additionally, we compute the median absolute deviation coefficient of variation (MAD-COV), a measure of data dispersion similar to standard deviation, but robust to data outliers (Rousseeuw and Croux 2012). The MAD-COV describes the median percent variation of a dataset from the median value. In equation (3), MAD-COV$_{EI,ALj}$ is defined as:

$$MAD\_COV_{EI,ALj} = \frac{median(|x_{i,EI,ALj} - median(X_{EI,ALj})|)}{median(X_{EI,ALj})} \quad (3)$$

Average values of CV are calculated as the arithmetic mean of CV values obtained at a certain level. For example, if we consider the exterior wall category, the average CV $_{EI,AL5}$ consists of the arithmetic mean of the 52 CV values obtained for a given EI at AL5. The same applies at AL4 and AL3, using CV values at the equivalent levels of specificity. At AL1, instead, the average CV is computed by using simultaneously all of the 52 results of the Monte Carlo simulation at AL3, for a total of 52,000 values. At AL2, the average CV is computed with the same procedure used at AL1, but with smaller samples, corresponding to the nested groups identified in the assembly classification; for instance, considering the exterior wall category, the 5 types of exterior walls were taken into account. The same procedure applies for the calculation of the average MAD-COV. Equations (1), (2), (3) can also be applied to the material database, when the index ML replace AL in these equations.

<heading level 3>Comparison of results

Since LCA results are usually interpreted in a comparative manner (Noshadravan et al. 2013), it is also important to include comparative metrics. To evaluate the difference between two alternative designs, we used a comparison indicator (CI). This is defined as the ratio of environmental impacts of two products (Huijbregts et al. 2003). Therefore, the overall uncertainty in individual assemblies is not the key driver, but the uncertainty in the ratio of the results of the two building assemblies (design A and design B, in equation (4)) becomes crucial. $CI_{EL,ALj}$ is the comparison indicator for a given EI at a given level of specificity:

$$CI_{EI,ALj} = \frac{X_{EI,ALj,design\ A}}{X_{EI,ALj,design\ B}} \tag{4}$$

We define $\beta_{EI,ALj}$ as the probability that design A has lower environmental burden than design B once again for a given EI, at a given AL. Given this, we can express $\beta_{EI,ALj}$ between A and B as:

$$\beta_{EI,ALj} = \mathrm{P}\left(CI_{EI,ALj} < 1\right) \tag{5}$$

where $P(CI_{EI,ALj < 1})$ means the probability of CI being less than 1. Results of the comparison are considered statistically significant if $\beta_{EI,ALj}$ exceeds some threshold, β_{crit}, (or if $\beta_{EI,ALj} \leq (1-\beta_{crit})$) established by the decision-maker. Cases that meet this criteria will be referred to as "resolvable".

<heading level 2>Comparison against available literature

A literature review was conducted to understand if other studies addressed LCA of building assemblies in a similar way. However, we realized that it is difficult to find studies that look at similar scopes, in the literature. Most of these studies focus on whole buildings or constituent materials, therefore the comparison of results appears to be unfeasible. A detailed discussion of related case work is described in the supporting information S2 on the Web, Section S9.

<heading level 1>Results

The first goal of this work is to characterize the overall performance of this under-specification structure for building materials and assemblies. Before reporting those results we first show two detailed case analyses – the performance of the material rock wool and the assembly of a wall created using an ICF wall. These analyses provide a clearer picture of how the data structure would impact analyses with various levels of available information.

<heading level 2>Case Analyses

<heading level 3>Materials

We use the rock wool insulation (Table 1) as a demonstration case for the materials datasets, to show result distributions at different levels of specificity. Figure 2 shows median, interquartile range, 5th and 95th percentiles of the EIs for the specific dataset (ML5) and for the corresponding hierarchical categories of which it is a member up through the most general – Thermal and Moisture Protection (ML1).

These plots depict the trend of generally decreasing uncertainty in results from ML1 to ML5 that is desirable within a well formed data structure. For this particular data, the dispersion of results drops for all four EIs when measured by either MAD-COV or CV with only one exception; the CV_{GW} grows from ML1 (145%) to ML3 (161%). This occurs because the ML3 category of which rock wool is a member (i.e., Thermal Insulation) contains both the highest (extruded polystyrene) and lowest (renewables-based insulation) within the ML1, while some of the central values within that ML1 are assigned to other ML3's. More details on this are provided in the supporting information S2 on the Web, section S7. EP results from ML4 and ML5 showed distributions at the extremes of the distributions associated with ML1 and ML3. This impact category had the highest variation of impacts already at the material database level, which was then analyzed in detail (see supporting information S2 on the Web, section S6). The behavior of the ML4 (Insulation blanket) and ML5 (rock wool) occurs because the values within category Insulation blanket are distributed across two uneven modes. The largest mode contains a number of low impact alternatives that pulls the median and interquartile range to lower values. The impact of the rock wool material, however, is positioned within the high impact mode. In fact, rock wool has the second highest mean EP impact among the insulation blanket materials. While such a distribution of impacts could always occur, this particular instance highlights an issue that needs to be considered in future taxonomy development. Namely, sometimes large differences in average impacts across databases. The two modes within the insulation blanket category are notable in part because each mode is dominated by materials from distinct databases. Future research should explore the implications of either limiting the application of this method to a single database or otherwise adjusting databases to make their results more comparable.

