THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Modeling Material Stocks and Embodied Emissions from the Built Environment

A machine learning and dynamic material stock and flow framework

Qiyu Liu

Department of Environmental and Energy Sciences

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2026

Modeling Material Stocks and Embodied Emissions from the Built Environment

A machine learning and dynamic material stock and flow framework

Qiyu Liu

ISBN: 978-91-8103-469-1

Acknowledgements, dedication, and similar personal statements in this thesis reflect the author's own views.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny Serie nr 5926
ISSN 0346-718X
https://doi.org/10.63959/chalmers.dt/5926

Department of Environmental and Energy Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone +46(0)31-772 1000

Printed by Chalmers Digitaltryck
Gothenburg, Sweden 2026

Modeling Material Stocks and Embodied Emissions from the Built Environment

A machine learning and dynamic material stock and flow framework

QIYU LIU

Department of Environmental and Energy Sciences

Chalmers University of Technology

Abstract

The building and construction sector is responsible for around 37% of global carbon emissions. It is therefore crucial to dicarbon the sector, and it is currently not on track to meet the Paris Agreement aligned decarbonization pathways. The carbon emissions from building and construction can be classified into operational emissions from direct energy consumption, and the embodied emissions associated with the material production process. Policy efforts have been mainly focused on reducing operational emissions through measures to reduce energy consumption by building renovation such as the Energy Performance of Buildings Directive (EPBD) in the European Union. However, prioritizing renovation could result in trade-offs such as increased embodied emissions.

It is therefore important to understand and model the embodied emissions from the building and construction sector, especially regarding maintenance and renovation. A key challenge in developing detailed bottom-up material stock and flow model for the built environment is the lack of data. This thesis addresses this challenge by developing a modeling framework that applies machine learning to predict and impute missing inventory data, and subsequently further develops material flow analysis models to better estimate embodied emissions from maintenance and renovation. This modeling framework is first developed using roads in Sweden as a case study and further developed and refined for residential buildings in Sweden. The machine learning approach is developed to circumvent the need for scarce height data to predict usable floor space and construction year for buildings. The building inventory dataset is then complemented by a detailed material intensity (MI) dataset that can be disaggregated into building layers, and a material flow model that ensures renovation is correlated with demolition is developed. Lastly, the modeling framework is integrated with a building energy model to analyze the total emissions reduction potential for energy efficiency renovation for the Swedish city of Uddevalla.

The results show that firstly machine learning can be effectively applied to predict inventory data in absence of height data, and such approaches should be tested in other geographical areas. The embodied emissions results show that in Sweden embodied emissions from maintaining roads and renovating residential buildings (both functional and energy efficiency renovations) is expected to exceed the embodied emissions from new construction. Therefore, more policy attention is needed to reduce embodied emissions from renovations.

Keywords: *Material stock, Material flow analysis, Machine learning, Embodied emissions*

List of publications

The thesis is based on the following appended papers, which are referred to in the text by their assigned Roman numerals:

I. Liu Q, Rootzén J, Johnsson F. Development of a machine learning model to improve estimates of material stock and embodied emissions of roads. *Cleaner Environmental Systems,* 2024, 14, 100211.

II. Liu Q, Lanau M, Rootzén J, Johnsson F. Predicting building age and floor space using feature-engineered 2D urban morphology. *Smart and Sustainable Built Environment, 2026,* 1-23.

III. Liu Q, Lanau M, Rootzén J, Cao Z, Johnsson F (2026). A Layered Dynamic Material Flow Framework for Modeling Building Renovations. *Accepted for publication at Journal of Industrial Ecology.*

IV. Liu Q, Lanau M, Somanath S, Rootzén J, Johnsson F. (2026). Dynamic assessment of the climate benefits of biobased insulation materials. *Manuscript to be submitted.*

Qiyu Liu is the principal author of all papers. Dr. Johan Rootzén contributed to method development, discussions, and editing Paper **I** to **IV**. Professor Filip Johnsson contributed to method development, discussions, and editing Paper **I** to **IV**. Assistant Professor Maud Lanau contributed to the conceptualization, method development, discussions, and editing Paper **II** to **IV**. Professor Zhi Cao contributed to the discussions and editing Paper **III**. Dr. Sanjay Somanath contributed to conceptualization, method development, discussions, and editing Paper **IV**.

Other publications not included in the thesis

- Tasseven U, Zachariadis T, Johnsson F, Savvidou G, Liu Q. Assessing material requirements, supply and circularity potentials for photovoltaic systems – The case of Cyprus. *Renewable Energy*, 2026, 256(f).

- Tasseven U, Zachariadis T, Taliotis C, Liu Q. Photovoltaics under constraints: Dynamic material flows, supply pressures and circularity in the Eastern Mediterranean and Middle East. *Resource, Conservation and Recycling*, 2026, 233, 108999.

- Savvidou G, Johnsson F, Ljunggren M, Liu Q, Tasseven U, Zachariadis T. Circular technology design and its potential influence on minor metals demand in wind and solar expansion. *Journal of Industrial Ecology,* 2026, 30, 589-605.

Acknowledgements

My deepest gratitude goes to my supervisors, Johan Rootzén, Filip Johnsson, and Maud Lanau. Thank you, Johan and Filip, for giving me this incredible opportunity to come to Sweden and conduct this exciting research, and of course all the discussions and support through out all five years. I am incredibly grateful and feel lucky that Maud agreed to become my supervisor later into my PhD process and has brought me so much knowledge with regards to Industrial Ecology as a field and buildings more specifically. I also would like to thank Ida Karlsson, who has done such an incredible job in her PhD and afterwards that forms a basis for the research in this thesis.

To my colleagues at Energy Technology, thank you for creating this friendly and supportive work environment that has made me feel very welcomed and supported. I have also made so many friends in the office that has become my second family in Sweden. Thank you so much to Anh, Farah, Henrik, Sina, and Tharun for always being there for me through all the highs and lows at work and in my personal life, this journey would not have been the same without you. I would also like to thank the Sustainability Group for creating such a welcoming and caring environment that I always look forward to attending each meeting. To the Energy System group, I have learned so much from all of you and I really appreciate it.

To my colleagues at the research area of Sustainable Built Environment, and especially at the material stock squad, I am happy I got to be part of the research group and get much more detailed knowledge about building and construction. Thank you to David, Emiline, Luca, Jonathan, Shuang, Sanjay, and Xin.

To my family, I am thankful for all the support you have given me and always trying to give me the best opportunities possible. Finally, to Alfrida, I would like to thank you for being by my side and supporting me through the highs and lows and always inspiring me to stay positive. I look forward to experiencing the future to the fullest with you.

Qiyu Liu

Göteborg, 10th of August 2026

Table of Contents

CHAPTER 1

Introduction

The global building and construction sector provides the foundation for economic activities and human wellbeing, but at the same time is responsible for around 37% of the global carbon emissions and approximately 50% of the global material extraction [1]. To achieve the goal set in the Paris Agreement, the building and construction sector needs to rapidly decarbonize [2]. However, the sector is currently not on track to meet Paris-aligned climate pathways [3].

Modelling the environmental impact of the material stock dynamics can thus help inform policy makers on how to design policies to reduce emissions [4]. Material stock and flow analysis (MSFA) has been developed over time as the most commonly used methodological framework to estimate and quantify these stock dynamics [5], [6]. The strength of the MSFA methodology that makes it useful for tracking resource flows and for informing policy making is its ability to capture the stock dynamics over time and space [7].

The first strength of MSFA that makes the model output useful for policy makers and decision makers is the temporal dynamics of the emissions, as most emission reduction measures in the built environment takes time to develop and diffuse [8]. Furthermore, national level scenario-based studies that assess the potential to reduce emissions (e.g., [9], [10]) and policy relevant studies that try to establish a carbon budget for the construction sector (e.g., [11], [12]) rely on building stock dynamics as the basis for emissions assessment. Therefore, the work included in this thesis is developed based on the rationale that detailed modeling of the material stock dynamics of the built environment is necessary to achieve decarbonization.

Another important strength of MSFA is the ability to capture the spatial dependence and dependence of material stock and flows [13]. In the beginning of the development of MSFA models, the predominant approach is the top-down approach that relies on statistical inflow data to calculate stock based on lifetime dynamics due to higher data availability, but the top-down approach does not capture the granular spatial details that is crucial for understanding material stock dynamics [14]. However, most bottom-up MSFA studies that accounts for the spatial aspect is usually limited to the urban level due to data availability [15].

MSFA studies that are intended for modeling emissions reduction scenarios are needed at the national level since regulatory measures are largely absent, especially for embodied emissions

[16]. The emissions from the building and construction sector are generally classified into two different components: operational emissions and embodied emissions [17]. Operational emissions refers to the emissions associated with the energy consumption in the use phase while embodied emissions refers to the emissions associated with the production and transport of the materials used in building the building or infrastructure stock [18]. The initial research question posed at the start of the research was focused on embodied emissions as it was much less studied compared to operational emissions, but as the research progressed the question shifted towards how to model maintenance and renovation at a national level.

Renovation of the building stock (to a much lesser extent for the maintenance of roads) represents the intersection of operational and embodied emissions as energy efficiency renovations reduce energy demand and operational emissions overtime while incurring significant upfront embodied emissions [19]. In MSFA studies, renovations are relatively understudied due to data constraints and most existing studies focus on the urban level (See **Paper III**). There is therefore a need to understand the magnitude and importance of embodied emissions from renovations to provide more complete pathways on how to decarbonize the building and construction sector instead of only focusing on new construction.

As part of the EU's efforts to decarbonize the built environment, the Energy Performance of Buildings Directive (EPBD) has been introduced in 2025, and as part of the directive each member state is required to submit a national renovation plan [20]. In Sweden, the preliminary report published states that Swedish Board of Housing, Building and Planning (Boverket in Swedish) plans to introduce limit value on energy demand per floor area in the future, but such limit values are not currently being considered [21]. In addition, there is no limit value for embodied carbon planned for energy efficiency renovation. The report also mentions the importance of modeling the development trajectory or the stock dynamics, but the methodology used in the report is based on renovation rates and archetypes taken from energy performance declarations that cannot capture embodied emissions trajectories from renovation.

In this context, there is a need to develop a methodological frame to estimate emissions from renovation and maintenance at the national level. As highlighted in a recent review study by Fu et al. [14], there is a need to combine MSFA that can track different materials at different scales of long time series with spatial analysis that can enable the utilization of detailed material intensity coefficients (such as "material passport" suggested by Lanau et al. [22]). Based on the knowledge synthesized before, there is a gap in both research and policy space for a model that can capture the stock dynamics in time and space from renovation activities.

1.1 Aim

The overarching aim of this thesis is to estimate the potential to reduce emissions from renovation and maintenance of the Swedish built environment, focusing on residential buildings and roads using bottom-up dynamic material stock and flow models. The modeling framework developed in this thesis addresses the challenge of lack of bottom-up data in MSFA modeling. This framework is first theorized and tested on roads and subsequently further refined for residential buildings, and residential buildings are thus the main focus of the thesis as it represents a much larger share of the stock and emissions. The main contribution of this work is to the development of methodology, while demonstrating the applicability of the methods in case studies. The aim is then addressed with the following research questions

RQI: How can machine learning models be leveraged to predict missing inventory data to facilitate bottom-up material stock modeling? (Addressed in **Paper I** and **II)**

RQII: What is the in-use material stock of road and residential buildings in Sweden? (Addressed in **Paper I** and **II)**

RQ III: How to better model the material flows from building renovations with detailed bottom-up material stock and material intensity data? (Addressed in **Paper III)**

RQ IV: What is the magnitude of embodied emissions from maintenance and renovation of Swedish roads and residential buildings? (Addressed in **Paper I and IV)**

RQ V: What is the potential of using alternative insulation materials to reduce emissions from energy efficiency renovations in Sweden?

1.2 Scope and contribution of the research

The work in this thesis is broadly situated in the field of industrial ecology and the subfield of material stock and flow modeling. The method development is done by borrowing, adapting and applying methods from other fields such as machine learning and building on existing dynamic material flow analysis (dMFA) models to synthesize a framework that is uniquely suitable for the Swedish context and data but could also be adapted to other countries or regions. **Figure 1** presents an overview of the four papers included in the thesis and their respective contributions to addressing the research questions.

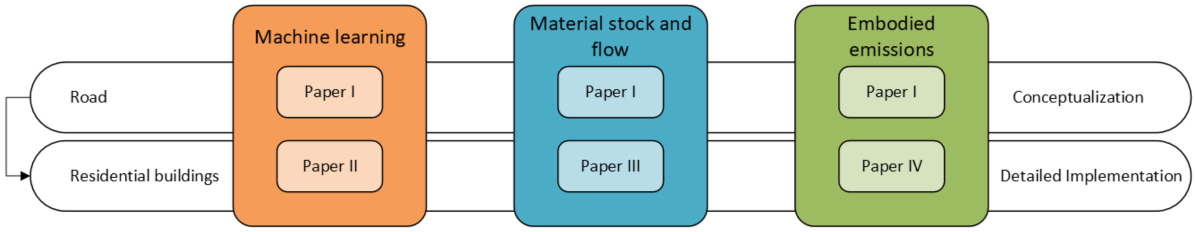

Figure 1. Overview of the three main topics shown in Chapter 5 – 7, and the papers which the results are taken from.

The modeling framework is first conceptualized and tested in **Paper I** by using roads as a case study. The framework is subsequently implemented in more depth and detail for residential buildings in Sweden due to the larger amount of data requirement for residential buildings and the more complex stock dynamics in Paper **II** to **IV**.

1.3 Outline of the thesis

This thesis is a combination of this introductory essay and the four appended papers. Following this introductory chapter, the background information on concepts in the thesis is explained in Chapter 2. Chapter 3 provides an explanation of the methodologies used in the thesis. Chapters 4, 5, and 6 show key results from the appended papers. Chapter 4 presents the training, testing, and application of machine learning models to predict missing inventory data. Chapter 5 presents the material stock and flow model results using the completed inventory data. Chapter 6 presents the embodied emissions results as well as the environmental impacts of biobased insulation materials compared to traditional insulation materials like mineral wool. Lastly, Chapter 7 discusses the results and provides some outlook for future work and Chapter 8 concludes the thesis. While the four appended papers model both roads and residential buildings, this thesis focuses on residential buildings.

The contribution of each appended paper is briefly summarized below.

Paper I present and applies the initial methodological framework of using machine learning models to predict missing inventory data as input for material stock and flow modeling. The framework is applied to model the material flow and embodied emissions from maintenance of Swedish road infrastructures. The machine learning model is used to predict missing road width data as road width is necessary to calculating the material stock of roads.

