Parametric Archetype: A Synthetic Digital Method of Building Material Stock Representation Based on Distance Measurement

Author(s):
Pei, Wanyu; Stouffs, Rudi

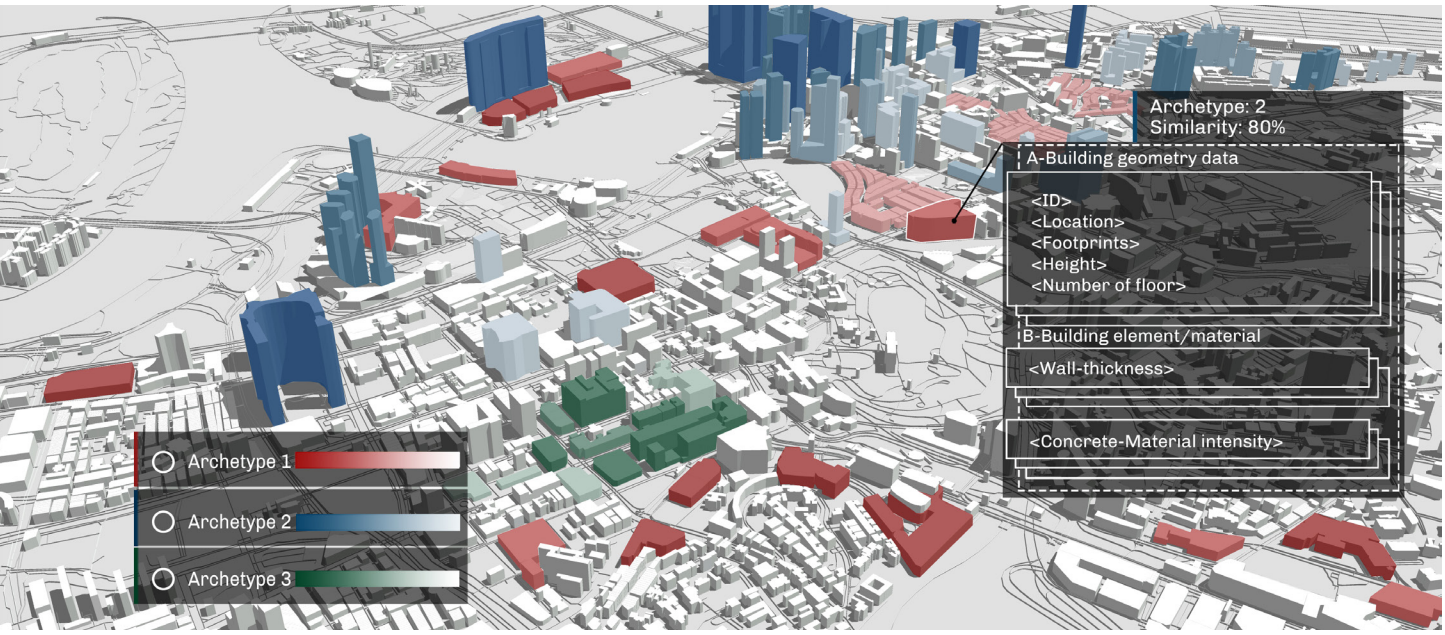
Publication date:
2023-10-21

Permanent link:
https://doi.org/https://doi.org/10.3929/ethz-b-000653906

Parametric Archetype

Wanyu Pei
National University of Singapore
Singapore-ETH Centre
Rudi Stouffs
National University of Singapore

A Synthetic Digital Method of Building Material Stock Representation Based on Distance Measurement



1

ABSTRACT

Building material stock (BMS) is a crucial inventory of secondary resources which contain comprehensive information for analyzing the potential of material reuse and urban harvesting. Due to the complexity of urban building systems and the large number of buildings, obtaining building information one by one is impractical. Existing methods for stock representation mainly start from data collection, and utilize techniques such as clustering, machine learning, computer vision, et cetera, to process and analyze large and complete datasets. However, it is noticed that data on urban buildings, especially for building materials, is very limited or rather inaccessible. Existing methods cannot be applied in data-scarce cities and are also challenging to update over time. Therefore, this study proposes a synthetic approach named *parametric archetype* for the digital representation of BMS. This approach combines distance measurement, which is a distance within dimensions describing building features, to match instance buildings dynamically to a *parametric archetype* with the highest similarity. The weight and types of different building features, which may influence building material (composition and properties) in distance measurement, can be determined by supervised, semi-supervised, or unsupervised learning, whether relying on ample available data or domain rules/expert knowledge when data is scarce. This way, the *parametric archetype* model can use data more efficiently to form a synthetic and extensible representation for urban-level BMS (Figure 1). The *parametric archetype* is anticipated to offer an approach for describing, quantifying, and modeling the real building material stock system incrementally and transparently.