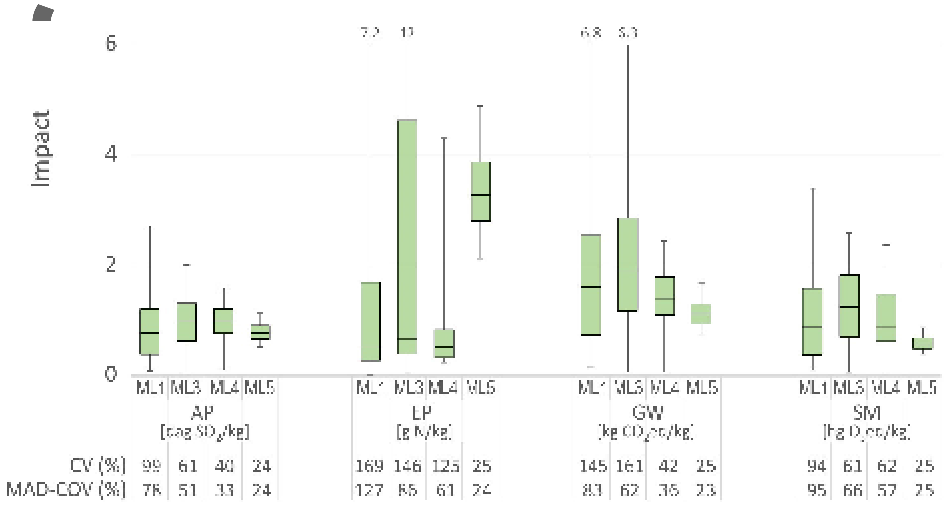

Figure 2. Probabilistic distributions of impact metrics (AP, EP, GW, SM) for the rock wool dataset, at ML1-ML5. ML1 = thermal and moisture protection; ML3 = thermal insulation; ML4 = blanket insulation; ML5 = rock wool packed, at plant.

Note: AP = acidification; EP = eutrophication; GW = global warming; SM = smog creation; ML = material level; CV = coefficient of variation; MAD-COV = median absolute deviation coefficient of variation.

<heading level 3>Assemblies

We use the ICF wall described in Table 2 as a demonstration case for the assemblies datasets. Figure 3 shows the median, interquartile range, 5th and 95th percentiles of the four EIs, progressing from a generic (AL1) exterior wall and finishing with the specific (AL5) ICF wall.

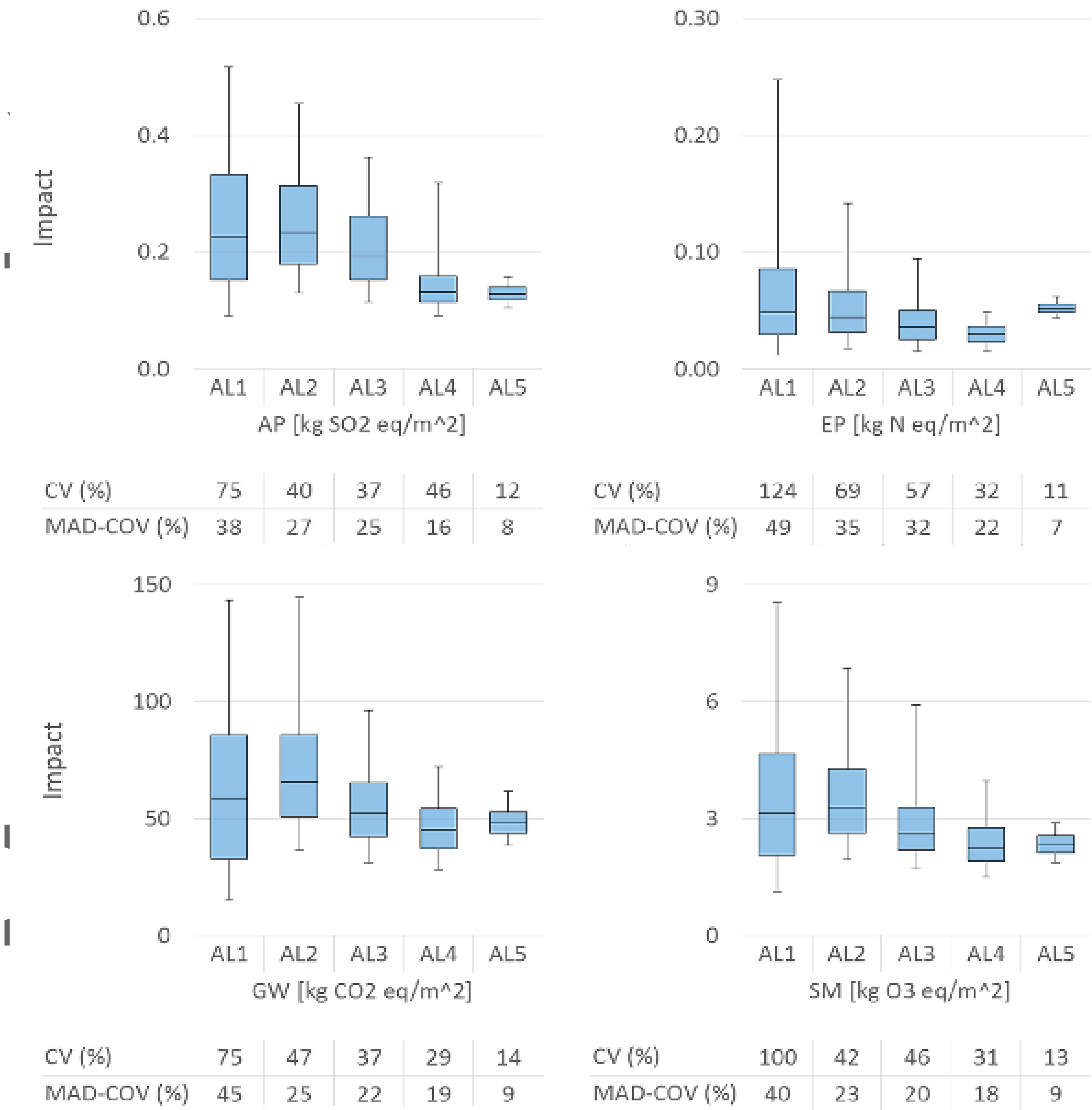

Figure 3. Probabilistic distributions of impact metrics (AP, EP, GW, SM) for an ICF wall (BOM in Table 2). AL1 = exterior walls; AL2 = ICF walls; AL3 = ICF wall (BOM as in Table 2, materials specification at ML3); AL4 = ICF wall (BOM as in Table 2, materials specification at ML4); AL5 = ICF wall (BOM as in Table 2, materials specification at ML5)

Note: AP = acidification; EP = eutrophication; GW = global warming; SM = smog creation; AL = assembly level; CV = coefficient of variation; MAD-COV = median absolute deviation coefficient of variation.