Paper II shifts the focus to residential buildings to tackle the challenge of incomplete bottom-up inventory data from the Swedish Land Survey. The main contribution of this work is to predict building usable floor space and construction year without height data as an input feature. Urban morphology features are shown to be predictive, and the resulting completed building inventory dataset is used as the basis for subsequent papers.

Paper III utilizes the building inventory dataset and calculates the material stock of residential buildings in Sweden using material intensity data based on the shearing layer concept. The paper further develops existing dynamic material flow methodology to explicitly link renovation lifetimes with structural lifetimes that tracks the temporal evolution of building cohorts based on building layers.

Paper IV expands the methodological framework developed in **Papers II** and **III** on a case study for the Swedish city of Uddevalla. The case study investigates the carbon benefit of biobased insulation materials by coupling the layer-based renovation model with building energy demand modeling using EnergyPlus. The analysis compares the emission reduction potential between biobased insulation materials and mineral wool by also accounting for biogenic carbon storage potential.

CHAPTER 2

Background

This chapter introduces and reviews key concepts in the thesis such as material stock modeling and material intensities, machine learning models and their application in MSFA studies, and building renovation modeling.

2.1 Built environment stock

Buildings and infrastructure contain large amounts of materials that remain in use for decades. As a result, they shape future demand for raw materials and determine how much waste and recyclable material becomes available over time. Therefore, the modeling of material stock is crucial to understand these dynamics and to better manage resource use and carbon emissions from the built environment [23]. In general, material stock models can be categorized into top-down or bottom-up approaches [24], [25]. Top-down models do not consider the differences between individual part of the stock and the approach relies on /aggregated data on material use (e.g., material trade, consumption statistics) and applies lifetime estimates to derive the quantities of material accumulated in the built environment stock [26]. In contrast, the bottom-up approach models a stock using “piece by piece” approach, e.g., at the building level. The challenge with the bottom-up approach is the requirement of substantial amounts of physical inventory data such as building floorspace as well as the need for detailed material intensity data [22].

The method to calculate material stock is a straightforward multiplication of material intensities (MI) data with inventory data. Both inventory and MI data are associated with challenges to obtain reliable data. An increasing amount of effort has been put into collecting and developing MI databases. At the international level, three datasets have been compiled: the open MI database by Heeren and Fishman [27], the global residential MI database by Marinova et al. [28], and the RASMI database by Fishman et al. [29]. For Sweden , Gontia et al. [30] have developed a material intensity database of residential buildings. However, the database does not provide detailed information on where different materials are located within building structures [30]. This thesis instead uses the newly developed BUD-MI (Bottom-up data: Material Intensity) database for Swedish residential buildings by Lanau et al. [31]. The main strength of the BUD-MI is that it allows MI data to be disaggregated across building layers, elements, and components. **Figure 2** illustrates how BUD-MI makes it possible to trace material use to specific parts of the building structure. In terms of roads, a similar challenge of lack of homogenized MI data is present [32]. The MI dataset used in the thesis is based on the emissions calculation tool (Klimatkalkyl) by the Swedish Transport Administration that has information on the material requirements for road infrastructures [33].

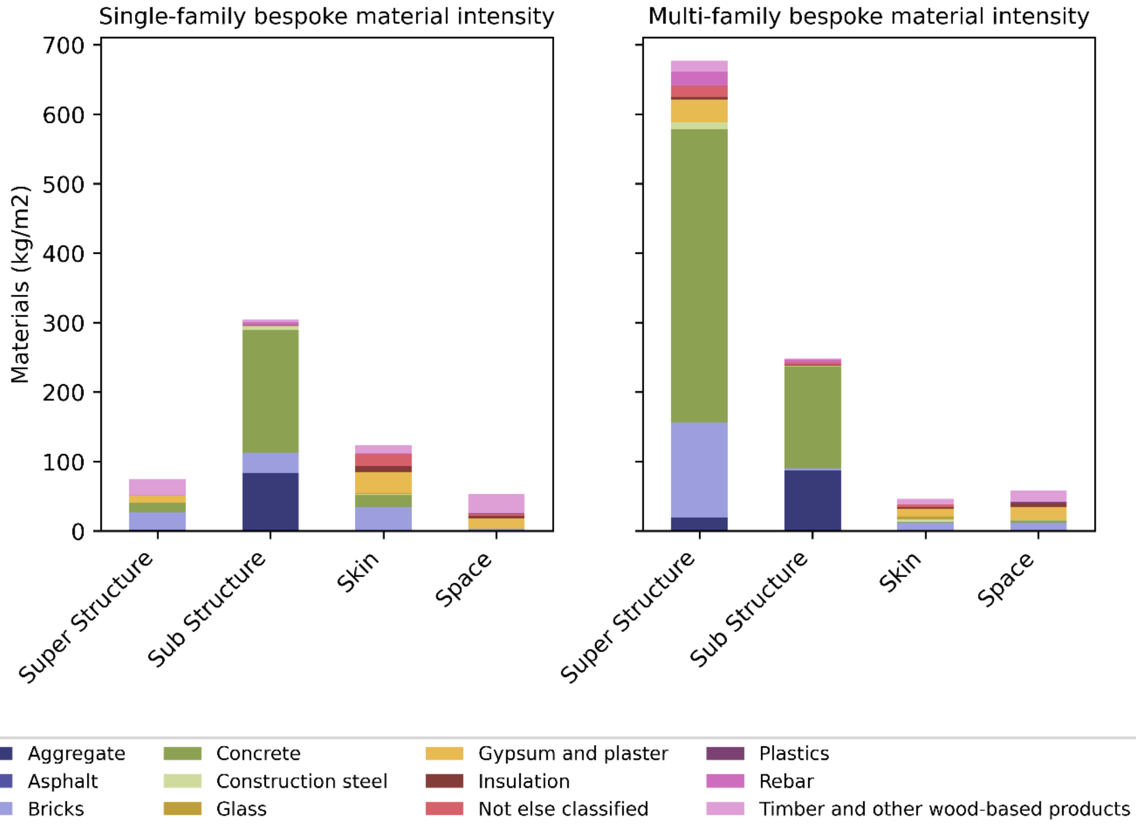

Figure 2. Bespoke material intensity based on BUD-MI for Swedish single-family (SF) and multi-family (MF) buildings.

The other key challenge for built-environment stock modeling is the lack of comprehensive inventory data. For buildings, building floor space dataset is hard to obtain and the diversity of building typologies and regional context further complicates the matter [34]. In addition, other attributes such as building footprints, heights, floor space, building type, and construction year are also important for stock modeling [35]. Open-source databases are fragmented, incomplete and official datasets can also often be hard to access. For example, a harmonized A review of European building inventory dataset using governmental data and OpenStreetMap has shown that for building height, construction year, and building type, only 73%, 24%, and 46%, respectively, of the data are available [35]. In the EU, only Spain, The Netherlands and France have national-level databases that contain building footprints and attributes [35]. Two tools have been developed by the EU to collect building inventory data, each with its own limitations. The Copernicus Reference Data Access (CORDA) node only contains data from selected countries and do not include Sweden [36]. In contrast, the EU Building Stock Observatory provides data at the country level for all of the EU Member States, but only in an aggregated fashion [37]. The Microsoft Corporation has developed an open-source building footprint dataset that uses a combination of neural networks and aerial images, but building height coverage in the dataset is not globally complete and is subject to methodological uncertainties [38].

For roads, the most important inventory data are the length, width, and archetype (a one lane road or a highway) [39]. The Global Roads Inventory Project (GRIP) has gathered and harmonized information related to road length and type of roads for 222 countries [14]. Rousseau et al. [5] have reported that in the GRIP dataset, more than 20% of the road length data is missing in many countries, while Central and South American and European countries have the lowest levels of missing data. Similarly, OpenStreetMap, which is another global-level dataset describing roads, suffers from data incompleteness [15]. In the European Union, many gaps exist in the official Eurostat statistics on transport infrastructure [16]. In addition, non-government-owned roads, such as municipal roads, have overall lower data quality and are sometimes not included in the national statistics [17].

2.2 Machine learning models

The recent advancement in machine learning algorithm and computational power means that machine learning has emerged as a promising tool to predict missing data and complete existing building inventory [34] and road inventory datasets [40]. Machine learning (ML) is the name given to algorithms that can learn from the attributes of an input dataset to generate a prediction based on the learning process (for an introduction to ML, see Hastie et al. [41]. To further simplify the concept, an ML model learns the patterns of the given training data using statistical methods and subsequently predicts the value of unseen data based on the learnt patterns.

Machine learning has been widely applied to studies within the field of Industrial Ecology in general and MSFA modeling since the start of the research for this thesis [42]. A trend that can be observed is that ML models can be applied to a wide range of research questions. However, the choice of method should be tailored to the specific research design and the availability of data [43]. This subsection briefly reviews the existing literature on applying ML to predict missing road and building inventory data.

For roads, machine learning has mainly been applied to predict depth of road layers. Ebrahimi et al. [10] estimated and predicted the material stocks and flows of the Norwegian road network by predicting the depth of roads with a decision tree-based ML algorithm. The strengths of this method are its abilities to incorporate the effect of traffic flows and to estimate the dissipative flows of materials. As the ML training process requires extensive data, the analyses were limited to national roads, for which all the input data were complete. Similarly, Wang et al. [30] estimated the material stocks and flows of road infrastructures in Belgium using a combination of ML models and the archetype-based approach. This work still requires road thickness data which is not widely available. These approaches can be seen as building upon and scaling up scarcely available data, but they do not address the lack of more fundamental inventory data for the purpose of estimating material stocks such as road width.

In comparison, building inventory data has received considerably more attention than roads [34]. In the context of building stock modeling, ML-based approaches are often used to predict building attributes such as age or height, to complete or enrich building inventory datasets. As shown in **Table 1**, existing work using ML-based approaches for building inventory development have been applied at various spatial scales. Most of the existing work was performed at the urban scale (neighborhood or city), with some studies carried out at the national scale [44] [45] [46].

Table 1. Comparisons of the different machine learning approaches used for building inventory modeling, sorted based on spatial scale.

Spatial scale	Geographic location	3D building geometry	Predicted attributes	Reference

Neighborhood	Vancouver, Canada	Yes	Age	[47]
Neighborhood	Trento and Bologna, Italy	Yes	Height	[48]
City	Rotterdam, The Netherlands	Yes	Age	[49]
City	Nottingham, UK	Yes	Age	[50]
City	Shenzhen, China	Yes	Age	[51]
City	North-Rhine Westphalia, Germany	Yes	Age	[52]
City	Rotterdam, The Netherlands	Yes	Height	[53]
City and National	The Netherlands, Germany, Italy, France	Partial	Height	[44]
National	Canada and USA	No	Height and floor space	[45]
National	The Netherlands, France, Spain	Yes	Age	[46]

The limitation with regards to spatial resolution is mainly due to the inherently high demand for data in ML-based approaches. This is especially true for studies that use 3D building geometry data. A 3D building dataset is usually derived using geoinformatics techniques or inspections of building plans, and depending on its level of detail, contains building information such as height and roof angle [54]. These 3D building data are informative for modeling a building stock, although official statistics with 3D data are rarely available.

As demonstrated previously ([45]), ML models are capable of providing good predictions of usable floor space in the absence of 3D building attribute data and with the benefit of a large geographic scope at the national level. However, Arehart and colleagues ([45]) have utilized satellite remote sensing imagery, which introduces additional uncertainty to the predictions and limits the generalizability of the approach. Furthermore, no previous studies have predicted building age at the national level without the usage of building height data. Nachtigall and coworkers [46] have predicted the age of residential buildings in The Netherlands, France and Spain using mainly urban form features with a mix of real height data and previously predicted height data from [44]. The results indicate that building and building neighborhood-level features have the greatest predictive power in terms of feature importance [46].

One of the most significant developments in the application of machine learning to building stock attributes since Paper II in this thesis was published has been the emergence of several large-scale building height data projects. For example, the World Settlement Footprint 3D dataset by Esch et al. [55], the 3D-GloBFP dataset by Che et al. [56], and the GlobalBuildingAtlas dataset by Zhu et al. [57]. These building height datasets could be used to estimate building usable floor space given 2D building footprints, but they do not address the issues of lack of building construction years. It is possible to use these height data as training data for ML workflows, but that carries the risk of error propagation.

2.3 Building renovation modeling

A common modeling assumption behind the simplification of renovation modeling is that buildings are represented as single, homogeneous units in dMFA studies. As reviewed by Mohammadiziazi & Bilec (2022), at the time of their writing, 80% of their reviewed studies did not distinguish between materials accumulated in structural and non-structural components. The exception being Stephan and Athanassiadis (2017, 2018) who quantified material stock and flow while differentiating between structural and non-structural materials without explicitly modeling renovation activities. If buildings are modeled as single units, the temporal dynamics between different components driven by renovation are not fully captured.

The temporal dynamics that is not captured is the differences in lifetimes for different building components [61]. Architectural theory and empirical evidence conceptualize buildings as assemblies of components with distinct functions, material compositions, and longevities [62].

Structural elements typically determine the end of a building's service life (EoL), whereas non-structural elements (such as façades, roofs, and interior partitions) are replaced repeatedly through renovations, without triggering demolition. The observed renovation frequencies vary significantly across these components, reflecting technological obsolescence, functional changes, and maintenance practices [63]. Capturing this heterogeneity is essential for the realistic modeling of renovation dynamics.

Despite this, most existing dMFA approaches experience limitations in simultaneously representing component-specific renovation dynamics and their dependencies on structural lifetime. Although the frameworks developed for other component-based products provide useful conceptual insights, they are not directly extrapolatable to buildings [64], where the replacement of non-structural components does not induce system-level obsolescence and demolition decisions are often driven by economic or planning considerations rather than material failure [65]. Building-specific approaches that introduce renovation cycles or renovation rates improve the temporal resolution but generally lack the ability to link renovation probability explicitly to demolition risk, while maintaining cohort-level detail for multiple building layers [66], [67]. There is therefore a need for dMFA models that can capture the dynamics of obsolescence between different building components.