1 The representation of Building Material Stock (BMS).

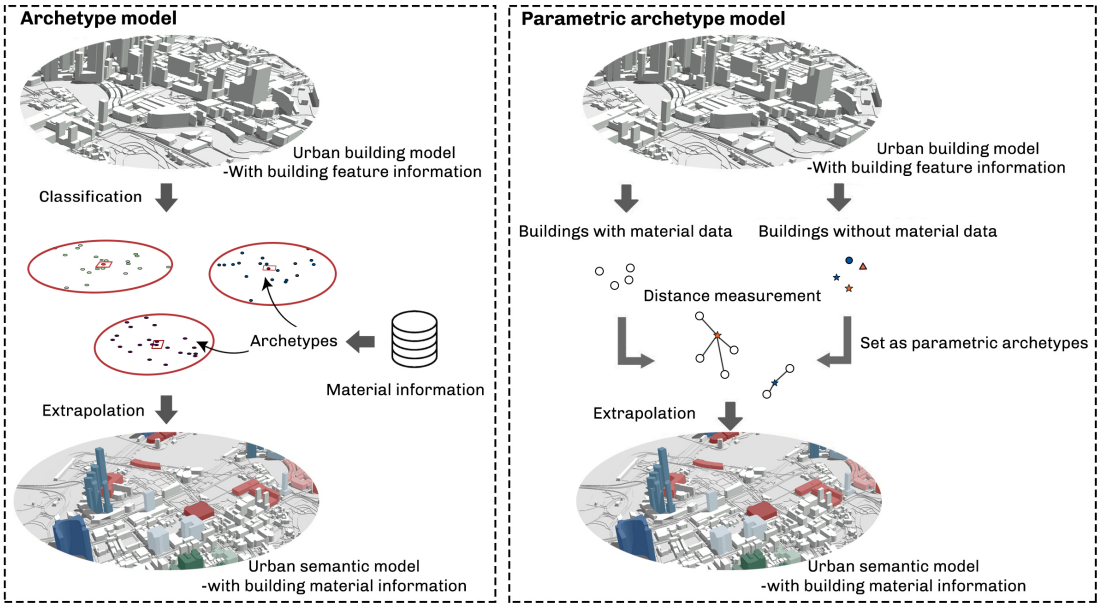
INTRODUCTION

Building material stock (BMS) is an essential resource in a city that can be reused and repurposed for future construction activities to reduce waste and promote sustainability (Kralj and Markič 2008). To effectively utilize BMS as secondary resources, comprehensive information and knowledge pertaining to the accumulated building materials are essential (Morlet et al. 2016). For example, building material stock analysis (MSA) and material flow analysis (MFA) can be adopted to track the composition, quantity, and location of existing building materials within an identified area (Mohammadiziazi and Bilec 2022). The primary step in conducting MSA and MFA is developing and characterizing a representation of the material stock and characterizing it. However, because of the complex nature of the urban building systems, the information on BMS is often scarce and disparate (Honic et al. 2023), making it challenging to account for and represent BMS.

Previous research studies applied "top-down," "bottom-up," and "demand-driven" approaches for BMS representation (Usman et al. 2021). The "top-down" approach involves obtaining data from a broader perspective, such as obtaining remote sensing or statistical data at a large level (Marinova 2020). The "demand-driven" approach utilizes socio-economic indicators, such as economic, population, and social factors, to model the demand for construction materials (Fishman et al. 2015). Compared to these two approaches, the "bottom-up" approach is more widespread, as it provides more detailed information and links stock dynamics with construction activities (Lanau et al. 2019). The basic "bottom-up" approach develops archetypes as representative buildings with the same building features, and extends the archetypes' information to the stock level (Mastrucci et al. 2015).

Recently, the archetype approach tried to incorporate techniques such as geographical information systems (GIS), machine learning (ML), mobile sensing, computer vision, clustering, et cetera, to enhance the accuracy and performance of representation results across various aspects and steps (refer to the "related work" section). However, there are limitations to these methods. Firstly, they rely on comprehensive and voluminous data sets, making data collection a crucial starting point. However, certain regions and cities may face challenges such as data scarcity, inaccessibility, or incompleteness (particularly for material data), due to privacy protection measures, policies, regulations, economic interests, or technological barriers. In such cases, conventional methods are not applicable. Secondly, data collection involves extracting data from diverse sources and storing all relevant information for analysis, which can be time-consuming and entail significant economic/human costs. Thirdly, the accuracy of BMS representation is contingent upon the quality of the data, which can be inexplicable and immeasurable. Moreover, updating the data and results is challenging as the arrival of new values necessitates a complete recalculation.