Figure 3 shows MAD-COV and CV $_{EI\ AL1\text{-}AL5}$ results for four EIs. Boxplots also demonstrate the reduced, but still present, uncertainty at AL5. Concerning GW at AL5, it is possible to observe a median impact of 48.2 kg CO_2 eq/m^2 and a $CV_{GW,AL5}$ of 14%. Using the under-specified categories described in Table 2, AL4 and AL3 results are characterized by a similar median GW (respectively 45.2 and 52.1 kg CO_2 eq/m^2) but a wider distribution, with $CV_{GW,AL4}$ of 29% and $CV_{GW,AL3}$ of 37%. AL2 and AL1 boxplots appear even larger because of the variation of different assemblies within the same category (12 ICF walls in AL2 and 52 exterior walls in AL1).

MAD-COV values with the highest levels of specificity (MAD-COV $_{EI,AL5}$) range from 7% to 9%, whereas they range from 38% to 49% for the most generic information (MAD-COV $_{EI,AL1}$) and from 20% to 32% for an average level of detail (MAD-COV $_{EI,AL3}$). EP showed the greatest dispersion of results. The same trend can be seen in the CV values, which decrease from AL1 (75-124%) to AL5 (11-14%).

When this approach is put into practice, the user would be able to quantify the uncertainty at each level of detail, from AL1 to AL5. While at AL5 a full fledged LCA is required, this is not the case for previous levels. At AL3 and AL4 less specificity is required for the different materials used for the building assembly (it can be possible to define a generic insulation board, for example). The burden of collecting data is drastically reduced when AL2 is considered (only the definition of the type of wall is needed). At AL1, in principle, the only piece of information required is the total area of the building assembly, and without any further specification, a range of environmental impacts is available.

<heading level 2>Material Categories

To understand the performance of the proposed data structure we explore several metrics of categorization performance and compare them to a previously published data structure. Figure 4 plots average MAD-COV $_{GW\ ML1\text{-}ML5}$ for each level of specificity when the data is organized by the ML1 levels (TSS and average MAD-COV plots for this and for the other impact categories are provided in the supporting information S2 on the Web, section S3). This and the corresponding plots in the supporting information S2 on the Web depict the trend of decreasing uncertainty in results from ML1 to ML5 that is desirable within an effective data structure. The average MAD-COV $_{GW,ML1}$ ranges from 51% for the Openings category to 134% for Basic materials. Also, this dispersion decreases when assessed by all three metrics for every increase in specificity across every category. This trend continues across the other EIs as well. For this particular data, the dispersion of results, as measured by TSS, drops for all eight categories and all four impact assessment metrics. When measured by either average MAD-COV or average CV, dispersion also decreases with only four exceptions; the average CV $_{SM}$ in the metals category grows slightly from CV $_{SM,ML3}$ (46%) to CV $_{SM,ML4}$ (49%) and the average MAD-COV $_{EP}$ grows for three categories (Concrete, Metals and Openings) in the transition from ML3 to ML4 (note: Monte Carlo sampling inherently generates some amount of error in estimating any parameter. To account for this, in this analysis we consider values equivalent if they are within 0.5% to account for error in estimation attributable to Monte Carlo sampling).

Considering each individual subgrouping of data, in 80% to 88% of cases, the ML3 category has an equal or lower MAD-COV than the corresponding ML1, and in 83% to 90% of cases MAD-COV $_{EI,ML4}$ is lower than the corresponding ML3. On average, the MAD-COV drops by around 33% (CV by approximately 40%) with each additional level of specificity across the four metrics.

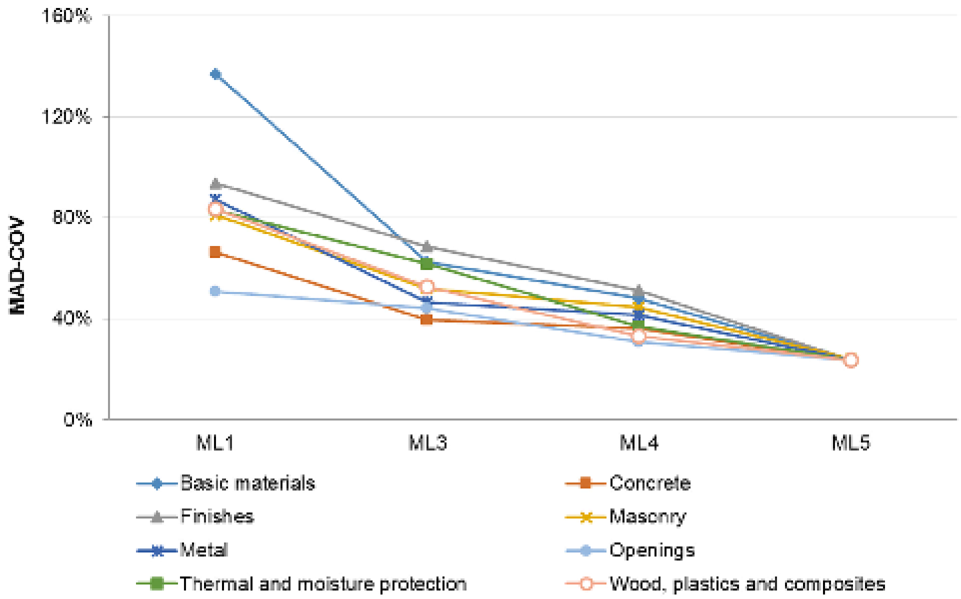

Figure 4. Average median absolute deviation coefficient of variation (MAD-COV) values for global warming (GW) for all building material categories (from material level ML1 to material level ML5)

To provide some benchmark of these results, the authors analyzed the product dataset (labeled P11) originally assembled by Patanavanich (2011) that assessed CED for a broad set of materials. In Olivetti et al. (2013), the P11 dataset proved sufficiently effective to identify the key drivers (probabilistic triage) of impact across seven case studies, even at the ML1 level of specificity. For P11, the MAD-COV ranged from 29% to 148% and average CV ranged from 34% to 282%. The span of dispersion within the data structure explored here is within that for P11 (it is worth noting that this is true even considering that P11, when compared to this study, was developed assuming a lower level of uncertainty for the database entry level (ML5)). This should imply that for similarly structured cases the dataset presented here should be at least as effective as P11 for probabilistic triage.