The common assumption made in the modeling of building material stock is that buildings behave as a single product and thus have a single lifetime, and this assumption is in strong disagreement with the architectural concept of a building layer, which has become an increasingly central theme in the circular construction literature, and which states that "There isn't any such thing as a building. A building properly conceived is several layers of longevity of built components" (Duffy, quoted in Brand, 1995 [68]). Indeed, the vast majority of buildings are not monolithic; rather, they are composed of multiple parts, each with their own lifetime [68]. Brand [68] has proposed the following six layers for a building: site, structure, skin, space, services, and stuff (See **Figure 3**). When it comes to material stock (MS) modeling of the built environment, three of the six layers are particularly relevant: the structure; the skin (i.e., the building envelope); and the space. The latter refers to the building elements that subdivide the space within the building and that are not part of the structure, such as partition walls, ceilings, and flooring. Each layer has a specific lifetime, making their differentiation particularly critical when investigating the renovation dynamics of a building stock. Indeed, buildings typically undergo multiple renovations over their lifetime, with each renovation focusing on a specific layer [68]. For example, renovation of the façade of a building (skin layer) does not always happen at the same time as the changing of the interior layout of a building (space layer). Therefore, to capture accurately the complex dynamics of building renovation activities, it is essential to model the material stock and flow at a higher level of granularity, so that the different longevities of the layers are represented. Thus, dividing a building into layers can make building dMFA models more representative of real-world conditions.

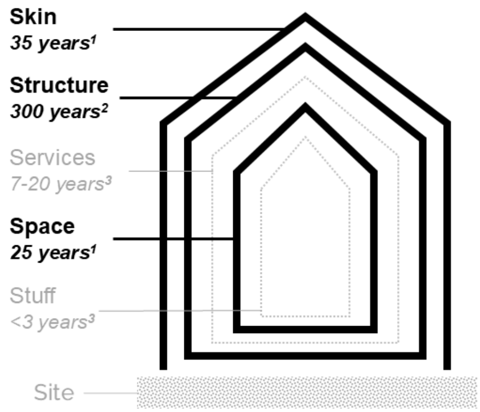

Figure 3. Demonstration of the shearing layers; adapted from [68]. 1. The skin and space lifetimes are derived from (Sartori et al., 2016). 2. The structure lifetime is assumed based on the empirical demolition rate in Sweden [70]. 3. The service and stuff lifetimes are taken from [71].

The concept of layers allows buildings to be treated as long-lived products composed of multiple layers. The product is the building, with the structure layer being the critical determinant of product obsolescence. In addition, there are two non-critical layers (skin and space), which can be renovated multiple times while the product remains in use. In addition, material intensities tend to differ by building construction age, which means that the materials required for the renovation of buildings in each cohort will differ. We have developed a model framework based on these unique characteristics of buildings functioning as group of layers instead of a single monolithic object.

CHAPTER 3

Case study and system boundary

This chapter briefly describes Sweden as a case study with a focus on the stock inventory data availability and the system boundary of the papers appended in the thesis. Sweden is chosen as a case study as the country has set climate goals to become carbon neutral by 2045 [72].

3.1 Road inventory data

This subsection describes the data sources for road and building inventory that are essential for modeling the material stocks and flows. The data source for road inventory data in Sweden is the Swedish Transport Administration (*Trafikverket* in Swedish). The road inventory data is retrieved from the data portal ‘Lastkajen’, which is an open-source platform for road and railway data in Sweden [73]. The data for all roads in Sweden are downloaded are processed, and the main inventory attributes extracted are the road lengths, widths, ownership type, archetypes (e.g., one-lane road or two-lane road), and the geographical location.

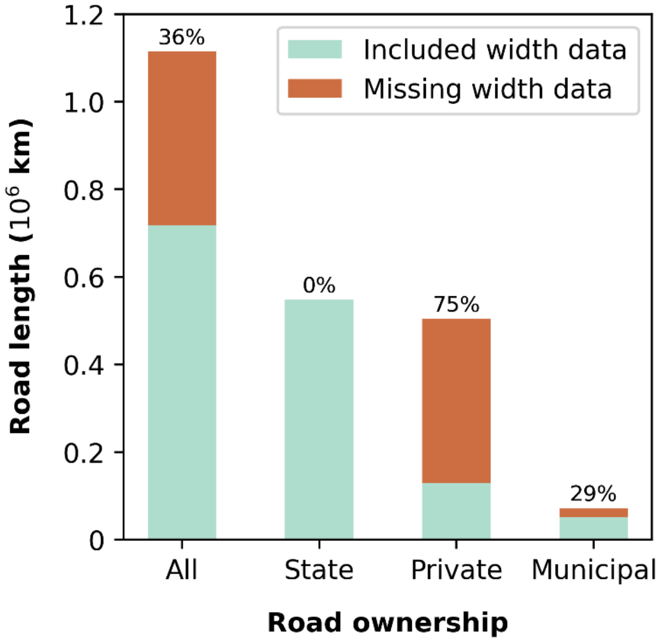

Figure 4. Lengths and percentages of roads with missing road width data according to ownership type, based on data from the Swedish Land Survey [47]. Source: **Paper I**.

The retrieved data shows that around 36% of roads are missing width data, which is essential for calculating the road material stock, as shown in **Figure 4**. Road width is an important attribute necessary to calculate the area of roads that can be multiplied with material intensity data to calculate road material stock and flow.

3.2 Building inventory data

The building inventory data is from the building registry of the Swedish Land Survey (*Lantmäteriet* in Swedish), retrieved in November 2022. This is the official database on land ownership that is used primarily for property tax assessment purposes. The registry contains approximately 8 million buildings, of which 2.8 million are residential buildings. The attributes obtained from the registry that are needed for material stock calculation include building type, construction year and usable floor space (here defined as the area of the building intended for occupancy, excluding basements and attics) [74]. Like the road dataset, many attributes in the building registry are incomplete. The key descriptive statistics of the dataset are summarized in **Table 2**.

Table 2. Key summary statistics for the dataset used in this study. The mean, standard deviation (std), 25% quantile (Q1), 75% quantile (Q3), and percentage of missing data for the usable floor space (m^2) and construction year of all residential buildings based on building type are shown.

		Usable floor space/building (m^2)					Construction year				
Building type	n	Mean	Std.	Q1	Q3	Missing	Mean	Std.	Q1	Q3	Missing
Detached houses	1,982,870	115.4	50.2	81	143	15%	1955	35	1930	1977	18%

Terraced houses	112,356	126.0	32.6	110	141	6%	1974	17	1969	1979	5%
Small houses	4,499	348.1	846.3	145	324	86%	1993	38	1977	2013	27%
Row houses	30,000	127.7	138.9	104	132	15%	1978	19	1970	1983	9%
Apartment buildings	50,389	1075.2	1927.6	292	1120	63%	1954	27	1936	1968	20%

The degree of incompleteness of the dataset with regards to construction year and usable floor space varies across building types. This is likely due to the facts that the data are collected and reported by individual municipalities and that there is a lack of national-level reporting requirements [75]. Statistics on the existing data show apartment buildings have significantly larger usable floor space per building than other building types. This is due to apartment buildings being registered as entire buildings, i.e., their usable floor area is the sum of the area of each apartment in the building. Apartment buildings and small houses exhibit the widest Interquartile Range (IQR) values of 828 and 179, respectively. They also have the highest proportion of missing data for usable floor space (63% and 86%, respectively) and construction year (20% and 27%, respectively). Therefore, machine learning models are used to predict the missing road widths and building attributes in **Paper I** and **II** respectively.

3.3 System boundary

This subsection provides an overview of the system boundaries of the papers appended in this thesis, the focus is on Sweden, but the methods are generalizable. For roads studied in **Paper I**, the types of roads include all paved roads and gravel roads, excluding bike lanes and pedestrian walkways. The road layers included in the material intensities are surface course, binder course, base, sub-base embankment (as defined in Lanau et al. [76]). We do not distinguish between above ground and underground in this study. The materials included in the MI are asphalt, steel, and aggregates. Concrete roads are excluded from the system boundary, as there are only 68 km of concrete roads in Sweden as of the year 2022 [63].

For buildings studied in **Paper III**, the scope of study includes all single-family (SF) buildings and multi-family (MF) buildings. The SF buildings are assumed to be one structure type across 12 age cohorts (each cohort is 10 years, from 1880 to 2000). The predominant structure type of SF buildings is assumed to be timber for all age cohorts. The MF buildings are assumed to include four structure types, and 13 age cohorts (each cohort is 10 years, from 1880 to 2010). The structure types include Wood multi-family (WMF) buildings, Wood-brick multi-family (WBMF) buildings, Brick multi-family (BMF) buildings, and Concrete multi-family (CMF) buildings. The data on building structure was retrieved from the study of Bian et al. [77] who used building morphological indicators to predict building structures of residential buildings in Gothenburg. Results showed high accuracy (more than 80%), and we use them in this study to be able to apply Swedish BUD-MIs (which follow a structure-use-age archetype) to the inventory [31]. The material intensity includes 12 materials for both SF and MF buildings (See **Paper III**). Furthermore, the scope of the MI dataset for both SF and MF buildings includes the superstructure (above ground), substructure (underground), and foundation's compact layers.

The scope of the study in **Paper IV** is residential buildings divided into single-family and multi-family buildings. The material intensity used in the study is from the TABULA database [78] instead of BUD-MI because the thermal properties of materials necessary for building energy model has not been fully implemented in the EnergyPlus simulation manager (EPSM) tool [79]. The material intensities are not divided based on structure types and six different age cohorts are modeled. The renovation package modeled assumes U-value improvements in the roof, wall, and floor. The insulation materials included are mineral wool as baseline, cellulose and wood fiber insulations as biobased alternatives.

The geographical scope of both **Paper I** and **Paper III** are Sweden, and the geographical scope of **Paper IV** is narrowed to urban level for the Swedish city of Uddevalla. The narrowing of geographical scope is due to data availability. The temporal scope of **Paper I** is from 2022 to 2045 while the temporal scope of **Paper III** and **Paper IV** are from 2025 to 2050. The dMFA in **Paper I** is conducted at the regional scale (Sweden divided into 6 regions based on different lifetime estimates) and the dMFA in **Paper III** is conducted at the municipal level (290 municipalities in Sweden). **Paper IV** is conducted at the urban level using the lifetime modeling approach developed in **Paper III** to simulate the amount of floor space renovated every year.

CHAPTER 4

Methodology

This chapter briefly describes the methodology used in this thesis, for more detailed descriptions the reader can refer to the appended papers. While the focus of this chapter is on how the methods are applied to residential buildings, the machine learning and material stock and flow analysis workflow is still applicable to roads. This chapter also aims to synthesize and present the methodologies that are appended in the papers into a single framework or workflow. The structure of the chapter follows the chronological sequence of development of the framework and shows how the different methods are combined to address the aim of the thesis.

4.1 Machine learning

The generalized workflow of using machine learning models to predict values can be divided into the following steps. The first step of the workflow after collecting data is to inspect and preprocess the dataset. The main purpose of this step is to understand the dataset, such as the statistical distribution, completeness of data, and the potential existence of outliers. After gaining an initial understanding of the dataset, it is important to decide on whether to treat the prediction target either as a discrete or continuous variable. If the prediction target can take any values between two real values, for example road width **(Paper I)** or the usable floor space of a building **(Paper II)**, regression models are more suitable. On the other hand, if the prediction target can only be natural numbers in a subset of numbers, for example the construction year of a building **(Paper II)**, then classification models are more suitable. The preprocessing step could also include removing outliers if they exist or data stratification which involves dividing the dataset into subgroups before sampling. To summarize, the data processing step is crucial for the ML workflow, but also highly dependent on the input dataset.

4.1.1 Data preprocessing

One aspect of the data that requires special preprocessing is class imbalance in construction years of buildings. A 'class' in a classification problem is one of the possible categories or labels to which an input data can be assigned to by the model. Class imbalance is a problem for classification models when the underlying data are skewed and certain classes are under-represented, making the ML model's ability to predict minority classes less-effective [80]. In the building dataset, a significant class imbalance is found for construction years of buildings, more precisely between 1960 and 1980. This imbalance of construction years can be attributed the Swedish 'Million Program' in which more than 1 million buildings were built in a ten-year span (1965 to 1974) [81]. For regression models, no specific data preprocessing is required.

4.1.2 Feature engineering

The second step of the ML workflow is feature engineering, which is the main area of method development in **Paper II**. In machine learning, a feature serves as a numeric representation of a specific aspect of the raw data. Since ML models require an adequate number of features to capture and reflect effectively the underlying data's characteristics, feature engineering is essential [82]. This process involves extracting features from the raw data and converting them into formats that are suitable for utilization by the ML model [83].

The challenge that needs to be addressed at this step is the lack of height data at the national scale for Sweden. To address this challenge, urban form features based on urban morphology are hypothesized as predictive features. The advantage of using urban form features is that these features can be calculated using only building footprints and road data. The quantitative analysis of urban morphology distinguishes between three key building elements: footprints, plots, and street-based blocks [84], as illustrated in **Figure 3**, where tessellations enclosed by streets are used as a proxy for building plots. Each element captures the unique attributes of the corresponding building, and the features are generated based on the elements.

Building elements refer to the 2D geometry of the footprints of buildings, which contain useful information that can be used for predicting construction age [50] and floor space [45] through ML models.

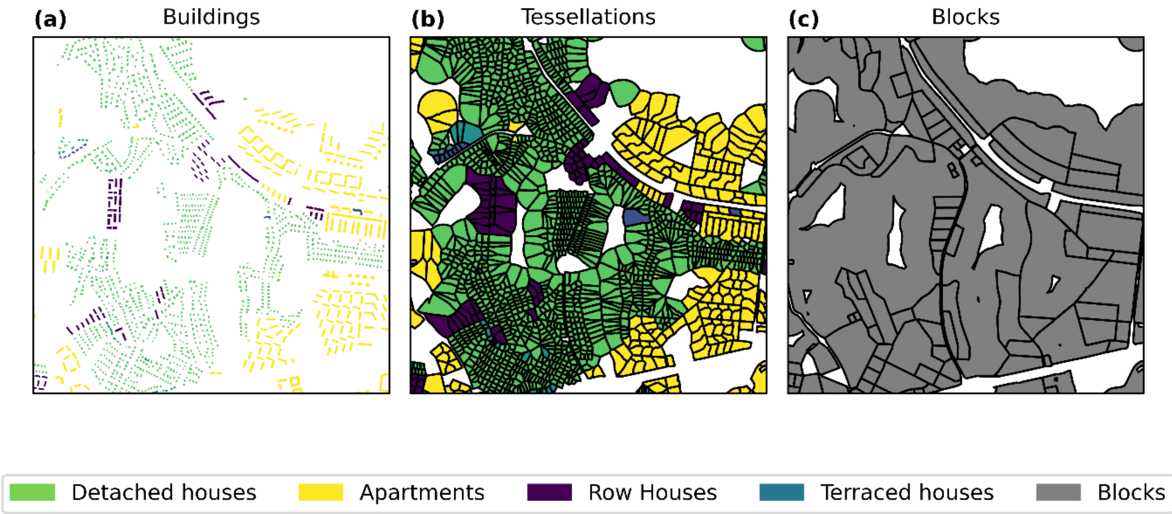

Figure 5. Illustration of the three elements of urban morphology using an area of Gothenburg (Sweden) as an example: (a) building footprints; (b) tessellations enclosed by streets, used as a proxy for building plots; and (c) building blocks generated using tessellations. Note that no color coding is used here because the blocks include a variety of building uses. Source: **Paper II**.