To address these issues, this study proposes a novel representation model, utilizing a distance measurement for similarity evaluation, providing a more flexible and incremental method for calculating and characterizing BMS. This model extends the existing archetype approach



2 Comparison of the archetype and *parametric archetype* model.

2

by setting buildings with complete attributes and material data as *parametric archetypes*. The *Euclidean distance* serves to quantify the similarity between building material use (related to types and properties), and is achieved by computing the distance between the attributes of the building (such as building age, function, height, shape, et cetera). This measures the similarity of material used between instance buildings (with attribute but no material data) and the *parametric archetypes* so that the most similar material data can be matched to instance buildings.

The objectives of this paper are to: 1) define the *parametric archetype* for BMS representation; 2) develop a distance measurement model for the *parametric archetype*; 3) evaluate the distance measurement model through a case study; and 4) discuss the potential of applying the model more generally. This paper is organized into six sections: The "related work" section reviews the existing methods for BMS representation; The "parametric archetype and distance measurement" section defines the concept and computational rules of *parametric archetypes* and distance measurement; The "case study" section tests the distance measurement model; and the "disscussion" section discusses the results, model limitations, and potential application. The conclusion is provided in the last section.

RELATED WORKS

Representing the BMS at the urban level is challenging due to limited building materials information, for which simplifying the complex and diverse BMS into a smaller set of representative types (archetypes) is widely adopted. The basic approach is conducting "expert-based" classification (Mollaei et al. 2021; Mastrucci et al. 2017), which groups buildings based on their physical features or other relevant variables (building age, function, height, et cetera) according to the expertise of professionals in the building industry. The "expert-based" classification can be seen as an unsupervised learning process that is particularly useful when data is limited or incomplete because no material information is required for preliminary classification. The experts only need to collect the real material data or make assumptions to identify geometry and material parameters, for these archetypes which stand for distinct groups. However, it may be subject to bias, and not always reflect the true diversity of the building stock. Hence, some research introduces data-driven approaches such as ML (Ali et al. 2019) or cluster technique (Tardioli et al. 2018) to conduct supervised learning, using material data to improve the accuracy of representation results.

Previous data-driven research has enhanced the BMS representation in the following aspects: 1) building features extraction: using cluster technique to extract the most representative and useful variables of buildings and remove those that do not influence the development of the archetypes (Yuan et al. 2023); 2) Data extension: Alaa et al. (2020) apply neural networks and random forests to identify the missing features (e.g., number of floors and construction periods) from building shape; 3) Building classification: Giovanni et al. (2018) use the cluster technique and form a predictive model to identify the building groups and archetypes based on a large dataset; 4) BMS quantification: Yuan et al. (2023) adopted ML regression to predict the BMS based on six easy-to-obtain building features and ground truth data of 71 buildings.

In summary, existing studies conduct BMS representation through: 1) an unsupervised learning method to identify the building features and archetypes by classification/clustering based on expert knowledge. Then, material data of each archetype is required to estimate the urban-level material stock; 2) A supervised learning method, a data-driven process, to quantify (predict) the BMS using clustering analysis and ML, which requires complete data set on building features and material to train the ML algorithm. A potential research topic with promise is developing a synthetic method that can select/integrate expert knowledge or/with data-driven approaches, adapting to BMS representation under different situations with varying amounts of data.

PARAMETRIC ARCHETYPE AND DISTANCE MEASUREMENT

Comparison of Archetype and *Parametric Archetype*

Extending upon the existing archetype method, this research proposes a novel *parametric archetype* (Figure 2) model. The "*parametric*" is included in this nomenclature to indicate that this approach utilizes the basic building geometric features as measurable parameters for describing the building material stock information. The existing archetype approach requires material data for all representative buildings (archetypes) that are decided by building stock segmentation rules. Small numbers of archetypes will result in poor representation as the variability within each category is difficult to capture, while large numbers make data collection challenging. Hence, the number of archetypes should balance the representative ability and the number of material data required. For the *parametric archetype* approach, the number of *parametric archetypes* has randomness, variability, and increment, as all buildings with complete features and material data can be set as *parametric archetypes*. A feature-based distance measurement is designed to quantify the similarity between the dimensions which describe the object buildings'

features related to material use. The similarity of material used between these archetypes and other instance buildings can be quantified through distance measurement. This way, instance, buildings lacking material information can be mapped, respectively, with a *parametric archetype* with the most similar material data until all buildings have material information, and a building-by-building BMS representation is achieved.

Determining the Building Features

Identifying the building features which are representative and related to material use is the core part of the distance measurement that can influence the accuracy of representation. The selection of these parameters mainly considers two factors: 1) directly related to building material use (material properties and composition); and 2) are easily accessible. These parameters are identified according to existing research (Mohammadiziazi and Bilec 2022) and professional knowledge, including building age, building type, building height, building footprint shape (ratio k), and footprint area.

- ***Building age:*** This is a prevalently utilized parameter to classify a building stock into archetypes in existing research, which provides valuable information about the construction methods and materials used.