Like the dataset explored here, metrics of dispersion decline for P11 for most additions of specificity. When considering the transition of individual categories within P11, however, in only 63% of cases was the MAD-COV of ML2 equal or lower than that of the corresponding ML1 and in only 58% of cases was ML3 lower than the corresponding ML2. As such, the currently proposed materials structure seems to be more efficient with gains in specificity than P11.

<heading level 2>Assembly Categories

Performance metrics were also calculated for the assemblies data structure and Figure 5 depicts average MAD-COV $_{GW\ AL1\text{-}AL5}$ for each level of specificity (complete set of plots in the supporting information S2 on the Web, section S4). As observed for materials, the assemblies data structure exhibits generally decreasing dispersion in results from AL1 to AL5. The average MAD-COV $_{GW,AL1}$ ranges from 33% for the Windows and Interior Walls categories to 109% for Finishes, and average CV ranges from 31% to 285% for Windows and

Finishes. These are well within the range of P11 suggesting that the assemblies data structure should be effective at probabilistic triage even at the AL1 level.

When measured according to average measures of error, the assemblies data structure also performs well. In terms of average CV, all additions of specificity lead to improvements in normalized error except for four including the transition from AL1 to AL2 for Exterior Walls (75% to 81%). Using the measure of MAD-COV, average error declines for all additions of specificity except for 10 transitions (seven in the transition from AL3 to AL4) with an average drop of nearly 25% for each additional level of specificity. Considering each individual category, in 65% to 88% of cases, the AL2 category has an equal or lower MAD-COV than the corresponding AL1 and in 82% to 97%of cases AL3 MAD-COV is lower than the corresponding AL2. The most populated category is Roofs and ceilings, and gets the most benefit in terms of dispersion reduction when moving from AL2 to AL3 (MAD-COV from 92% to 53%); this drop is mainly due to the classification used at AL2, which considered only two subcategories.

Given these measures, it would appear that the assemblies data structure should be at least as efficient as P11 with gains in specificity. Furthermore, it appears that most of the challenges appear in the transition from AL3 to AL4 and with the EP metric.

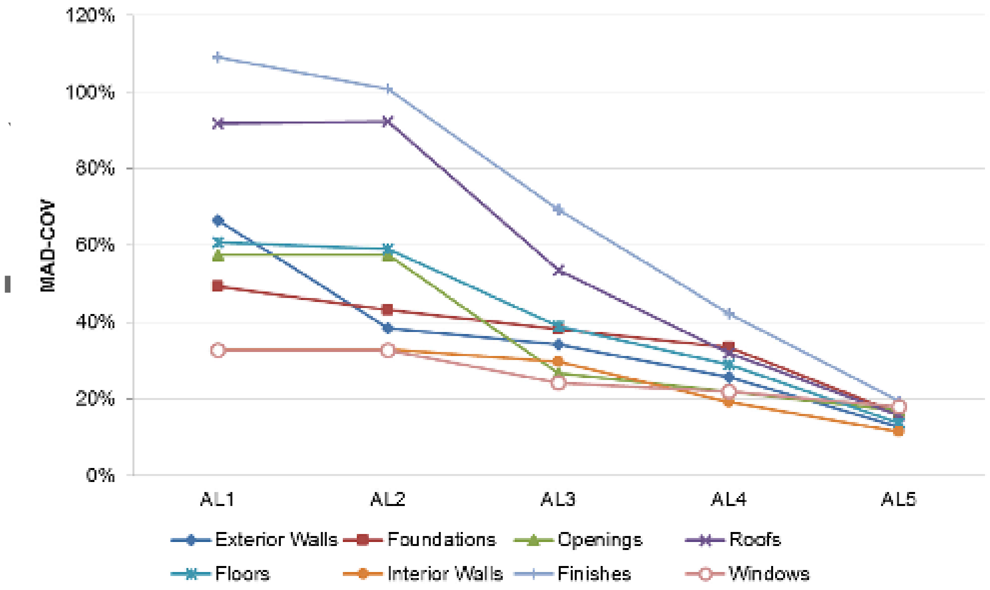

Figure 5. Average median absolute deviation coefficient of variation (MAD-COV) values for global warming (GW) for all building assembly categories (from assembly level AL1 to assembly level AL5)

<heading level 2>Comparisons

We further explore the utility of the data structure by calculating CIs (formulas (4) and (5)) for the ICF wall described in Table 2 (design A) in reference to the other results of exterior walls analyzed in this study, for 51 different designs B, and therefore a total of 51 comparisons. Each exterior wall was assigned with a code, from 1 to 52. Codes and specifications of the 52 exterior walls are detailed in the supporting information S2 on the Web, Table S1-2. It is worth noting that this is a demanding test of data structure performance. The authors do not make a claim that underspecified data is sufficient to reach most LCA conclusions. The role of the underspecified data is to identify the most likely large contributors for further, targeted specification. Furthermore, in the context of a whole building LCA the energy performance of alternatives would also contribute to their differences. Nonetheless, this provides an interesting test of performance.

Figure 6 provides the results of these comparisons performed at different levels of specificity (AL3: black dots & AL5: green dashes, AL4 results are described in the supporting information S1 on the Web). Results show that the higher the specificity level (AL5), the better the resolution of results: a β value close to 50% implies a situation in which it is not possible to determine the design alternative with the lower environmental impact, while β values approaching 100% or 0% are clearly identifying the lower impact designs (respectively design A or design B). Thus, in order to verify the resolution of results at different levels, probabilities can be divided in two groups: 1) $\beta > \beta_{crit}$ or $\beta < (1 - \beta_{crit})$ – green background color on the plot, and 2) $(1 - \beta_{crit}) \leq \beta \leq \beta_{crit}$. For this analysis, β_{crit} was assumed equal to 90%. The first group contains well-differentiated situations (resolvable), while the second group contains ambiguous results (unresolved).