For the ML predictions of usable floor space and construction year of residential buildings, 16 urban morphology features are generated and applied to the three elements of urban morphology for a total of 48 features. All features are calculated using the Momepy package developed in Python using a combination of the building's GIS footprint data and road network GIS data [85]. The same set of features are used for both the classification and regression model. For the ML prediction of width of roads, 12 road network features are calculated using only road GIS data using Momepy. In addition, social economic features such as population and building features such as distance to the nearest buildings are used as features.

4.1.3 Classification models training and validation

The next step in the ML workflow after generating features is the model training and validation process. Classification models are used to predict building construction years (**see Paper II**). As previously discussed in **Section 4.1.1**, there is a significant class imbalance of construction years. Existing ML algorithms can mitigate some of the class imbalance problems, but not always in a satisfactory manner. Therefore, an additional step is required to counteract the class imbalance and maximize the ML model's performance [86].

There are multiple available techniques to tackle the class imbalance problem, and the under-sampling technique is chosen due to the underlying distribution of the data (See **Paper II**). Therefore, the open-source Python package 'Imbalanced-learn' is used to test four under-sampling techniques to identify which technique can produce the optimal prediction performance [87]. The first technique is the Random Under-Sampler (RUS), which randomly under-samples the majority classes without replacement. The second is the Near Miss (NM) method, which selects a subset of the majority classes samples that are closest to the minority classes samples [88]. The third is the One-sided Selection (OSS) technique, which initially finds the observations that are hard to classify and then removes noisy samples [89]. The fourth is the Neighborhood Cleaning Rule (NCR) method, which uses a combination of edited-nearest-neighbor and a k-nearest-neighbor to remove noisy samples from the dataset [90]. Subsequently, in Paper **II** we use the XGBoost algorithm for the classification task and train the model on the four different under-sampled datasets. XGBoost is chosen as it has been successfully applied to predict various attributes of the building stock with high levels of precision [46], [91]. Therefore, XGBoost algorithm is tested with different resampling methods to address class imbalance issues in the dataset.

To evaluate the result from the under-sampling techniques, a set of evaluation metrics are used. The metrics used in the study and their associated equations and descriptions are shown in **Table 3** in **Paper II**. There is currently no clear consensus in literature on what are the most accurate evaluation metrics, so we use the most used metrics. The best performing under-sampling technique is then used to train the final classification model that predicts the construction year of buildings.

4.1.4 Regression model training and validation

The road width prediction problem in **Paper I** and the usable floor space prediction problem in **Paper II** are tackled with regression ML models. Since regression models do not face class imbalance problems, we compare the four most widely used algorithms to test which has the best performance instead. One of these is XGBoost [92], which is also used for the classification model. XGBoost is an ML algorithm that is a member of the family of gradient-boosting techniques [55]. XGBoost is an ensemble learning method that combines the predictions of multiple individual decision trees, known as weaker learners, to create a strong final predictive model. In gradient boosting, new models are built sequentially by correcting the mistakes of the previous model. Each new model is trained to predict the residual errors of the ensemble of previous models.

In addition to XGBoost, we test three ML algorithms that are more commonly used for regression problems. Random forest (RF) is an ensemble learning algorithm that builds a collection of decision trees and combines their predictions to create a more-accurate and robust model [93]. The RF algorithm is implemented using the scikit-learn package in Python [94]. LightGBM is another gradient-boosting decision tree algorithm developed by Microsoft that has higher training efficiency and lower memory usage [95]. CatBoost is yet another gradient-boosting algorithm developed to offer more support for categorical features [96]. For both the regression and classification models, a five-fold cross-validation method is used to prevent the model from overfitting to the training data. Overfitting refers to an ML model that learns the training data too well, including the noise to the point that it performs poorly on test data. The consequence of overfitting is that the trained model performs exceptionally well on the training data but cannot generalize to make accurate predictions on new data. The hyper-parameters of the models are optimized using the Optuna Python package [97].

To validate the results of the ML models, several evaluation metrics are used. The regression models use a set of more traditional evaluation metrics. The first metric is the Mean squared error (MSE) as the evaluation metric during the model training process. MSE measures the average squared differences between the predicted and actual values and is defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{1}$$

where n is the total number of data-points, y_i is the actual value of the i-th observation, and $\hat{y}_i$ is the predicted value of the i-th observation.

Two additional evaluation metrics are used to compare the different ML algorithms for regressions: Mean absolute error (MAE), and R-squared score (R^2). MAE measures the average absolute differences between the predicted and actual values, defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

where n is the total number of data-points, y_i is the actual value of the i-th observation, and $\hat{y}_i$ is the predicted value of the i-th observation. MAE penalizes errors to a lesser degree than the MSE, but it is a more-intuitive evaluation metric.

The R-squared score measures the proportion of the variance in the dependent variable that is predicted from the independent variables and is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3}$$

where n is the total number of data-points, y_i is the actual value of the i-th observation, $\hat{y}_i$ is the predicted value of the i-th observation, and $\bar{y}$ is the mean of the target variable.

Like the classification model training and validation process, the best performing regression algorithm is chosen and then used to predict the road width and usable floor space.

4.2 Material stock and flow analysis

This section provides an overview of the methods and equations used for the modeling of the material stocks and flows of the Swedish roads and residential buildings. Equations 4-6 are generalized for both roads and residential buildings, and specific equations can be found in the appended papers.

4.2.1 General material stock and flow modeling

This subsection describes the general material stock and flow modeling approach used in **Paper I** and is the basis for the modeling framework developed in **Paper III**. The material stock and flow modeling done in this work is all based on the stock-driven approach [5]. The dMFA models are coded in Python using the Open Dynamic Material Systems Model (ODYM) [98]. The first step is to calculate the in-use material stock, and the stock is calculated by multiplying the inventory with material intensity, as shown in equation (4):

$$MS_i(t) = MI_i(t) * Inv_i(t) \tag{4}$$

where $MS_i(t)$ is the material stock of built environment type i at time t in tons, $MI_i(t)$ is the material intensity of built environment type i at time t in tons, and $Inv_i(t)$ is the inventory of built environment type i at time t. The inventory used in **Paper I** is road surface area in m^2 (length multiplied by width), with both existing and ML-predicted width data. The inventory used in **Paper III** is the usable floor space of buildings in m^2. The inventory is the result of **Paper II**.

The first step of the stock-driven DMFA model is to calculate the outflow from existing stock, as shown in equation (5):

$$outflow_{i,c}(t) = \sum_{\tau=t_0}^{t} inflow_i(\tau) * (1 - Survival_c(t-\tau)) \tag{5}$$

where $outflow_{i,c}(t)$ is the outflows or demolition of stock type i from age-cohort c (construction year) at the end-of-life (EoL) at time t in tons, and $Survival_c(t-\tau)$ is the survival function that represent the share of the inflow from age-cohort c remaining in the stock at time t. The survival table is constructed from an age-cohort dependent lifetime distribution. There is no consensus in literature on functional forms of lifetime distributions [99]. The Weibull distribution is chosen as it is the most commonly used and has the strongest empirical data support [100].

The next step is to calculate the inflow of building type i at time t using the mass-balance principle, as shown in the following equations:

$$S_{survive_{i,c}}(t) = \sum_{\tau=t_0}^{t} (inflow_i(\tau) - outflow_{i,c}(\tau))$$

$$inflow_i(t) = S_i(t) - S_{survive_{i,c}}(t) \tag{6}$$

where $S_i(t)$ is the stock type i at time t in tons and $S_{survive_{i,c}}(t)$ is remaining or surviving stock type i at time t in tons after $outflow_{i,c}(t)$ is removed.

4.2.2 Road maintenance

A key assumption behind the modeling of road maintenance in **Paper I** is that no roads are demolished in Sweden. Therefore, only one stock-driven DMFA is conducted to estimate the material flow from the maintenance of roads. In addition, it is assumed that only 50mm of asphalt is removed and repaved from maintenance according to expert opinion (from PEAB construction company that works with road maintenance in Sweden).

4.2.3 Building renovation modeling

This subsection describes the modeling framework developed in **Paper III** to model building renovation based on the concept of building layers described above. The logic behind the framework is summarized into the following steps:

(1) The gross floor area (GFA) dynamics (stocks and flows) are calculated from construction and demolition activities, using a first stock-driven dMFA (Section 3.2.1);
(2) The GFA dynamics (flows) from renovations activities are then estimated, using two stock-driven dMFAs (Section 3.1.2) running in parallel for the skin and space layers;
(3) Subsequently, the flow results are converted into material quantities, using material intensities (MIs) (Section 3.1.3);

The main concept behind the layer renovation model is demonstrated in **Figure 6**. In the monolithic model, renovations happen completely independently from the demolition probabilities or structural lifetime. On the other hand, the layer renovation model couples the renovations with structural lifetime by only renovating the surviving cohort of buildings after demolition. The result of this coupling or convolution is that renovation curves follow closely with the demolition probabilities.

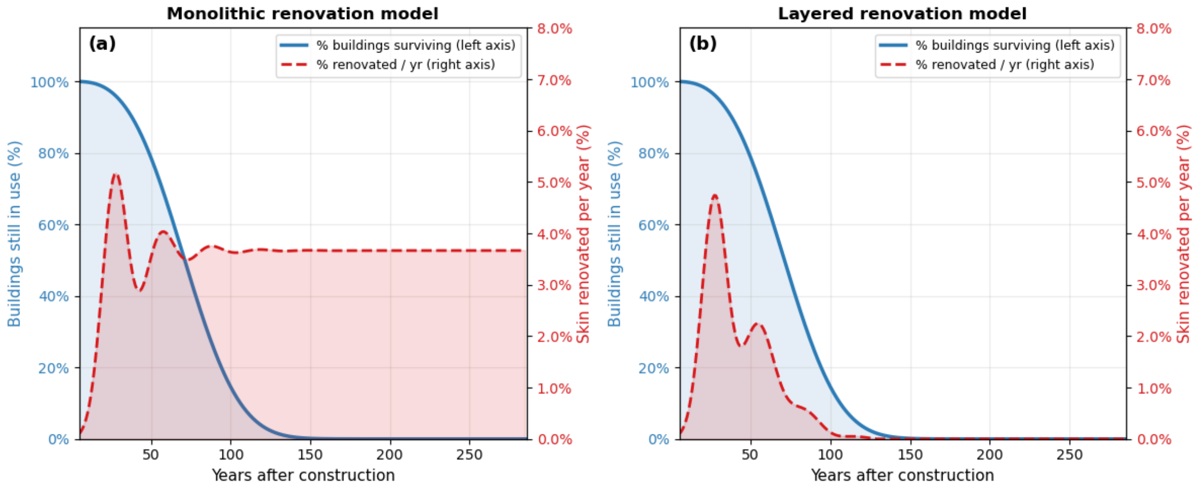

Figure 6. Annual skin renovation rate versus cohort age under a normal skin lifetime distribution (mean μ_{skin}, CV = 0.3). The left axis (blue) shows the fraction of buildings still in use; the right axis (red, dashed) shows the percentage of skin area renovated per year. In the monolithic model (a) the skin DSM is driven by constant gross floor area, producing persistent full-amplitude renovation cycles. In the layered model (b) the skin DSM is driven by surviving floor area $S_b(t)$, so renovation demand reduces as the cohort is gradually demolished.

A key data input to operationalize the framework is the BUD-MI for Swedish residential buildings that is disaggregated across the building layers. The concept behind the framework is that different building layers undergo renovations at different rates and thus should be represented with different lifetimes [101]. At the same time, renovation probabilities should be related to the lifetime of the entire building instead of being independent.

The first stock-driven dMFA model is used to model demolition of the stock represented by outflow. The lifetime distribution to represent demolition is chosen to be a Weibull distribution, with a scale of 300 and shape of 3 for all buildings built after 1920, with buildings built before 1920 being excluded from renovation as they are assumed to be historical buildings. The Weibull distribution is chosen as it is commonly used in literature, and the choice of lifetime distribution does not significantly impact results [102]. The rationale for choosing these parameters is the need to calibrate the demolition rate based on statistical data. There is no official data for SF building demolition rate, and 1,400 MF buildings were demolished in 2024, which corresponds to a demolition rate of approximately 0.05% of the existing building stock [70]. Equations 4 – 6 are used for this first stock-driven model.

The next two parallel stock-driven models that are described in Step 2 are collectively referred to as the layer renovation model (LRM). In the LRM, renovation activities are modeled with an additional stock-driven MFA applied to the surviving stocks from each age-cohort. First, only the surviving stock from each age-cohort is subject to renovation. Therefore, the lifetime convolution method proposed by Sartori et al. (2016) – which ensures that no demolished buildings are renovated – is adapted. The main concept behind the approach is that demolition happens only once, but renovation happens multiple times in cyclical fashion. The implementation of this approach is by shifting the lifetime profile by the mean renovation cycle (e.g., 25 years) to ensure that a building's expected lifetime after renovation is at least as long as the renovation lifetime itself. We adapt and apply this approach to the skin and space layer using specific survival curves, using the following equations:

$$O_{i,c,L}^{R}(t) = \sum_{\tau=t_0}^{t} I_{i,c}^{R}(\tau) \times \left(1 - surv_{L,c}(t-\tau)\right) \tag{7}$$

$$S_{S,i,c}^{R}(t) = \sum_{\tau=t_0}^{t} (I_{i,c}^{R}(\tau) - O_{i,c}(\tau)) \tag{8}$$

$$I_{i}^{R}(t) = \sum_{c} S_{S_{i,c}}(t) - S_{S,i,c}^{R}(t) \tag{9}$$

where $O_{i,c,L}^{R}(t)$ is the outflow of renovated floor area of layer L of building type i from age-cohort c at time t (in m^2), $I_{i}^{R}(t)$ is the inflow of floor area or renovation of building type i (in m^2), and $Surv_{L_c}(t-\tau)$ is the survival function that represents the share of the inflow from age-cohort c remaining in the stock at time t. These equations are applied to both the skin and space layer in parallel, as we assume that skin renovation for functional maintenance and space renovation for changing internal layout happens independently.