- ***Building type:*** As per the requirements of most codes and standards, buildings' architectural and structural designs are determined based on their function, directly affecting the composition and the amount of materials. Building age and building type are widely used parameters for building classification in BMS characteristics.

- ***Building height:*** This parameter influences the quantity and composition of materials (Mohammadiziazi and Bilec 2022). A straightforward example is the different material compositions of low-rise and high-rise building structural systems. Besides, multiplying Material Intensity (MI) by building floor area (m^2) or building volume (m^3) is a widely used approach to quantity BMS. Lower-story buildings usually have lower MI (kg/m^2 or kg/m^3) values. In contrast, the lower floors of a high-rise building have high MI because they require additional material to support loads of stronger structural systems.

- ***Building footprint shape:*** The plan shape determines the wall/floor (or enclosing ratio), impacting the cost of materials for external walls and windows. It can be described using the ratio k of building footprint perimeter ($P_{footprint}$) and footprint area ($A_{footprint}$) (Mac-Barango 2012) (Equation 1). Generally, the smaller the ratio k, the less material is required for external walls. It is related to building shape because the k value of "circular," "square," "rectangular," "T-shaped," and "L-shaped" footprint areas increase in order.

$$k = \frac{P_{footprint}}{A_{footprint}} \qquad \text{(Equation 1)}$$

- ***Building footprint area:*** For instance, two industrial buildings with identical shapes, but differing footprint areas, would require different structural systems and material compositions.

Distance Measurement

For distance measurement, this research utilizes the *Euclidean Distance* method, treating buildings as points in Euclidean Space with identified parameters as several dimensions. Assume we have N buildings (with complete features and materials data) to be set as *parametric archetypes*, i.e., $D_A = \{a_1, ..., a_N\}$, where $a_i \in \mathbb{R}^{M+4}$. Specifically, we define a_i using Equation 2:

$$a_i = [t_i, g_i, h_i, f_i, k_i] \qquad \text{(Equation 2)}$$

where $t_i \in \mathbb{R}^M$, which uses a series of zeros and a single one (or several ones for mixed-use buildings) to identify the exact building type, $g_i \in \mathbb{R}$, $h_i \in \mathbb{R}$, $f_i \in \mathbb{R}$, and $k_i \in \mathbb{R}$ denote the one-hot building type encoding vector, building age, height, footprint area, and ratio k, respectively. M refers to the total number of building types. Given a building instance $b \in \mathbb{R}^{M+4}$ (without material data) which is also encoded in the same way as described in Equation 3, we can define the *Euclidean Distance* $d_i \in \mathbb{R}$ between b and the *parametric archetype* a_i:

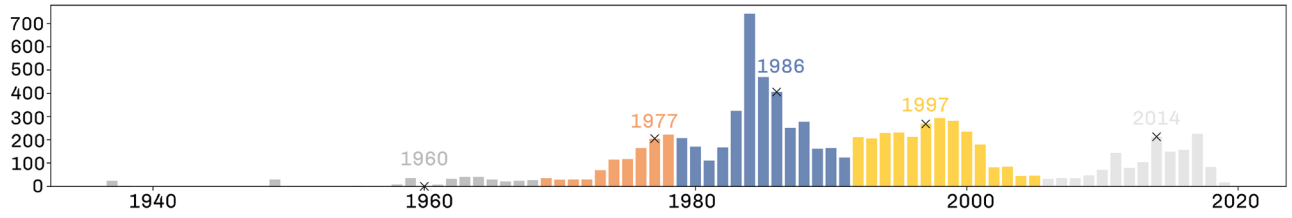
$$d_i = \sqrt{\sum_{j=1}^{m+4} \left(b_j - a_{ij}\right)^2} \qquad \text{(Equation 3)}$$

where $b_j \in \mathbb{R}$ denotes the j-th parameter of b and $a_{ij} \in \mathbb{R}$ represents the j-th parameter of a_i. Considering that the contribution of each feature parameter is often different, we further define a weighted *Euclidean Distance* (Equation 4):

$$d_i^* = \sqrt{\sum_{j=1}^{m+4} w_j \left(b_j - a_{ij}\right)^2} \qquad \text{(Equation 4)}$$

where w_j denotes the weight of the j-th feature parameter. To identify w_j, we define the relationship between the building material m and the five featured parameters in Equation 5 as:

$$m = F(t_i, g_i, h_i, f_i, k_i) \qquad \text{(Equation 5)}$$



4

4 Choose centroids based on clustering result of HDB buildings according to building age.

In practice, we can instantiate F to obtain the contribution w_j. For example, if the relation F is linear, we can represent Equation 6 as:

$$m = [w_1, \cdots, w_m] t_i^\top + w_{M+1} g_i + w_{M+2} h_i + w_{M+3} f_i + w_{M+4} k_i + b' = w^{diag} a_a^\top + b' \quad \text{(Equation 6)}$$

where $w^{diag} = [w_1, \ldots, w_{M+4}]$ is a vector including all the diagonal parameters and b' is the bias parameter. The F can be identified based on expert knowledge or domain rules first, when no building material data is available.