For these designs and this critical threshold, it is notable that 18 (35%) of the 51 comparisons remain unresolved at AL5. That means that the impacts of the designs are similar enough that the level of uncertainty present in the LCI database is sufficient to prevent a statistically defensible conclusion. Of these 33 resolvable comparisons, at AL3 it is possible to identify 16 resolvable comparisons (50%). This grows to 21 (65%) at AL4. The results highlight that the method as a whole can be quite efficient. With very low resolution information (AL3) it is possible to resolve half of the resolvable (33) comparisons. With a little more information (AL4), nearly two-thirds are resolvable. As this is a statistically based assessment, we would assume some fraction of false conclusions. In fact for these threshold values, up to 10% of outcomes represent an opposite conclusion. For these cases and this threshold, there are 4 cases (exterior walls 48, 49, 50, and 51) for which the AL3 suggests that the cases are resolvable, while the AL5 result finds them unresolved. There is also one case (exterior wall 52), that is nearly resolvable as A preferred at AL3, but then shifts to B preferred for AL5. All of these occur because at least part of the assembly is made of a material at the extreme of its class's impact distribution.

As noted earlier, these results alone cannot diagnose whether the data structure is sufficient for a specific case, but they provide a benchmark against which future data structures or modifications to this one can be compared. The full set of results is available in supporting information S2 on the Web, Section S8.

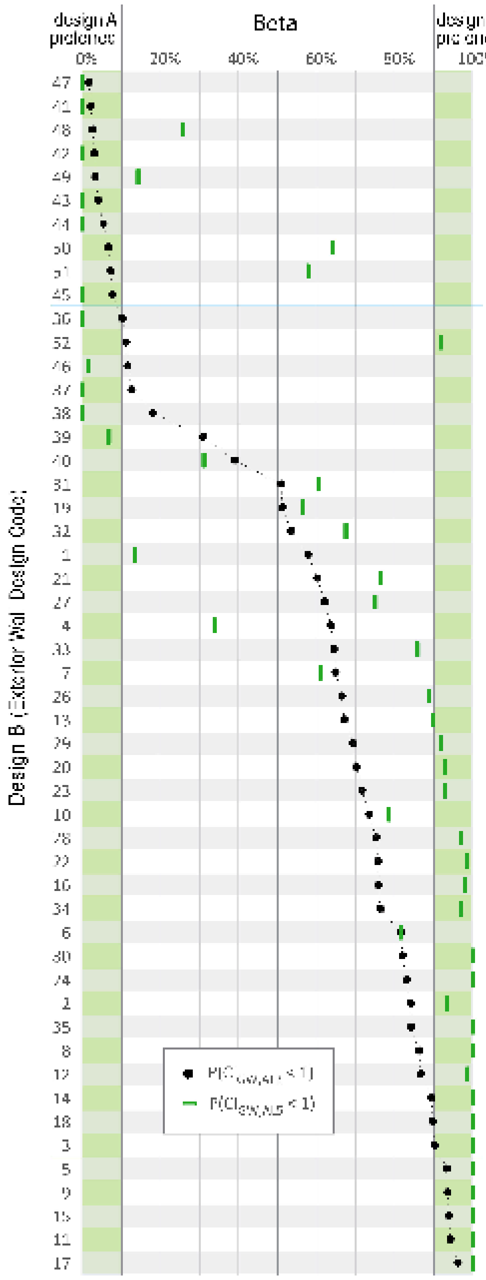

Figure 6. Probability β that design A has lower environmental impact than design B, evaluated at AL3 and AL5. The EI considered for this evaluation was GW. AL1 and AL2 were not considered as representative of groups of assemblies. Full set of results is available in supporting information S2 on the Web, Section S8.

Note: CI = comparison indicator, AL = assembly level

<heading level 1>Discussion

In this document we have proposed an approach that can be used to implement structured under-specification for building materials and assemblies. Altogether the data structure is a hierarchical classification that comprises 530 materials, 8 categories and 4 specificity levels (ML1-5, ML2 has been excluded), and 297 building assemblies, 8 categories and 5 specificity levels (AL1-5). This set should be sufficiently comprehensive to characterize the structural, envelope, and finishing materials for most residential construction in the USA. The hierarchical nature of the data structure means that individuals can get estimates of impact for any level of knowledge about the materials or assemblies.

Additionally, we have proposed some initial metrics by which such data structures can be evaluated. Specifically, we have examined overall TSS, progress of average dispersion (average CV, average MAD-COV), and the frequency of reduced dispersion with additional classification information. These metrics were applied to the data structures described here, as well as in a novel analysis of the only previously characterized under-specification data structure originally assembled by Patanavanich (2011) referred to here as P11, as in Olivetti et al. (2013). The results presented indicate that the materials and assemblies data structure should be at least as effective at probabilistic triage as the under-specification data structure presented by P11 for general materials. Furthermore, it would appear that the materials data structure described here gains resolution much more rapidly than P11. Finally, through a case analysis of 52 wall assemblies, we have demonstrated that a significant number of comparisons can be statistically resolvable even at low levels of fidelity, and therefore that structured under-specification has real potential to streamline building-related LCA.

These results suggest that the data structures presented here, which are based on a widely used standard within the construction industry, could be used to produce quantitative environmental results, including uncertainty, using low-fidelity information about building materials and assemblies. The method is the foundation of a streamlined LCA approach that will enable building LCAs to be conducted even when few details about the design have been decided.

There are ample opportunities to improve this approach in future work. One area is the use of different and more efficient ways to structure the taxonomy for materials and for assemblies (e.g., using cost or material properties). Figure 5 shows how data dispersion at AL1 and AL2 is wider for some assembly categories, indicating that a more structured approach could possibly be used to improve taxonomies. In this work, we made use of LCI data from several sources. Differences in the data collection strategies across those sources is a key source of observed impact variation in this present work. Further research should be done to evaluate the performance of this method when LCI datasets are gathered from only one database. This would bring benefits from data harmonization, but has the downside of limiting the number of options available to the user.