This specific survival function is used to represent the lifetime of the layer (skin or space), and a building requires renovation if the layer (skin or space) reaches its EoL. As a stock-driven model is applied to each age-cohort c, the total amount of floor area in need of renovation at time *t* is calculated by summing across all the age-cohorts for each time *t*, as shown in Equation (6).

The in-use MS is calculated by multiplying the floor area by its relevant material intensity, as shown in Equation (10):

$$\begin{cases} MS_{i,m,L}(t) = \sum_{c} MI_{i,c,m,L} \times S_{i,c}(t) \\ \mathrm{L} = \{\text{structure; skin; space}\} \end{cases} \tag{10}$$

where $MS_{i,m,L}(t)$ is the material stock in building type i for material type m at time t for layer L, $MI_{i,c,m,L}(t)$ is the material intensity of material type m for layer L in building type i of cohort c. The rationale for this formula is that buildings from different cohorts have different material intensities, and within these buildings, the building layers also have different material intensities.

Lastly, the inflows associated with renovations are calculated using Equation (11):

$$\begin{cases} m_{I^R_{i,m,L}}(t) = \sum_{c} MI_{i,c,m,L} \times I^R_{i,L}(t) \\ \mathrm{L} = \{\text{skin; space}\} \end{cases} \tag{11}$$

where $m_{I^R_{i,m,L}}(t)$ is the inflow of construction material from renovations of building type i of material type m for layer L at time t (in kg/year), $MI_{i,c,m,L}(t)$ is the material intensity of material type m of layer L of building type i of cohort c (in kg/m^2), and $I^R_{i,L}(t)$ is the inflow of GFA of layer L being renovated at time t (in m^2).

4.3 Building heat demand modeling

This subsection describes the modeling of building energy (more specifically heat) demand modeling in **Paper IV**. In the paper, the LRM model is used first to model the amount of floor area in need of renovation from each age cohort. The renovated buildings are then assigned a new U-value for each building component to simulate energy efficiency renovations, with a focus on building envelopes. The building age cohorts, components, and U-values are shown in **Table 3**.

Table 3. Description of the current U-value and the target U-value after energy efficiency renovation for each building component based on age cohort and building type from TABULA for Sweden.

Age cohort	Building component	Baseline U-value SF building (W/mK)	Target U-value SF building (W/mK)	Baseline U-value MF building (W/mK)	Target U-value MF building (W/mK)
Before 1960	Roof	0.29	0.11	0.36	0.12
	Wall	0.6	0.33	0.58	0.29
	Floor	0.28	0.21	0.32	0.24
1961 - 1975	Roof	0.21	0.1	0.2	0.1
	Wall	0.31	0.22	0.41	0.24
	Floor	0.32	0.24	0.26	0.2
1976 - 1985	Roof	0.15	0.08	0.17	0.09
	Wall	0.21	0.16	0.33	0.21
	Floor	0.27	0.21	0.27	0.21
1986 - 1995	Roof	0.12	0.07	0.15	0.08
	Wall	0.17	0.14	0.22	0.16
	Floor	0.24	0.19	0.24	0.19
1996 - 2005	Roof	0.12	0.07	0.13	0.08
	Wall	0.2	0.16	0.2	0.15
	Floor	0.18	0.15	0.21	0.17

The building heat demand modelling is carried out using the Energy Performance Simulation Manager (EPSM) package [79]. EPSM is developed as a containerized package that can be used to simulate energy renovation strategies across large building stocks in a streamlined fashion. The core simulation engine used in EPSM is the EnergyPlus model [104], and the function of EPSM is that the user can interact with EnergyPlus in a more consistent fashion. Although EPSM is primarily designed as a web application, in this study EPSM's backend code is used in Python to conduct simulations.

The first step is to select the study area and then download the building footprints and building heights calculated using LiDAR data through the DTCC package contained in EPSM [105]. The next step is to model the building dynamics to simulate how many buildings are renovated using the method outlined in Section 2.2. The building dynamics are modelled in yearly intervals and subsequently grouped into three renovation year intervals (2030, 2040, 2050) to reduce the computational complexity for EPSM. Each building is then assigned a baseline set of U-value. Subsequently, all buildings that are modelled to be renovated are assigned updated U-values as shown in **Table 3**. The energy simulations are then carried out using a set of predefined parameters for EnergyPlus that are selected as default value in EPSM. The results are then parsed to extract the heating demand. EPSM is not developed as part of the thesis, but further adaption is carried out to couple EPSM with LRM.

4.4 Biobased insulation materials and carbon storage accounting

Another important aspect of **Paper IV** is the quantification of carbon storage potential of biobased insulation materials. The dynamic approach and more specifically the GWPbio framework proposed in [106], [107] is chosen to account for the emissions of biobased insulation materials. The rationale for choosing the GWPbio method and a brief review of the other accounting methods can be found in **Paper IV**.

The GWPbio method considers the storage period of biomass as well as rotation period, to dynamically estimate the biogenic GWP of a bio-based material used in buildings. Rotation period refers to the number of years required for a biomass to become harvestable. The essence of this methodology is that biogenic CO_2 emissions are compensated by the growth in biomass while fossil CO_2 emissions have no such replacements, thus temporarily storing biogenic carbon results in overall a lower cumulative radiative forcing (i.e. GWP) over the time horizon of assessment. The two main factors that affect the GWP of biobased insulation are the storage period (i.e., lifetime) and the rotation period. Storage period affects how long the carbon is stored, and rotation period affects when the sequestration from regrowing starts. The implication is that longer storage periods and shorter rotation periods lead to lower GWP.

The formula for calculating the biogenic GWP from storage ($GWP_{storage}$) is shown in:

$$GWP_{storage} = CO_{2,biogenic} \times GWP_{bio} \tag{12}$$

where $GWP_{storage}$ is the amount of biogenic carbon storage, $CO_{2,biogenic}$ is the amount of biogenic carbon in the insulation material and GWP_{bio} is the GWP factor that is based on Priore et al. (2026). The biogenic carbon values are taken from Boverket's climate database [109].

Under current Swedish regulations, biogenic carbon is not included in the building climate declaration (developed based on LCA methodology) as the greenhouse gases associated with end of life of the buildings (stages C1 – C4) are not included in the LCA system boundary [110]. This methodology utilizes concepts and values from the Life Cycle Assessment (LCA) literature but does not aim to directly engage with the methodology discussion on how to use life-cycle based data to set limit values for renovations in Sweden [111]. Instead, the method and results are used to demonstrate the modeling framework and its transferability and generalizability to other countries or regions.

4.5 Carbon emissions quantification

This subsection briefly describes the methods used to quantify carbon emissions in the thesis. Embodied emissions are the focus of this thesis, and embodied emissions are quantified in **Papers I**, **III**, and **IV**. The embodied emissions are calculated using equation (13):

$$CO_{2,embodied} = \sum_{i,t,m} M_{inflow_{total_{i,t,m}}} * EF_m \tag{13}$$

where $CO_{2,embodied}$ is the total embodied CO_2 emissions in year t in tons, $M\_inflow\ \_total_{t,m}$ is the total inflow of material m from stock type i at time t in tons, and EF_m is the emission factor for material m in tons/ton. The emission factors (*EF*) used here are based on calculation done by Karlsson et al. [68], and the EFs account only for life cycle stages A1–A3, thereby excluding emissions associated with later stages such as transportation to site and construction (A4–A5). The emission factors calculated by Karlsson et al. are the main sources of emissions data used in this study to ensure that the results are consistent across papers.

Similarly, the operational emissions in **Paper IV** are calculated by multiplying the heat demand with emission factors, as shown in Equation (14):

$$CO_{2,operational}(t) = D_{hp}(t) * EF_{hp} + D_{dh}(t) * EF_{dh} \tag{14}$$

where $CO_{2,operational}(t)$ is the operational emission in year *t*, $D_{hp}(t)$ and $D_{dh}(t)$ are heat demand for SF and MF building at year *t*, and EF_{hp} and EF_{dh} are emission factors for heat pumps and district heating. The heating supply systems of buildings are modeled based on Energy Performance Certificates (EPC) data, where 7% of SF buildings and 49% of MF buildings are supplied by district heating (DH) while other buildings are assumed to be supplied by heat pumps (HP)

CHAPTER 5

Results from the machine learning training

The results from the machine learning training and testing process are presented in this chapter. The regression model results are synthesized from **Paper I** for roads and **Paper II** for residential buildings. The classification results are only from **Paper II**. Lastly, the completed building inventory dataset that is used as input data for **Papers III** and **IV** is presented. The results presented in this chapter answer **RQ I**.

5.1 Regression model

This following section describes the results of the regression ML models that are trained and tested to predict missing road with data and building usable floor space in **Papers I** and **II** respectively. For both prediction tasks, the same sets of ML algorithms are tested, including Random Forest (RF), extreme gradient boosting (XGBoost) [92], CatBoost [96], and Light gradient-boosting machine (LightGBM) [95].

5.1.1 Road widths prediction

Machine learning algorithms are trained and tested to predict missing road width data in **Paper I**. The evaluation metrics used are MAE, MAPE and R^2. The performance levels of the trained models are shown in **Table 4**. The results from the LightGBM and XGBoost models are very similar, with XGBoost producing slightly lower MAE, MAPE, and higher R^2 value. The similarity in the results is not unexpected, as the two models are variations of the gradient-boosting algorithm.

Table 4. Results of the evaluation metrics for each regression machine learning algorithm. For MAE, the lower the absolute value the better the performance, and for R2 the closer to 1 the better the results.

Algorithm	**MAE**	**MAPE (%)**	R^2
Random forest (RF)	0.623	13.4	0.748
XGBoost	0.567	12.5	0.784

CatBoost	0.669	14.4	0.728
LightGBM	0.576	12.6	0.781

MAE, mean absolute error; MAPE, mean absolute percentage error; R^2, coefficient of determination.

The RF algorithm has fewer hyperparameters available for tuning and, therefore, performs slightly poorer in terms of the MAE, MAPE and R^2 parameters. The CatBoost model results in the highest MAE and MAPE values and the lowest R^2 value of all four tested algorithms. This is most likely due to the relatively low number of categorical features (6) in the dataset, which does not fully take advantage of the algorithm's implementations on categorical features. Therefore, the XGBoost model is chosen to predict the missing road widths.

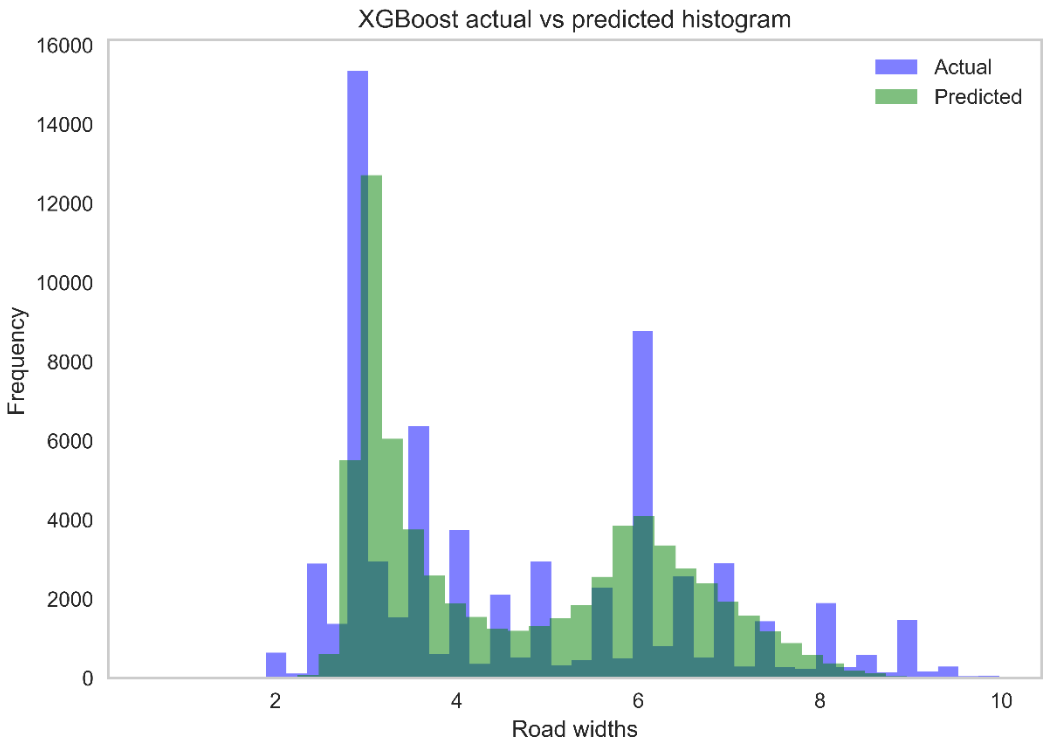

Figure 7. Comparison of actual versus predicted value in meters, the number of bins used is 40. Source: **Paper I**.

The results of the XGBoost model are presented in **Figure 7**. Overall, the model does capture the distribution of real data relatively well in terms of filling in missing data. Actual values (blue) refer to the validation dataset used in the training process and the predicted values (green) are the results produced by the XGBoost model using the same training dataset. The results show that the model is less capable of predicting extreme values, especially road widths exceeding 10 meters. Furthermore, the model under predicts roads with 6 meters in width and results in a normalized distribution for road widths between 4.5 and 8 meters.

5.1.2 Building usable floor space prediction

This subsection presents the results of the ML training and testing process to predict building usable floor space in **Paper II.** The results for each regression model are presented in **Table 5**.

Table 5. Results of the evaluation metrics for each regression machine learning algorithm. For MAE and RMSE, the lower the absolute value the better the performance, and for R^2 the closer to 1 the better the results.

ML models	MAE	RMSE	R^2
XGBoost	28.74	101.04	0.789
LightGBM	29.43	101.43	0.788
CatBoost	31.19	102.92	0.781
Random Forest	32.16	108.81	0.756

MAE, Mean absolute error; RMSE, Root mean-square error

The overall best-performing model is XGBoost, while LightGBM achieves similar results. Random forest is the worst-performing model by a relatively wide margin. This is an expected result, since the other three algorithms are improvements over RF in their implementation. An R^2 value of 0.789 can be considered as good performance. For comparison, in the study of Milojevic-Dupont and coworkers, the best-performing model for predicting building heights achieved an R^2 of 0.72 [91].