Material Information Mapping

Based on the *Euclidean Distance* between building b and each *parametric archetype* in D_A, using Equation 7, we can find the most similar *parametric archetype* with b:

$$y_b \in \{1, \cdots N\}, d_{y_b} = \frac{min}{i \in \{1, \cdots N\}} \{d_i^*\} \quad \text{(Equation 7)}$$

where y_b denotes the assigned label to b. For example, y_b = 2 if the 2nd *parametric archetype* is assigned to b. In this way, the material data of the 2nd *parametric archetype* can be mapped to building b.

Optimization Potential and Incremental Nature of Model

Two aspects influence this model and representation result: (1) what the features are; and (2) the weight of these features. They can be identified according to expert knowledge (e.g., domain expertise and existing research) when there are limited buildings with material data m for BMS modeling. When more material data m becomes available, these features and their weight can be checked and revised manually or automatically using the ML algorithm. This way, the model and the representation result could be optimized continually with increasing material data volume.

With limited building data, the existing *parametric archetypes* may be too few to represent all the given buildings to some degree of accuracy. Therefore, we propose progressively updating the *parametric archetype* set D_A to improve its representative ability. Concretely, considering a new dataset for building b, we first calculate the minimum distance d_{min}^* between a given building b and all the existing *parametric archetypes* by $d_{min}^* = \min_{i \in \{1, \cdots, N\}} \{d_i^*\}$. We can decide to update D_A by inserting b into it if d_{min}^* it is larger than some threshold λ. The update strategy can be formulated as:

$$D_A = \begin{cases} D_A \cup \{b\}, d_{min}^* > \lambda, \\ D_A, \end{cases} \quad \text{(Equation 8)}$$

CASE STUDY

This study utilized 1) an open dataset of Singapore Housing & Development Board (HDB) buildings (Biljecki 2020; HDB property information 2023) containing feature information; and 2) several building projects with both feature and concrete (CUI) stock information (Keung 2012). Our objective was to conduct a case study that tests the effectiveness of distance measurement (*Euclidean*) for evaluating the similarity between 16 instance buildings and five *parametric archetypes*.

Data Collection and Processing

Table 1 Data normalization.

No.	Building Age	BuildingHeight	Footprint Area	Ratio k	Building type			
					R_0	R_1	NR_0	NR_1
1	0.40	0.31	0.05	0.33	0	1	0	1
2	0.46	0.27	0.06	0.31	0	1	1	0
.....								
10081	0.62	0.02	0.06	0.30	1	0	0	1
10082	0.78	0.31	0.05	0.22	0	1	1	0

Table 2 Feature data of five parametric archetypes.

ID.	Block	Street	Building Age	Building Type	Height (m)	Footprint Area (m²)	Ratio k (m⁻¹)
a	10A	BOON TIONG RD	2014	R	113.3	637.70	0.33
b	10B	LOR 7 TOA PAYOH	1997	R	71.3	839.43	0.14
c	10	JLN RUMAH TINGGI	1986	R	12.5	844.18	0.20
d	2	BALESTIER RD	1977	NR	12.5	3135.85	0.15
e	12	MERPATI RD	1961	R	29.3	1337.24	0.17

Various data sources were combined in this research to gather all the feature information for HDB buildings, which was then normalized (Table 1). Building age and height are directly extracted from the open dataset, while the ratio k is calculated based on the area and perimeter of footprints using *Quantum Geographic Information System (QGIS)*. The building types are divided into residential (R), non-residential (NR), and mixed (M) buildings according to the data set. Following one-hot coding, the value range of building types differs from other categories. This research empirically chose a scaling factor of 0.2 to maintain the balance between the categories.

Choose Centroids

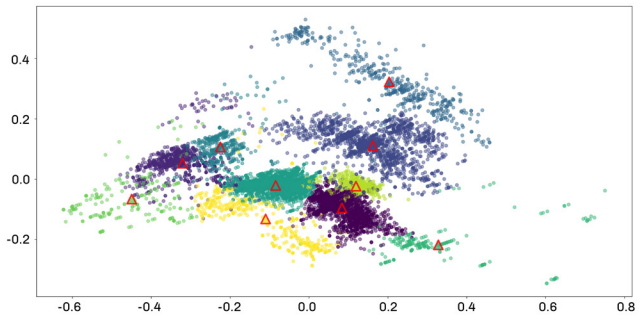
This research randomly selected centroid points as *parametric archetypes* for distance measurement by clustering HDB buildings based on their age. The number of clusters is five because the HDB buildings can be categorized into five periods, considering the evolution of structural systems and material composition. At each clustering center (or near the center) (Figure 4), a representative building was randomly chosen as the *parametric archetype* for the distance measurement (Table 2).