Future work should explore the possibility of using an alternative classification at AL2 for the roof and ceiling category, which should then reduce dispersion of results. Lee (2013) highlighted that poor data structures reduced accuracy in streamlining, while Reis (2013) found that data mining techniques support structured under-specification, providing guidance for developing taxonomies.

Another area that will be included in future work is the use of probabilistic triage in combination with structured under-specification. This approach determines the specific parameters of influence for a system and illuminates which data should be specified further in an LCA model. It involves the identification of the set of interest of the BOM that should be specified at ML5 in order to obtain reliable results with less effort (Olivetti et al. 2013).

The final area worthy of future work is testing the approach with building professionals to determine how best this approach can be used to support design decisions early in the design process.

Acknowledgments

This research has been conducted at the Concrete Sustainability Hub (CSHub), with support provided by the Portland Cement Association, the Ready Mixed Concrete Research & Education Foundation, and the Compagnia di San Paolo.

Supporting Information

Additional supporting information may be found in the online version of this article:

Supporting Information S1: This supporting information contains two main tables. Table S1-1 contains the materials database, classified into ML1, ML3, ML4, ML5 and CSI categories. The source of the dataset is included. Table S1-2 contains the list of assemblies, classified into AL1 and AL2 categories, with a description of the building assembly. Assemblies are hyperlinked to specific spreadsheets (1-297) for the definition of the bill of materials at ML3, ML4 and ML5.

Supporting Information S2: This supporting information includes sections: Section S1: materials. ML-level result distribution calculation, Section S2: assemblies. AL-level data creation, Section S3: TSS and average MAD-COV for the materials data structure, Section S4: TSS and average MAD-COV for the assemblies data structure, Section S4: TSS and average MAD-COV for the assemblies data structure, Section S5: TSS and average MAD-COV for the assemblies data structure, Section S6: EP trends, Section S7: Exploring the Behavior of GW in the category Thermal and Moisture Protection, Section S8: Comparison indicators, Section S9: Comparison against available literature.

References

Ajayi, S.O., L.O. Oyedele, B. Ceranic, M. Gallanagh, and K.O. Kadiri. 2015. Life cycle environmental performance of material specification: a BIM-enhanced comparative assessment. *International Journal of Sustainable Building Technology and Urban Development* 6(1): 14–24.

Alvarenga, R., I. Lins, and J. Almeida Neto. 2016. Evaluation of Abiotic Resource LCIA Methods. *Resources* 5(1): 13.

Anand, C.K. and B. Amor. 2017. Recent developments, future challenges and new research directions in LCA of buildings: A critical review. *Renewable and Sustainable Energy Reviews* 67: 408–416.

ASMI. 2014. Athena Sustainable Materials Institute. http://www.athenasmi.org/. Accessed April 23, 2015.

Bala, A., M. Raugei, G. Benveniste, C. Gazulla, and P. Fullana-i-Palmer. 2010. Simplified tools for global warming potential evaluation: when "good enough" is best. *The International Journal of Life Cycle Assessment* 15(5): 489–498.

Bare, J. 2012. *Tool for the Reduction and Assessment of Chemical and other Environmental Impacts (TRACI) User's manual*. Cincinnati: U.S. Environmental Protection Agency.

Bare, J., T. Gloria, and G. Norris. 2006. Development of the Method and U.S. Normalization Database for Life Cycle Impact Assessment and Sustainability Metrics. *Environmental Science & Technology* 40(16): 5108–5115.

Basbagill, J., F. Flager, M. Lepech, and M. Fischer. 2013. Application of life-cycle assessment to early stage building design for reduced embodied environmental impacts. *Building and Environment* 60: 81–92.

Basbagill, J.P., F.L. Flager, and M. Lepech. 2014. A multi-objective feedback approach for evaluating sequential conceptual building design decisions. *Automation in Construction* 45: 136–150.

Baumann, H. and A.-M. Tillman. 2004. *The Hitch Hiker's Guide to LCA. An orientation in life cycle assessment methodology and application*. Lund (Sweden): Studentlitteratur.

Bayer, C., M. Gamble, R. Gentry, and S. Joshi. 2010. *AIA Guide to Building Life Cycle Assessment in Practice*. Washington DC: The American Institute of Architects.

Benini, L. and S. Sala. 2016. Uncertainty and sensitivity analysis of normalization factors to methodological assumptions. *The International Journal of Life Cycle Assessment* 21(2): 224–236.

Blengini, G.A. and T. Di Carlo. 2010. Energy-saving policies and low-energy residential buildings: an LCA case study to support decision makers in Piedmont (Italy). *The International Journal of Life Cycle Assessment* 15(7): 652–665.

Bragança, L., S.M. Vieira, J.B. Andrade, S.M. Vieira, and J.B. Andrade. 2014. Early stage design decisions: the way to achieve sustainable buildings at lower costs. *TheScientificWorldJournal* 2014: 365364.

Buyle, M., J. Braet, and A. Audenaert. 2013. Life cycle assessment in the construction sector: A review. *Renewable and Sustainable Energy Reviews* 26: 379–388.

Cabeza, L.F., L. Rincón, V. Vilariño, G. Pérez, and A. Castell. 2014. Life cycle assessment (LCA) and life cycle energy analysis (LCEA) of buildings and the building sector: A review. *Renewable and Sustainable Energy Reviews* 29: 394–416.

Clavreul, J., D. Guyonnet, and T.H. Christensen. 2012. Quantifying uncertainty in LCA-modelling of waste management systems. *Waste Management* 32(12): 2482–2495.

Clavreul, J., D. Guyonnet, D. Tonini, and T.H. Christensen. 2013. Stochastic and epistemic uncertainty propagation in LCA. *The International Journal of Life Cycle Assessment*

18(7): 1393–1403.

Construction Specifications Institute, C. 2014. *MasterFormat® Numbers & Titles*. The Construction Specifications Institute and Construction Specifications Canada.

Curran, M.A. 2013. Life Cycle Assessment: a review of the methodology and its application to sustainability. *Current Opinion in Chemical Engineering* 2(3): 273–277.