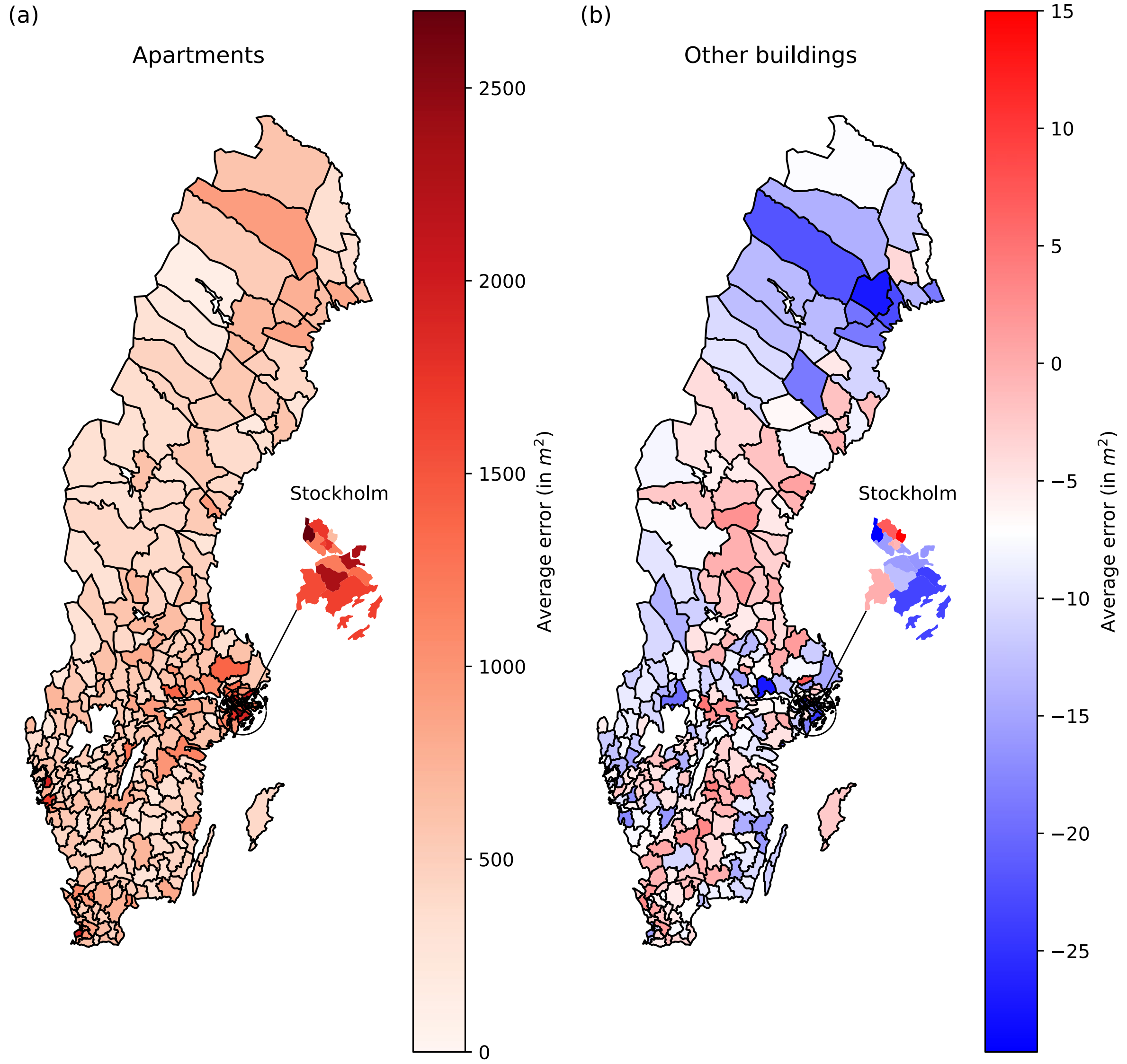

Figure 8. Average errors of the predictions of floor space in all municipalities in Sweden for: (a) apartment buildings: and (b) all other types of residential buildings (in m^2). Source: **Paper II**.

To understand the performance of the model at a more-granular level, the average prediction errors for floor space in all municipalities in Sweden are visualized in **Figure 8**. Prediction errors or residuals are defined as the ground truth floor space minus the predicted floor space. The average prediction errors for apartments are visualized separately from other types of residential buildings because the range of floor space values is much broader for apartments. The average prediction errors show large variations, with the minimum average error being 38.6 m^2 and the maximum average error being 2,644.4 m^2, with a standard deviation of 361.7 m^2. The three municipalities with average prediction errors >2,000 m^2 are all located in the

Greater Stockholm area (Järfälla, Huddinge, Nacka). This is likely due to the large amounts of missing data in these municipalities, with 92.4%, 80.9% and 82.3%, respectively, of the data being unavailable. The average percentage of missing apartment floor space data for the eleven municipalities in the greater Stockholm area is 76.4%, which likely accounts for the high average error observed for these municipalities. The median prediction error is 383.4 m^2 and the 75th percentile of the average errors is 544.2 m^2, indicating that the model performs relatively well, apart from the municipalities with lots of missing data.

The average prediction errors for other types of buildings (excluding apartments) show a much narrower range of variations, with minimum average error of -29.3 m^2 and maximum average error of 7.99 m^2 and a standard deviation of 6.05 m^2. The model on average underpredicts the floor space areas of other types of buildings, with a mean of -7.39 m^2 across all municipalities. There are also weaker spatial correlations for the prediction errors, unlike the case for apartment buildings.

5.2 Classification model

This subsection presents the result from the classification model to predict building construction year. Unlike the regression models that compare the performance between different ML algorithms, the classification model training and testing is focused on comparing the performance of different under-sampling methods.

The performance levels of the trained models are shown in **Table 6.** The results show that the choice of under-sampling method has a relatively potent impact on the performance level of the classification model. The difference between the MMCCs of the best-performing NCR under-sampling method and the worst-performing sampling method is 70.1%. The reason for this disparity in performance level is likely due to the random nature of RUS in removing valuable and informative data-points from the sample. Based on all the evaluation metrics, the NCR produces the best results. Thus, the results obtained from the classification model that uses NCR for resampling are used for the subsequent regression model training.

Table 6. Results of the evaluation metrics for each under-sampling method. For each of the evaluation metrics, a higher score (closer to 1) represents a better prediction performance.

Under-sampling method	AUPRC	MAUC	G-mean	MMCC
RUS	0.631	0.783	0.532	0.243
Near Miss	0.663	0.922	0.760	0.564
OSS	0.671	0.912	0.777	0.571
NCR	0.823	0.974	0.911	0.814

AUPRC, Area Under the Precision-Recall Curve; MAUC, Mean Area Under the Curve; MMCC, Mean Mathews Correlation Coefficient; RUS, Random Under-sampler; OSS, One-sided Selection; NCR, Neighborhood Cleaning Rule.

To understand more in detail the classification model's performance for each class, the normalized confusion matrix for the best-performing model is presented in **Table 6**. A confusion matrix evaluates the performance of a classification model by providing a detailed breakdown of the model's predictions compared to the actual outcomes to see where the model is making errors. The confusion matrix is row-wise normalized to increase the readability of the plot. In a multiclass confusion matrix, the diagonal line represents the true predictions. In the confusion matrix, each row represents the true age cohort, while each column represents the age cohort predicted by the model. As shown in the confusion matrix, the two age cohorts for which the model performs poorly are 1910-19 and 2020-23. The 1910-19 cohort has significantly fewer samples compared to the previous cohort (1900-09), and the two cohorts likely have very similar features and, thus, are more difficult for the model to predict. Similarly, the 2020-23 cohort has a relatively low number of samples, although the model shows the worst performance for this cohort because around 16% of the wrong predictions are for the 1900-09 cohort.

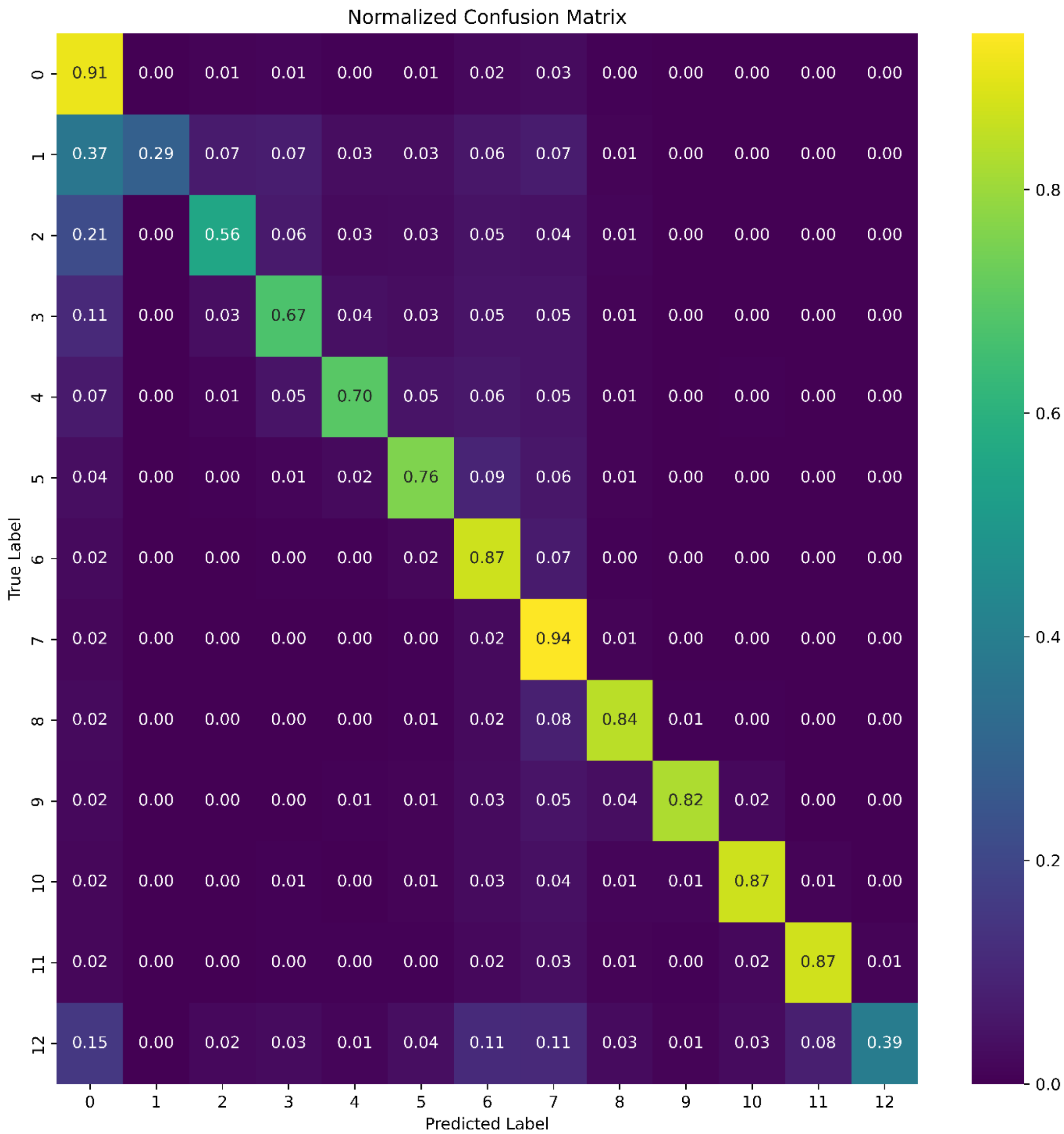

Figure 9. Row-wise normalized confusion matrix of prediction accuracy based on the model trained on the Neighborhood Cleaning Rule for the sampled dataset. Source: **Paper II**.

The errors of the classification model are comparable to a similar study by Dabrock et al. that predicted construction years of residential buildings in Germany with morphological features but with height data as input data [112]. The results of Dabrock et al. show that their model achieved an F1-score (similar evaluation metric) of approximately 74%, and this accuracy is comparable to the results from our models.

5.3 Building stock composition

The resulting building stock composition from the machine learning models is shown in **Figure 10**. For the input data to the machine learning models refer to Paper **II**. Only buildings built after 1920 are shown as in **Paper III** buildings built before 1920 were excluded from renovations.

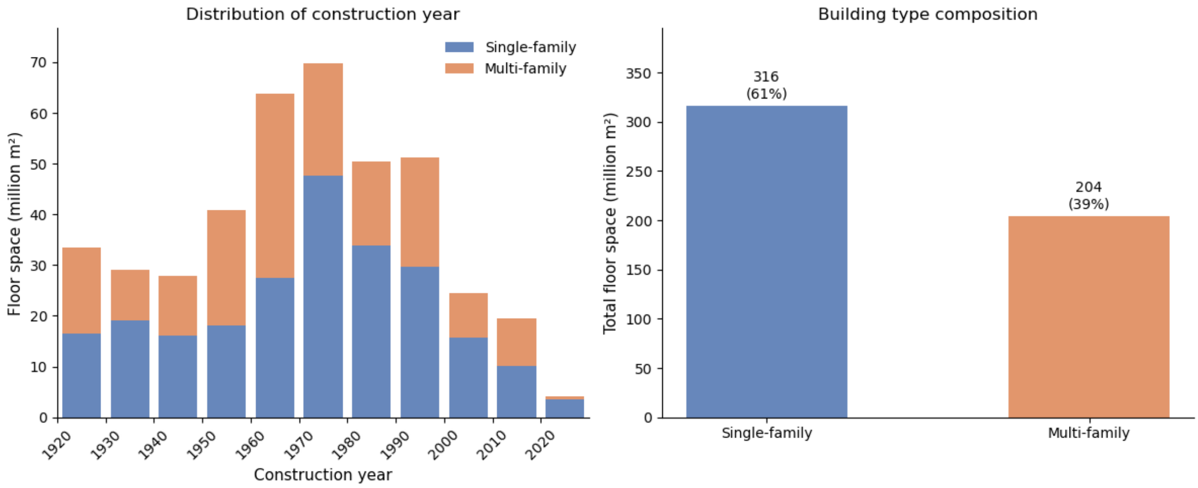

Figure 10. Distribution histogram of construction years and building type split between single family buildings (SF) and multi-family (MF) buildings for the building stock in **Paper III,** which is for residential buildings in Sweden.

In total, there are 2,55,406 buildings included in the stock, with SF buildings representing 61% and MF buildings representing 39% of the total usable floor space, at 316 million m^2 and 204 million m^2 respectively. The mean construction years for SF buildings are 1955 and 1958 for MF buildings. The two largest cohorts of buildings are built in 1960 and 1970s, which is a result of the 'Million Program', which was a public housing program implemented to build one million new housing dwellings [81].