Distance Measurement Model

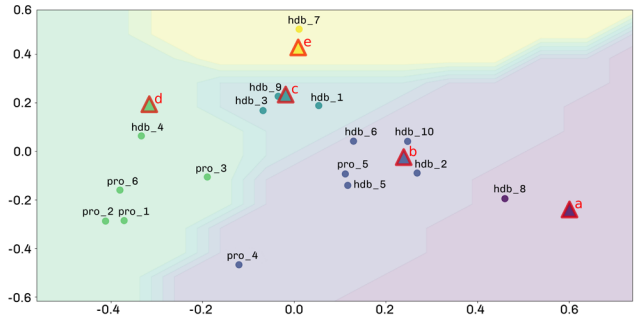
In this case study, the distance d_i between *parametric archetypes* and other instance buildings is calculated based on five equally weighted features. However, material information is currently unavailable for the five *parametric archetypes* and ten HDB building samples (see Appendice A). Only six building projects contain material information, which is the CUI defined by the Building and Construction Authority (BCA) to describe the amount of concrete used per unit floor area (Keung 2012). Ten HDB samples from the data set are chosen through clustering based on five parameters (Figure 5) to ensure the representation of various types.

The 16 sample buildings were individually measured for *Euclidean distance* d_i from 5 expert-defined *parametric archetypes*. Figure 6 illustrates distance measurements result, which can help examine building relationships from intra-cluster and inter-cluster perspectives:

• ***Intra-cluster:*** Buildings within the same cluster typically exhibit similar material properties, allowing for the extension of material information from centroids (*parametric archetypes*) to other buildings. For instance, the material properties of *parametric archetype* b can be mapped to *hdb_6*, which lacks material information, as they belong to the same cluster.



5



6

5 Selecting samples from HDB buildings through clustering.

6 Distance measurement result.

Available material data can be used to examine and adjust the parameters and their weight. If the actual data confirms similar material properties, it validates the reliability of the distance measurement. Conversely, different material properties would indicate the need to 1) remove or add certain parameters; and/or 2) modify the parameters' weights. For example, *parametric archetype* b and *hdb_6* have the same building age (1997) and type (R) but differ in height; the former is taller at 71.3m compared to the latter's 43.3m. While the footprint areas of b (839.43 m^2) and *hdb_6* (710.47 m^2) are similar, their ratio k values are different (0.14 for b and 0.22 for *hdb_6*). Thus, to address discrepancies between material properties obtained from the distance measurement model and actual material data, we could adjust the weights assigned to the building height and ratio k or consider removing these parameters.

• ***Inter-cluster:*** If the building samples belong to different clusters, it indicates that there should be significant variations in their material properties. For instance, *hdb_3* and *hdb_5* belong to separate clusters. Despite having similar building heights and functions, they differ in certain aspects. For instance, the footprint area for *hdb_3* is 1,593.62 m^2, while *hdb_5* has a footprint area of 1,356.64 m^2. Additionally, the ratio k is 0.17 for *hdb_3* and 0.15 for *hdb_5*, which are also similar. Another difference is their building age, with *hdb_3* constructed in 1983 and *hdb_5* in 2014. If the actual material data indicates insignificant

differences, adjusting the weight assigned to the building age parameter may be necessary.

Besides, *pro_1*, *pro_2*, *pro_3*, and *pro_6* belong to cluster (***d***), while *pro_5* and *pro_4* are in another cluster (***b***). *Pro_1*, *pro_2*, and *pro_3* share identical CUI values of 0.47, 0.46, and 0.46, indicating the same material intensity of concrete in their structural systems. However, *pro_5* in cluster ***b*** has a different CUI of 0.42 compared to the buildings in cluster ***d***. These materials' data align with the results from the distance measurement model.

However, some abnormal values exist. For instance, *pro_6* in cluster ***d*** shows a significant difference in CUI value (0.31) compared to *pro_1*, *pro_2*, and *pro_3*. The lower building height of *pro_6*, consisting of only three floors, may explain this deviation within cluster ***d***. Similarly, *pro_4* in the same cluster exhibits a distinct CUI value (0.46) compared to *pro_5*. This discrepancy could be attributed to notable differences in building footprint area and ratio ***k***. To address outliers, we can further adjust parameter weights and/or increase the number of centroids (*parameter archetypes*) to enhance the model's representative ability.