Ecoinvent. 2013. ecoinvent database. http://www.ecoinvent.ch/. Accessed April 23, 2015.

Finnveden, G., M.Z. Hauschild, T. Ekvall, J. Guinée, R. Heijungs, S. Hellweg, A. Koehler, D. Pennington, and S. Suh. 2009. Recent developments in Life Cycle Assessment. *Journal of Environmental Management* 91(1): 1–21.

Flager, F., M. Basbagill, M. Lepech, and M. Fischer. 2012. Multi-objective building envelope optimization for life-cycle cost and global warming potential. In *eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2012*, ed. by Gudni Gudnason and Raimar Scherer, 193–200. CRC Press.

Frischknecht, R., N. Jungbluth, H.-J. Althaus, G. Doka, R. Dones, T. Heck, S. Hellweg, et al. 2004. The ecoinvent Database: Overview and Methodological Framework. *The International Journal of Life Cycle Assessment* 10(1): 3–9.

Ghattas, R., J. Gregory, T.R. Miller, and R. Kirchain. 2015. *The decision-making process in the design of residential structures*. Cambridge (MA): CSHub - MIT Concrete Sustainability Hub.

Ghattas, R., J. Gregory, M. Noori, T.R. Miller, E. Olivetti, and S. Greene. 2016. *Life Cycle*

Assessment for Residential Buildings: A Literature Review and Gap Analysis (Revised). Cambridge (MA):CSHub - MIT Concrete Sustainability Hub.

Haapio, A. and P. Viitaniemi. 2008. A critical review of building environmental assessment tools. *Environmental Impact Assessment Review* 28(7): 469–482.

Heeren, N., C.L. Mutel, B. Steubing, Y. Ostermeyer, H. Wallbaum, and S. Hellweg. 2015. Environmental Impact of Buildings—What Matters? *Environmental Science & Technology* 49(16): 9832–9841.

Heijungs, R. and M. Huijbregts. 2004. A review of approaches to treat uncertainty in LCA. In *Transactions of the 2nd Biennial Meeting of the International Environmental Modelling and Software Society*, 332−339. Manno, Switzerland: iEMSs.

Heijungs, R. and R. Kleijn. 2001. Numerical approaches towards life cycle interpretation five examples. *The International Journal of Life Cycle Assessment* 6(3): 141–148.

Heijungs, R. and S. Suh. 2002. *The Computational Structure of Life Cycle Assessment*. Dordrecht (NL): Kluwer Academic Publishers.

Hischier, R. and M.A. Achachlouei. 2014. Evaluating the sustainability of electronic media: Strategies for life cycle inventory data collection and their implications for LCA results. *Environmental Modelling & Software* 56: 27–36.

Hochschorner, E. and G. Finnveden. 2003. Evaluation of Two Simplified Life Cycle Assessment Methods. *The International Journal of Life Cycle Assessment* 8: 119–128.

Hollberg, A. and J. Ruth. 2016. LCA in architectural design—a parametric approach. *The*

International Journal of Life Cycle Assessment 21(7): 943–960.

Houlihan Wiberg, A., L. Georges, T.H. Dokka, M. Haase, B. Time, A.G. Lien, S. Mellegård, and M. Maltha. 2014. A net zero emission concept analysis of a single-family house. *Energy and Buildings* 74: 101–110.

Huijbregts, M.A.J., W. Gilijamse, A.M.J. Ragas, and L. Reijnders. 2003. Evaluating Uncertainty in Environmental Life-Cycle Assessment. A Case Study Comparing Two Insulation Options for a Dutch One-Family Dwelling. *Environmental Science & Technology* 37(11): 2600–2608.

Hunt, R.G., T.K. Boguski, K. Weitz, and A. Sharma. 1998. Case studies examining LCA streamlining techniques. *The International Journal of Life Cycle Assessment* 3(1): 36–42.

Iddon, C.R. and S.K. Firth. 2013. Embodied and operational energy for new-build housing: A case study of construction methods in the UK. *Energy and Buildings* 67: 479–488.

ISO 14040. 2010. ISO 14040 - Environmental management -- Life cycle assessment -- Principles and framework.

ISO 14044. 2010. ISO 14044 - Environmental management -- Life cycle assessment -- Requirements and guidelines.

Kellenberger, D. and H.-J. Althaus. 2009. Relevance of simplifications in LCA of building components. *Building and Environment* 44(4): 818–825.

Khasreen, M.M., P.F.G. Banfill, and G.F. Menzies. 2009. Life-Cycle Assessment and the Environmental Impact of Buildings: A Review. *Sustainability* 1(3): 674–701.

Koch, A.L. 1966. The logarithm in biology 1. Mechanisms generating the log-normal distribution exactly. *Journal of Theoretical Biology* 12(2): 276–290.

Lee, Y.J. 2013. Streamlined carbon footprint computation : case studies in the food industry. Master's thesis, Massachusetts Institute of Technology, Cambridge (USA).

Lewandowska, A., A. Noskowiak, G. Pajchrowski, and J. Zarebska. 2015. Between full LCA and energy certification methodology—a comparison of six methodological variants of buildings environmental assessment. *The International Journal of Life Cycle Assessment* 20(1): 9–22.

Loh, W.-Y. 2002. Regression trees with unbiased variable selection and interaction detection. *Statistica Sinica* 12(2): 361–386.

Loh, W.-Y., Loh, and Wei-Yin. 2008. Classification and Regression Tree Methods. In *Encyclopedia of Statistics in Quality and Reliability*. Chichester, UK: John Wiley & Sons, Ltd, March 15.

Malin, N. 2005. Life Cycle Assessment for Whole Buildings: Seeking the Holy Grail. *Building Design & Construction*.

Malmqvist, T., M. Glaumann, S. Scarpellini, I. Zabalza, A. Aranda, E. Llera, and S. Díaz. 2011. Life cycle assessment in buildings: The ENSLIC simplified method and guidelines. *Energy* 36(4): 1900–1907.