In addition, the building stock composition of the urban area of Uddevalla that is the input data for **Paper IV,** is shown in **Figure 11**. Similarly, only buildings built after 1920 are shown in the plot.

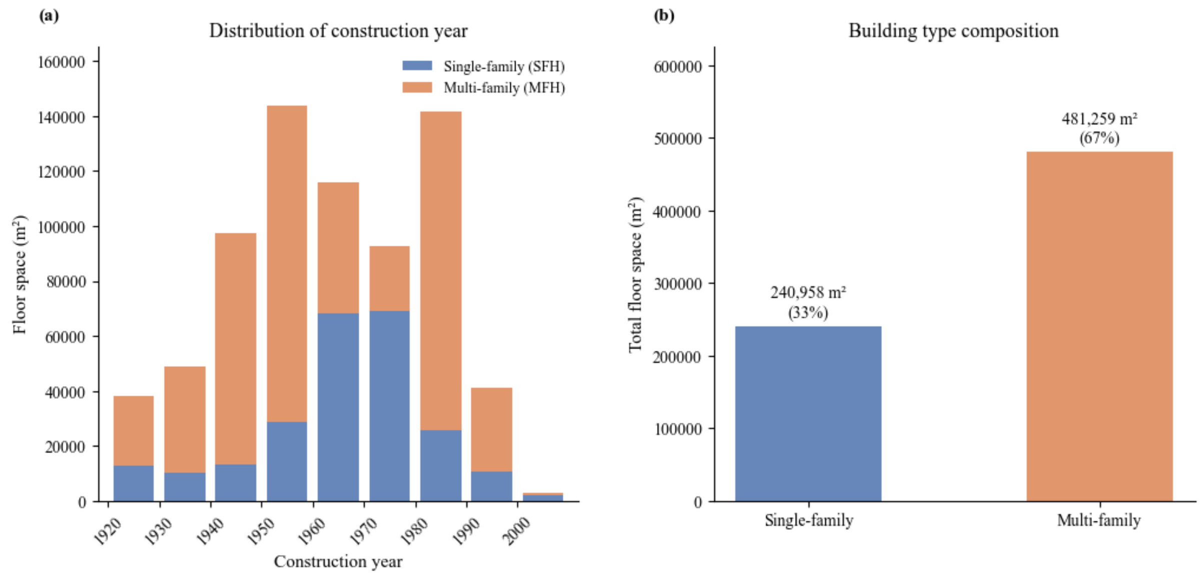

Figure 11. Distribution histogram of construction years and building type split between SF and MF buildings for the building stock of the City of Uddevalla investigated in **Paper IV**.

There are 2,334 buildings in the Uddevalla building stock, with 1,738 SF buildings and 596 MF buildings. The total amount of usable floor space is 722,217 m^2, of which 33% are SF buildings and 67% are MF buildings. The mean year of construction for MF buildings is 1954 while the mean year of construction for SF buildings is 1966.

CHAPTER 6

Material stock and flow

This section presents the road and residential building material stock and flow results at the national level. First, the total material stock of both roads and buildings synthesized from **Paper I** and **Paper III** are presented, which answers **RQ II**. The detailed material stock and flow results based on BUD-MI building layers and the LRM from **Paper III** are presented in subsequent subsections, which answer **RQ III**.

6.1 Total material stock

The material stock of roads and residential buildings for each county in Sweden is shown in **Figure 12**. In total, the material stock accumulated in roads in Sweden amounts to 1.95 megatons (Mt) and the material stock accumulated in residential buildings to 267 million tons.

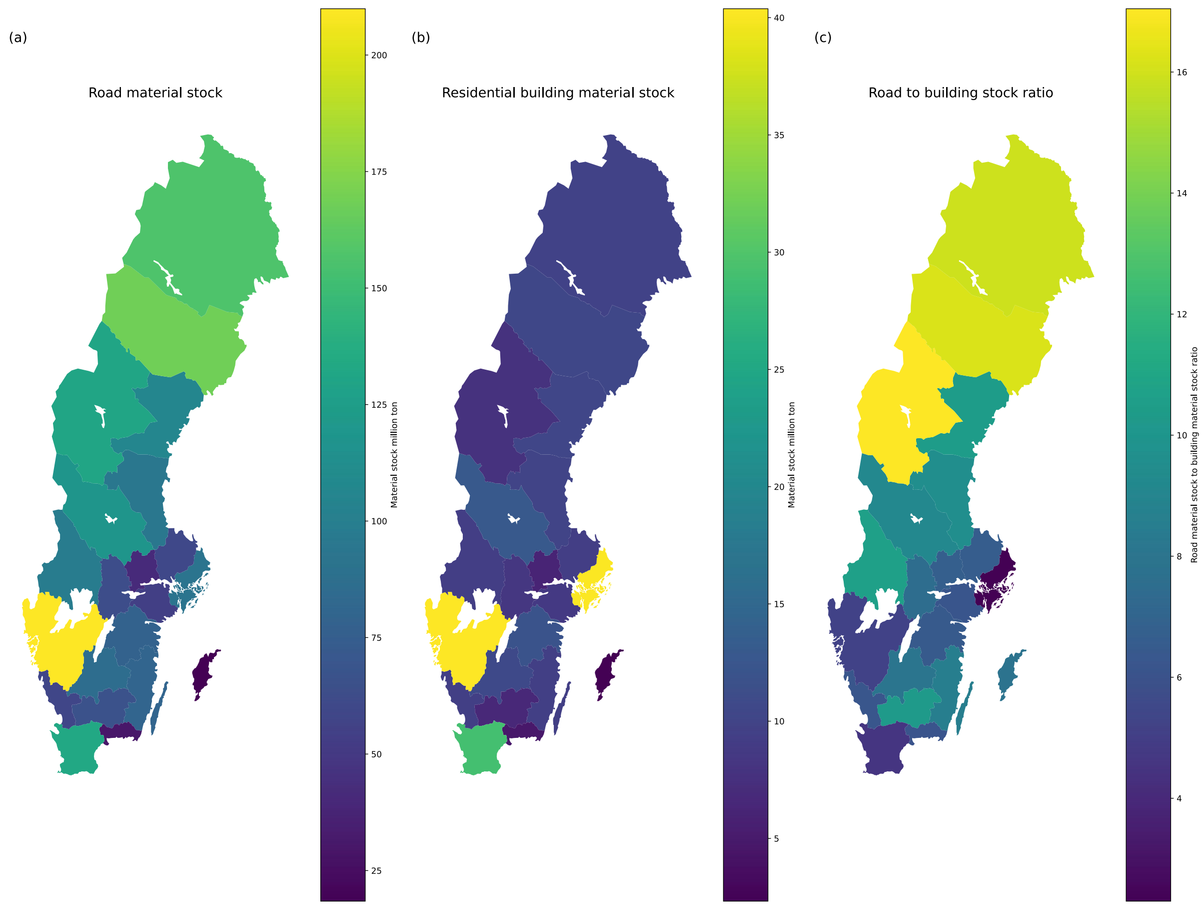

Figure 12. Material stock accumulated in roads and residential buildings for each county in Sweden, the unit is million tons (10^6). Plotted specifically for the thesis.

The county with the most absolute amount of material stock accumulated in roads is Västra Götaland where Gothenburg, the second largest city in Sweden is located. The counties with the second and third largest amount of material stock accumulated are Västerbotten and Norrbotten, which are both counties in the far north of Sweden with vast geographical areas and thus the need for more roads. Stockholm as the largest city in Sweden does not have the most amount of absolute material stocks accumulated in roads due to the relatively small geographical area of the county.

The counties with the most material stock accumulated in buildings are Västra Götaland, Stockholm, and Skåne, where the three largest cities in Sweden (Gothenburg, Stockholm, and Malmö) are located. The Västra Götaland county has 0.26 Mt more material stocks in residential buildings compared to Stockholm, which is due to the county of Västra Götaland encompassing a much larger rural area than the county of Stockholm.

Due to the large geographical area and sparsely distributed population, absolute material stock does not show the full picture of the distribution and accumulation of material stock. Therefore, the material stock per square meter for each county in Sweden is shown in **Figure 13**.

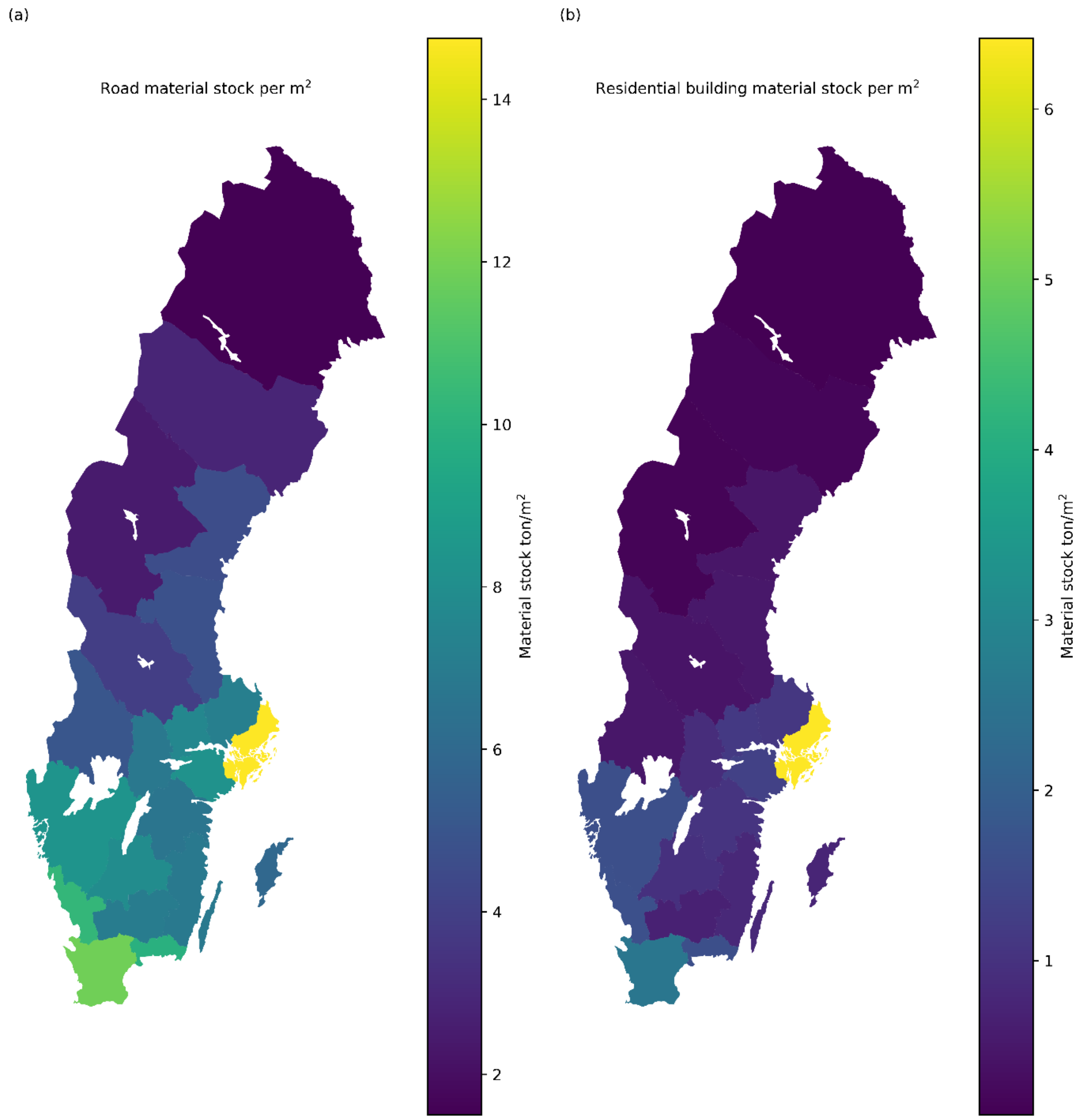

Figure 13. Material stock accumulated in roads and residential buildings for each county in Sweden, in t/m^2. Plotted specifically for the thesis.

By dividing the material stock by the area of each county, it can be seen much more clearly that material stocks are mainly accumulated in population centers. Moreover, there is a clear North-South divide, where most of the Swedish population resides in the southern part of the country.

6.2 Building material stock

This subsection presents the building material stock results that are disaggregated into building layers.

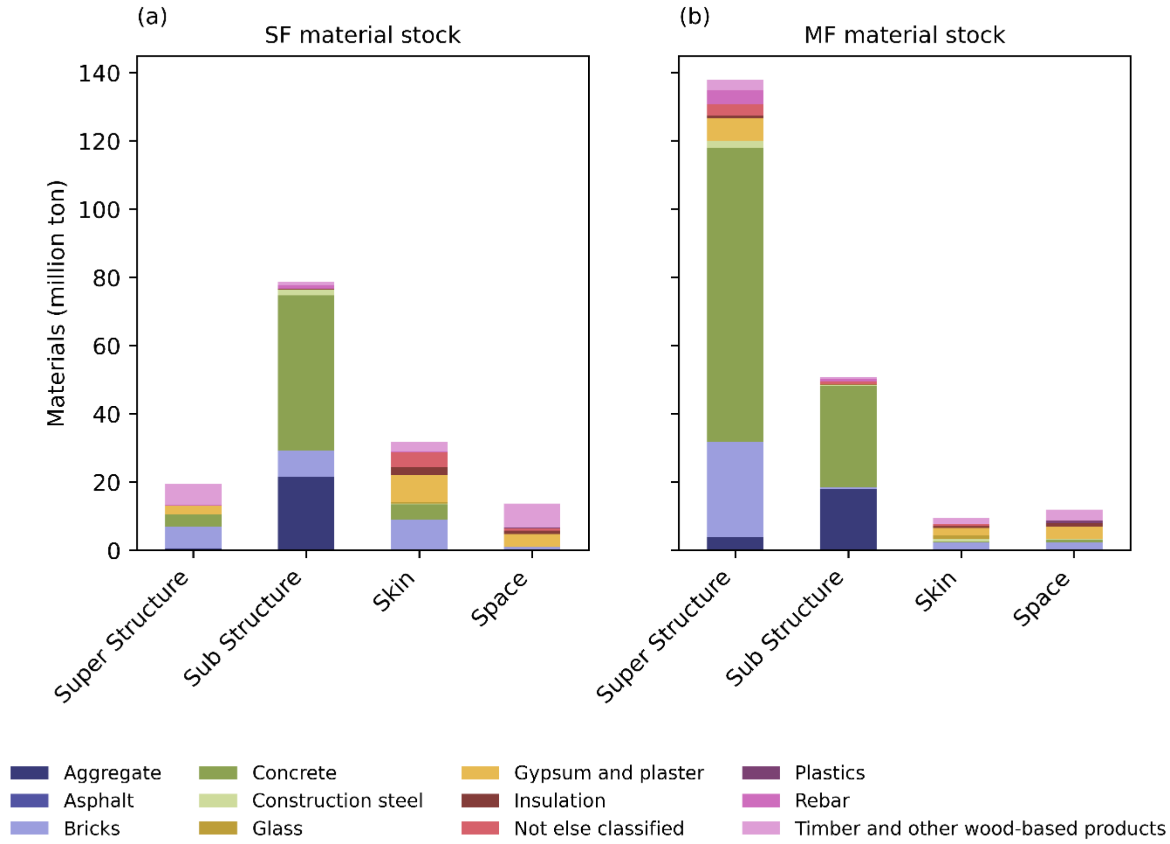

Figure 14. Material stocks for Swedish residential buildings. The SF material stock (a) and MF material stock (b) are shown to demonstrate the differences between the material stock distributions across layers. Source: **Paper III**.