DISCUSSION

This study innovates existing archetype methods and proposes a novel approach for urban-level BMS representation with incremental nature to accommodate the material data limitation situation. Also, this study defines the concept of the *parametric archetype* and conducts a preliminary test of the distance measurement model through case studies. The test results indicate that models utilizing building feature parameters and *Euclidean distances* can effectively determine the similarity between building materials. Nevertheless, the current proposed model also has some limitations:

First, the types and weights of parameters in distance measurement are based entirely on expert knowledge, existing research, and domain rules, which are somewhat subjective. Thus, in the next stage, a number of real building cases with materials data ***m*** will be integrated to revise parameter types and weights, enhancing the accuracy of the representation model. As the amount of available building material information ***m*** increases, this correction process can be automated by designing algorithms. Secondly, the current case study uses CUI as the property related to concrete stock to verify material similarity is insufficient. The following step requires the development of a material library on BMS, which contains material types and related properties to enable a more accurate correlation between building parameters and material properties.

Finally, we assume that adding the number of *parametric archetypes* can improve the representation capability of the model. Thus, several issues remain undefined in this model at its current stage, including the relationship between accuracy and distance, distance threshold λ, the criteria for setting a new *parametric archetype*, and the optimal balance between the number of *parametric archetypes* and the model's representative capability.

CONCLUSION

This study proposes an innovative distance measurement-based approach named *parametric archetype* for urban-scale BMS representation. The *parametric archetype* starts by incorporating expert knowledge and developing mathematical models, allowing the gradually incorporating and updating of available data for material stock accounting. Moreover, the initial parameters and their weights, which are determined based on expert knowledge, can be manually or automatically corrected as additional material data becomes available. This correction process model offers greater flexibility. Unsupervised learning can be performed when only feature parameters are available without material data, and clustering analysis can be conducted based on expert-based parameters. The model can be updated iteratively as material data becomes available through continual learning. When sufficient building material data is available, supervised learning can be utilized in this model to more accurately investigate the relationship between building features and material properties. This way, when material data is unavailable, the distance measurement model can approximate the material based solely on building feature data. This method is expected to overcome the limitations of the conventional archetype approach by mitigating its dependence on exhaustive material data sets and addressing the difficulties associated with updates.

ACKNOWLEDGEMENTS

The research was conducted partly at the Future Cities Lab Global at the Singapore-ETH Centre, which was established collaboratively between ETH Zurich and the National Research Foundation (NRF) Singapore. This research is supported by the NRF Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. The financial support provided by China Scholarship Council (CSC) to the first author, Wanyu Pei, is acknowledged.

REFERENCES

Ali, Usman, Mohammad Haris Shamsi, Cathal Hoare, Eleni Mangina, James O'Donnell. 2019. "A data-driven approach for multi-scale building archetypes development." *Energy and Buildings* 202: 109364.

Biljecki, Filip. 2020. "Exploration of open data in Southeast Asia to generate 3D building models." In ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, VI-4/W1-2020: 37-44.

Fishman, Tomer, Heinz Schandl, and Hiroki Tanikawa. 2015. "The Socio-economic Drivers of Material Stock Accumulation in Japan's Prefectures". *Ecological Economics* 113: 76-84.

"HDB Property Information." Data.gov.sg. Accessed September 8, 2023. https://beta.data.gov.sg/datasets/150/view.

Honic, Meliha, Peter Ferschin, Dominik Breitfuss, Oliver Cencic, Georgios Gourlis, Iva Kovacic, and Catherine De Wolf. 2023. "Framework for the Assessment of the Existing Building Stock Through BIM and GIS." *Developments in the Built Environment* 13: 100110.

Keung, John. 2012. "SUSTAINABLE CONSTRUCTION: A Guide on CONCRETE USAGE INDEX." Building and Construction Authority.

Kralj, Davorin and Mirko Markič. 2008. "Building Materials Reuse and Recycle." *WSEAS Transactions on Environment and Development* 4(5): 409-418.

Lanau, Maud, Gang Liu, Ulrich Kral, Dominik Wiedenhofer, Elisabeth Keijzer, Chang Yu, and Christina Ehlert. 2019. "Taking Stock of Built Environment Stock Studies: Progress and Prospects." *Environmental Science & Technology* 53 (15): 8499-8515.

Mac-Barango, Dumo O. 2012. "An investigative study of the effect of design decisions on the cost of buildings." In West Africa Built Environment Research (Waber) Conference 24-26 July 2012 Abuja, Nigeria 2: 819.

Marinova, Sylvia, Sebastiaan Deetman, Ester van der Voet, Vassilis Daioglou. 2020. "Global Construction Materials Database and Stock Analysis of Residential Buildings between 1970-2050". *Journal of Cleaner Production* 247: 119146.

Mastrucci, Alessio, Antonino Marvuglia, Emil Popovici, Ulrich Leopold, and Enrico Benetto. 2017. "Geospatial characterization of building material stocks for the life cycle assessment of end-of-life scenarios at the urban scale." *Resources, Conservation and Recycling* 123: 54-66.