Motuzienė, V., A. Rogoža, V. Lapinskienė, and T. Vilutienė. 2016. Construction solutions for energy efficient single-family house based on its life cycle multi-criteria analysis: a case study. *Journal of Cleaner Production* 112: 532–541.

Muller, S., P. Lesage, and R. Samson. 2016. Giving a scientific basis for uncertainty factors used in global life cycle inventory databases: an algorithm to update factors using new information. *International Journal of Life Cycle Assessment* 21(8): 1185–1196.

Noshadravan, A., M. Wildnauer, J. Gregory, and R. Kirchain. 2013. Comparative pavement life cycle assessment with parameter uncertainty. *Transportation Research Part D: Transport and Environment* 25: 131–138.

NREL. 2014a. NREL: U.S. Life Cycle Inventory Database Home Page. http://www.nrel.gov/lci/. Accessed April 23, 2015.

NREL. 2014b. NREL: Building Energy Optimization. https://beopt.nrel.gov/. Accessed April 23, 2015.

Olinzock, M.A., A.E. Landis, C.L. Saunders, W.O. Collinge, A.K. Jones, L.A. Schaefer, and M.M. Bilec. 2015. Life cycle assessment use in the North American building community: summary of findings from a 2011/2012 survey. *The International Journal of Life Cycle Assessment* 20(3): 318–331.

Olivetti, E., S. Patanavanich, and R. Kirchain. 2013. Exploring the viability of probabilistic under-specification to streamline life cycle assessment. *Environmental Science & Technology* 47(10): 5208–16.

Ortiz, O., F. Castells, and G. Sonnemann. 2009. Sustainability in the construction industry: A review of recent developments based on LCA. *Construction and Building Materials*

23(1): 28–39.

Patanavanich, S. 2011. Exploring the Viability of Probablistic Underspecification as a Viable Streamlining Method for LCA. Master's thesis, Massachusetts Institute of Technology, Cambridge (USA).

Pelton, R.E.O. and T.M. Smith. 2015. Hotspot Scenario Analysis. *Journal of Industrial Ecology* 19(3): 427–440.

Pesonen, H.-L. and S. Horn. 2012. Evaluating the Sustainability SWOT as a streamlined tool for life cycle sustainability assessment. *The International Journal of Life Cycle Assessment* 18(9): 1780–1792.

Pizzol, M., P. Christensen, J. Schmidt, and M. Thomsen. 2011. Impacts of "metals" on human health: a comparison between nine different methodologies for Life Cycle Impact Assessment (LCIA). *Journal of Cleaner Production* 19: 646–656.

Reis, L. 2013. An exploration of materials taxonomies to support streamlined life cycle assessment. Master's thesis, Massachusetts Institute of Technology, Cambridge (USA).

Rousseeuw, P.J. and C. Croux. 2012. Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association* 88(424): 1273–1283.

RSMeans. 2013. *RSMeans Green Building Cost Data 2013*. Rockland (USA): Construction Publishers & Consultants.

Schulz, M., M.D. Short, and G.M. Peters. 2012. A streamlined sustainability assessment tool for improved decision making in the urban water industry. *Integrated Environmental*

Assessment and Management 8(1): 183–93.

Schwartz, Y. and R. Raslan. 2013. Variations in results of building energy simulation tools, and their impact on BREEAM and LEED ratings: A case study. *Energy and Buildings* 62: 350–359.

Seo, S., S. Tucker, and P. Newton. 2007. Automated Material Selection and Environmental Assessment in the Context of 3D Building Modelling. *Journal of Green Building* 2(2): 51–61.

Soete, W. De, S. Debaveye, S. De Meester, G. Van der Vorst, W. Aelterman, B. Heirman, P. Cappuyns, and J. Dewulf. 2014. Environmental Sustainability Assessments of Pharmaceuticals: An Emerging Need for Simplification in Life Cycle Assessments. *Environmental Science & Technology* 48(20): 12247–12255.

Soust-Verdaguer, B., C. Llatas, and A. García-Martínez. 2016. Simplification in life cycle assessment of single-family houses: A review of recent developments. *Building and Environment* 103: 215–227.

Soust-Verdaguer, B., C. Llatas, and A. García-Martínez. 2017. Critical review of bim-based LCA method to buildings. *Energy and Buildings* 136: 110–120.

Steubing, B., C. Mutel, F. Suter, and S. Hellweg. 2016. Streamlining scenario analysis and optimization of key choices in value chains using a modular LCA approach. *The International Journal of Life Cycle Assessment* 21(4): 510–522.

Tecchio, P., P. Freni, B. De Benedetti, and F. Fenouillot. 2016. Ex-ante life cycle assessment approach developed for a case study on bio-based polybutylene succinate. *Journal of Cleaner Production* 112: 316–325.

The American Institute of Architects. 2010. *Architectural Graphic Standards for Residential Construction, 2nd Edition*. *American Institute of Architects*. Hoboken (USA): John Wiley and Sons, Inc.

Thinkstep. 2014. GaBi Software Professional Database. http://www.gabi-software.com/databases/professional/. Accessed April 23, 2015.

Upton, G. and I. Cook. 2008. *A Dictionary of Statistics*. Third Edit. Oxford (UK): Oxford University Press.

Wang, W., R. Zmeureanu, and H. Rivard. 2005. Applying multi-objective genetic algorithms in green building design optimization. *Building and Environment* 40(11): 1512–1525.

Weidema, B.P., C. Bauer, R. Hischier, C.L. Mutel, T. Nemecek, J. Reinhard, C.O. Vadenbo, and G. Wernet. 2013. *Data quality guidelines for the ecoinvent database version 3: Overview and methdology (final)*. St. Gallen (CH): the ecoinvent centre.

Weidema, B.P. and M.S. Wesnaes. 1996. Data quality management for life cycle inventories—an example of using data quality indicators. *Journal of Cleaner Production* 4(3–4): 167–174.

Zabalza Bribián, I., A. Aranda Usón, and S. Scarpellini. 2009. Life cycle assessment in buildings: State-of-the-art and simplified LCA methodology as a complement for building certification. *Building and Environment* 44(12): 2510–2520.