Mastrucci, Alessio, Emil Popovici, Antonino Marvuglia, Luís De Sousa, Enrico Benetto, and Ulrich Leopold. 2015. "GIS-based Life Cycle Assessment of Urban Building Stocks Retrofitting-A Bottom-up Framework Applied to Luxembourg." In EnviroInfo and ICT for Sustainability. 47-56. Atlantis Press.

Mohammadiziazi, Rezvan, and Melissa M Bilec. 2022. "Building Material Stock Analysis is Critical for Effective Circular Economy Strategies: A Comprehensive Review." *Environmental Research: Infrastructure and Sustainability* 2(3): 032001.

Mollaei, Aida, Nadine Ibrahim, and Komal Habib. 2021. "Estimating the construction material stocks in two Canadian cities: A case study of Kitchener and Waterloo." *Journal of cleaner production* 280: 124501.

Morlet, Andrew, Jocelyn Blériot, Rob Opsomer, Mats Linder, Anina Henggeler, Alix Bluhm, and Andrea Carrera. 2016. "Intelligent Assets: Unlocking the Circular Economy Potential." *Ellen MacArthur Foundation*: 1-25.

Tardioli, Giovanni, Ruth Kerrigan, Mike Oates, James O'Donnell, and Donal P. Finn. 2018. "Identification of representative buildings and building groups in urban datasets using a novel pre-processing, classification, clustering, and predictive modelling approach." *Building and Environment* 140: 90-106.

Yuan, Liang, Weisheng Lu, Fan Xue, and Maosu Li. 2023. "Building feature-based machine learning regression to quantify urban material stocks: A Hong Kong study." *Journal of Industrial Ecology* 27(1): 336-349.

IMAGE CREDITS

All drawings and images by the authors.

Wanyu Pei is a PhD candidate in the Department of Architecture, College of Design and Engineering, National University of Singapore. She is also a PhD researcher in the Circular Future Cities (CFC) project, Future Cities Lab (FCL) Global, Singapore-ETH Centre. Her research interests include Circular City, Building Material Stock Analysis (MSA), Domain Analysis and Semantic Modeling, and Urban Information Modeling.

Rudi Stouffs is Dean's Chair Associate Professor in the Department of Architecture and Assistant Dean (Research) in the College of Design and Engineering at National University of Singapore. He is Principal Investigator in the "Urban Climate Resilience" module in the Future Resilient Systems II research programme, and of the "Circular Future Cities" and "Sea-City Interface" modules in the FCL Global research programme (Singapore ETH Centre).

APPENDICES A

The information on sample buildings

ID.	Block	Street	Building Age	Building Type	Building Height	Footprint Area	Ratio k	CUI(kg/m²)	Distance Measurement				
									Shortest	2nd	3rd	4th	5th
hdb_1	403	JURONG WEST ST 42	1983	R	34.90	1620.80	0.17	NA	c	e	b	d	a
hdb_2	265A	PUNGGOL WAY	2013	R	48.90	682.92	0.31	NA	b	c	a	e	d
hdb_3	150	TAMPINES ST 12	1983	R	32.10	1593.62	0.17	NA	c	d	e	b	a
hdb_4	417	YISHUN AVE 11	1992	NR	6.90	2609.88	0.13	NA	d	c	e	b	a
hdb_5	315B	YISHUN AVE 9	2014	R	43.30	1356.64	0.15	NA	b	c	a	d	e
hdb_6	323	SEMBAWANG CL	1997	R	43.30	1269.27	0.19	NA	b	c	e	d	a
hdb_7	18	DAKOTA CRES	1958	R	20.90	710.47	0.22	NA	e	c	d	b	a
hdb_8	28D	DOVER CRES	2010	R	99.30	885.79	0.24	NA	a	b	c	e	d
hdb_9	43	LENGKOK BAHRU	1986	R	12.50	969.39	0.18	NA	c	e	d	b	a
hdb_10	10F	BEDOK STH AVE 2	1992	R	68.50	894.50	0.18	NA	b	c	e	a	d
pro_1	8	COLLEGE AVE WEST	2011	NR	22.16	7337.17	0.06	0.47	d	c	b	e	a
pro_2	63	TUAS SOUTH AVENUE 1	2012	NR	15.00	7698.24	0.06	0.46	d	c	b	e	a
pro_3	217	UPPER BUKIT TIMAH RD	2011	NR	20.03	776.67	0.17	0.46	d	c	b	e	a
pro_4	308B	PUNGGOL WALK	2014	R	50.16	14482.41	0.09	0.46	b	c	d	a	e
pro_5	97W	COAST DR	2013	R	36.00	754.32	0.18	0.42	b	c	a	d	e
pro_6	11	TAMPINES CONCOURSE	2009	NR	8.10	4184.74	0.07	0.31	d	c	b	e	a