The detailed material composition of the material stock in each building layer for the SF and MF buildings in Year 2024 is shown in **Figure 14**. The distribution of material stocks across the layers displays different patterns for the SF and MF buildings. For both, the structure layer stocks contain most of the construction materials: 68% of the SF stocks (98 Mt), and 90% (189 Mt) of the MF stocks. It is noteworthy that the vertical distribution differs depending on the building's use. The *sub*structure is the largest contributor to the SF stocks (79 Mt, 55% of the SF stocks), while the *super*structure dominates the MF stocks (138 Mt, 66% of the MF stocks). Regarding composition, unsurprisingly, the substructures are dominated by concrete and aggregate in both the SF and MF stocks. The superstructures show greater material variety: in the SF, bricks and timber dominate the composition, whereas bricks and concrete predominate in the MF.

Among the non-structural layers, the skin layer is the second-highest weight contributor to the SF stocks, amounting to 31 Mt (22% of the SF stocks), which is heavier than the superstructure (19 Mt, 14% of the SF stocks). There are two reasons why the skin layer contributes more than the superstructure layer to the SF stocks. First, buildings have a larger roof to usable floor area ratio and, thus, have proportionally more material in the roofs. Second, we see a large share of timber in SF superstructures. This results in the skin layer weighing more than the superstructure layer in the Swedish SF stock. In contrast, the skin layer contributes minimally to the total MF stock (9.5 Mt, 4% of the MF stocks).

As regards the space layer, it is a weak contributor to both the SF and MF stocks, amounting to 10% of the SF stocks (14 Mt) and 6% of the MF stocks (12 Mt). Material-wise, the space layer is dominated by timber, gypsum and plaster in the SF, and by gypsum and plaster in the MF (with timber having a smaller share). This reflects the function of the space layer, which includes building elements such as ceilings (timber and plaster), drywalls (timber and plaster), and flooring (timber).

The material stock for residential buildings in Sweden in 2024 was 354 Mt, which is equivalent to 33 tonnes per capita (t/cap). As shown in **Figure 15**, MF buildings make up most of the residential material stock at 210 Mt (59%), as compared with 144 Mt (41%) for SF buildings. In total, the building structures amount to 81% (286 Mt) of the total amount of Sweden's residential material stock. However, one needs to highlight the impacts of MI-related assumptions on these results.

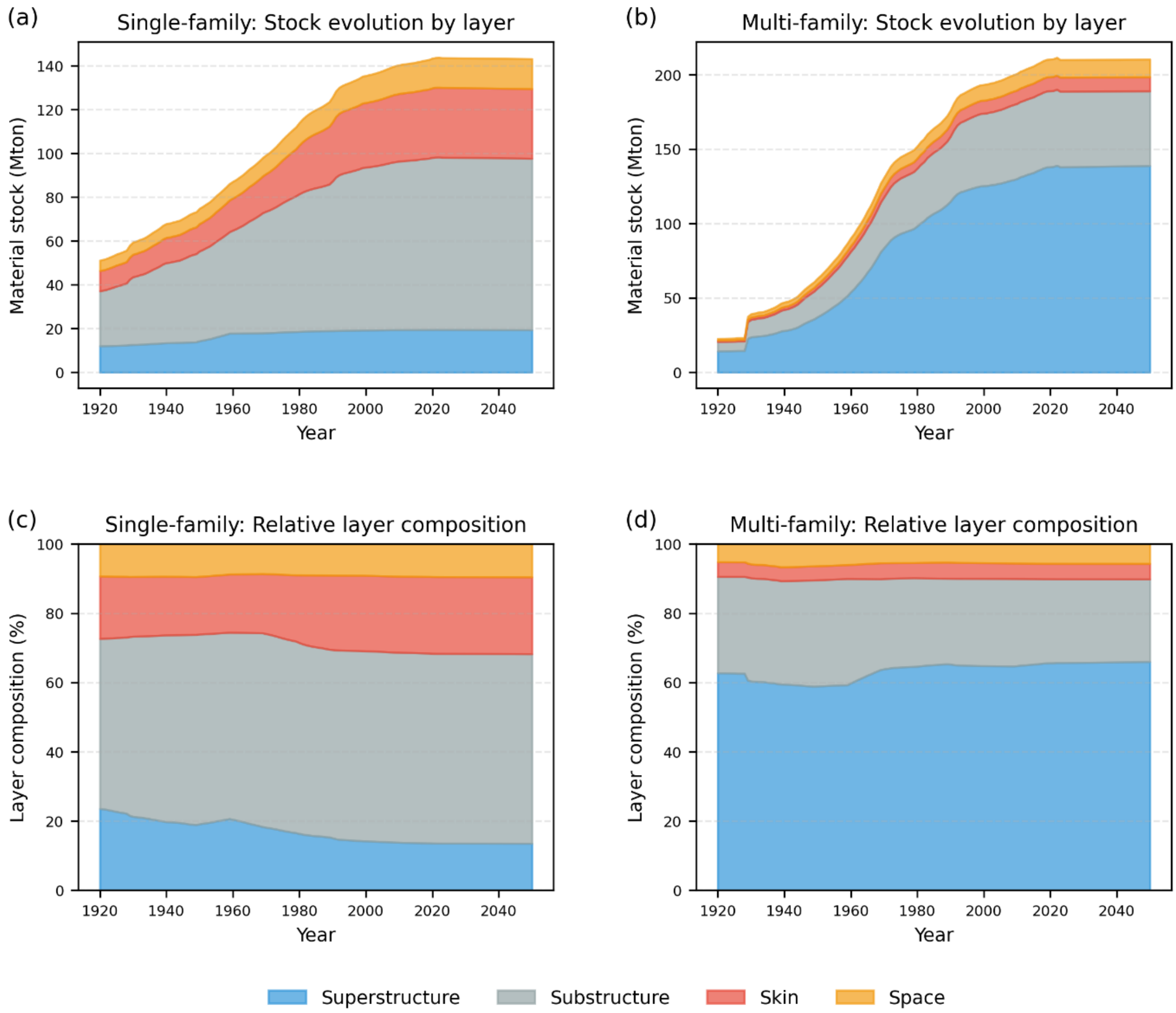

Figure 15. Evolution of the material stock over time by layer for single-family buildings (a) and multi-family buildings (b), and evolution of the material stock composition by percentage for each layer for single-family buildings (c) and multi-family buildings (d). Source: **Paper III**.

The material stock for SF buildings has increased by 112 Mt, from 65 Mt in Year 1920 to 177 Mt in Year 2024, which represents a 279% increase. In comparison, the material stock for MF stock has increased by 81 Mt, from 6 Mt in Year 1920 to 87 Mt in Year 2024, which represents a 927% increase over the same period. The differences in the percentage increases in MS stock indicate that Sweden has been building more MF buildings due to an increasing urbanization rate. The evolutions of layer composition of the material stock for SF and MF buildings are shown in **Figure 15 (c, d)**. The share of each layer as a percentage of the total MS for both SF and MF buildings is relatively stable, as compared with the evolution in total MS. For SF buildings, the share of space layer MS as total MS has remained stable, from 9.35% in 1920 to 9.6% in Year 2024, while the superstructure layer's share has declined from 24% in 1920 to 14% in 2024. The share of substructure layer of material stock increased from 49% in 1920 to

54% in 2024, and the share of skin layer has increased from 18% to 22%. The evolution of layer composition for SF buildings can be explained by the increased use of concrete in substructures, and the increased use of gypsum and plaster in the space layer. The layer composition of MF buildings is more stable compared to that in SF buildings, as the shares of superstructure, substructure, skin, and space layer have only changed from 63%, 28%, 4%, and 5% in 1920 to 66%, 24%, 4%, and 6% in 2024, respectively. The stability of the layer composition can be explained by the relatively similar techniques used for the construction of MF buildings in Sweden.

6.3 Building material flow

Figure 16 shows the material flow results for the skin and space layers in the SF and MF buildings.

The inflows of material from renovation activities from Year 2024 to Year 2050 show significantly different trends. Overall, the results indicate that SF buildings require substantially more renovation materials than MF buildings in absolute terms. In total, the levels of materials required for the renovation of SF buildings range from 1.4 Mt in 2025 to 1.6 Mt in 2050, while the levels of materials required for the renovation of MF buildings range from 587 kt in 2025 to 589 kt in 2050. However, the materials required for skin and space layer renovations differ between the SF and MF buildings. For SF buildings, the skin layer requires more materials for renovation at 880 kt in 2025 than the 511 kt required by the space layer. On the other hand, for MF buildings, the skin layer requires less material for renovation, at 262 kt in Year 2025 compared to the 325 kt required by the space layer in Year 2025. This difference can be explained by the difference in MI for the skin and space layers. In total, the MI of the skin layer for SF buildings contains 123 kg/m^2 of materials, while the MI of the space layer is only 53 kg/m^2 of materials. In contrast, the MI of the skin layer for MF buildings contains 47 kg/m^2 of materials, while the space layer contains 60 kg/m^2 of materials.

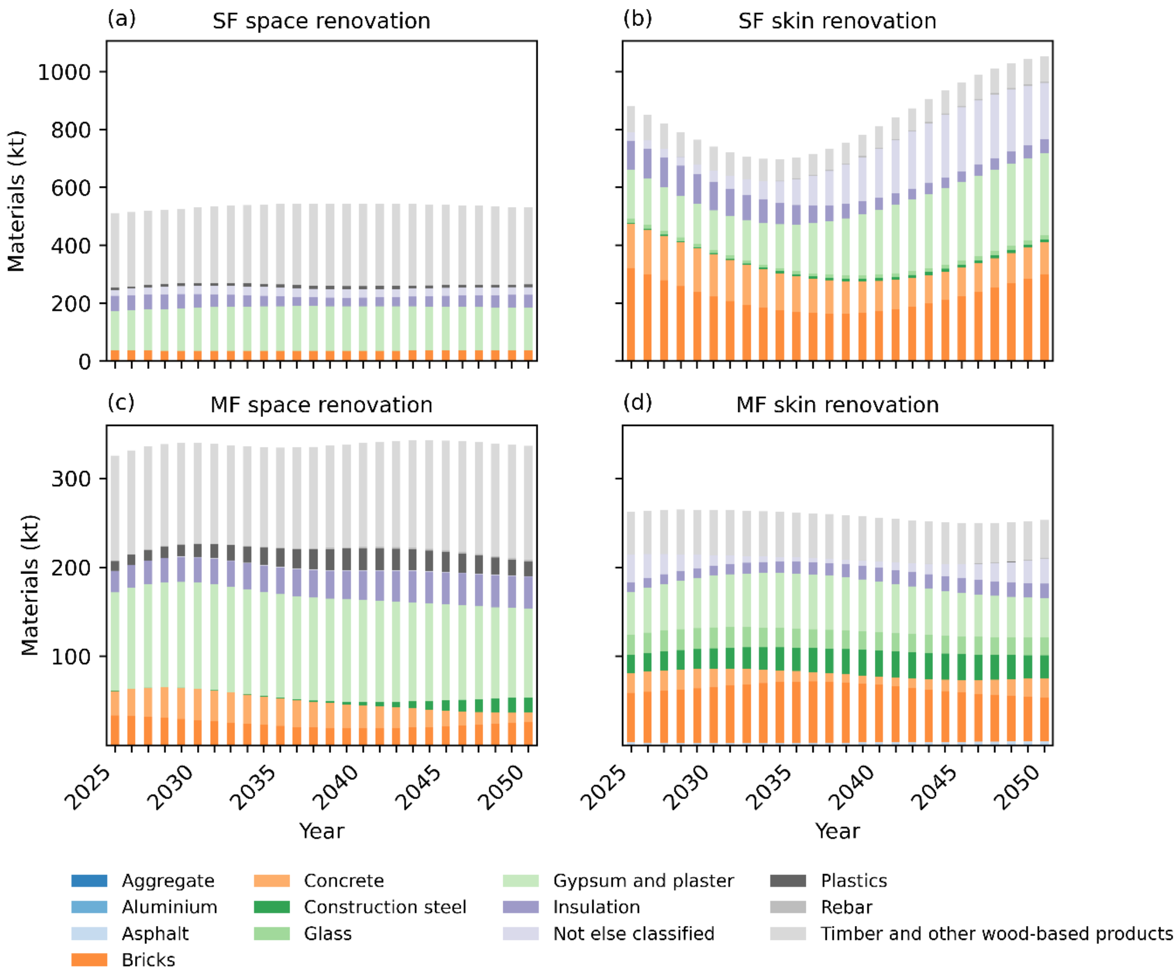

Figure 16. Inflows of materials from renovation activities for the skin and space layers of SF buildings (a and b) and MF buildings (c and d), respectively. Note that the y-axis is shared between each row and is different for the SF and MF buildings. Source: **Paper III**.

Another notable finding is that the material flows from the renovation of the space layer in SF buildings exhibit less variability compared to the other building layers. Since the same lifetime distribution is used for the space layers of the SF and MF buildings, this variability difference is likely due to the construction year distributions of the SF buildings.

To test the hypothesis of this study, we compare the inflows of floor area for renovations between the monolithic model and the LRM, as shown in **Figure 17**. The results are presented separately for SF and MF buildings, to highlight the effect of the age-cohort composition on the framework, given that the magnitude of the differences in results of renovation activities differs based on the age-cohort composition of the model building stock.

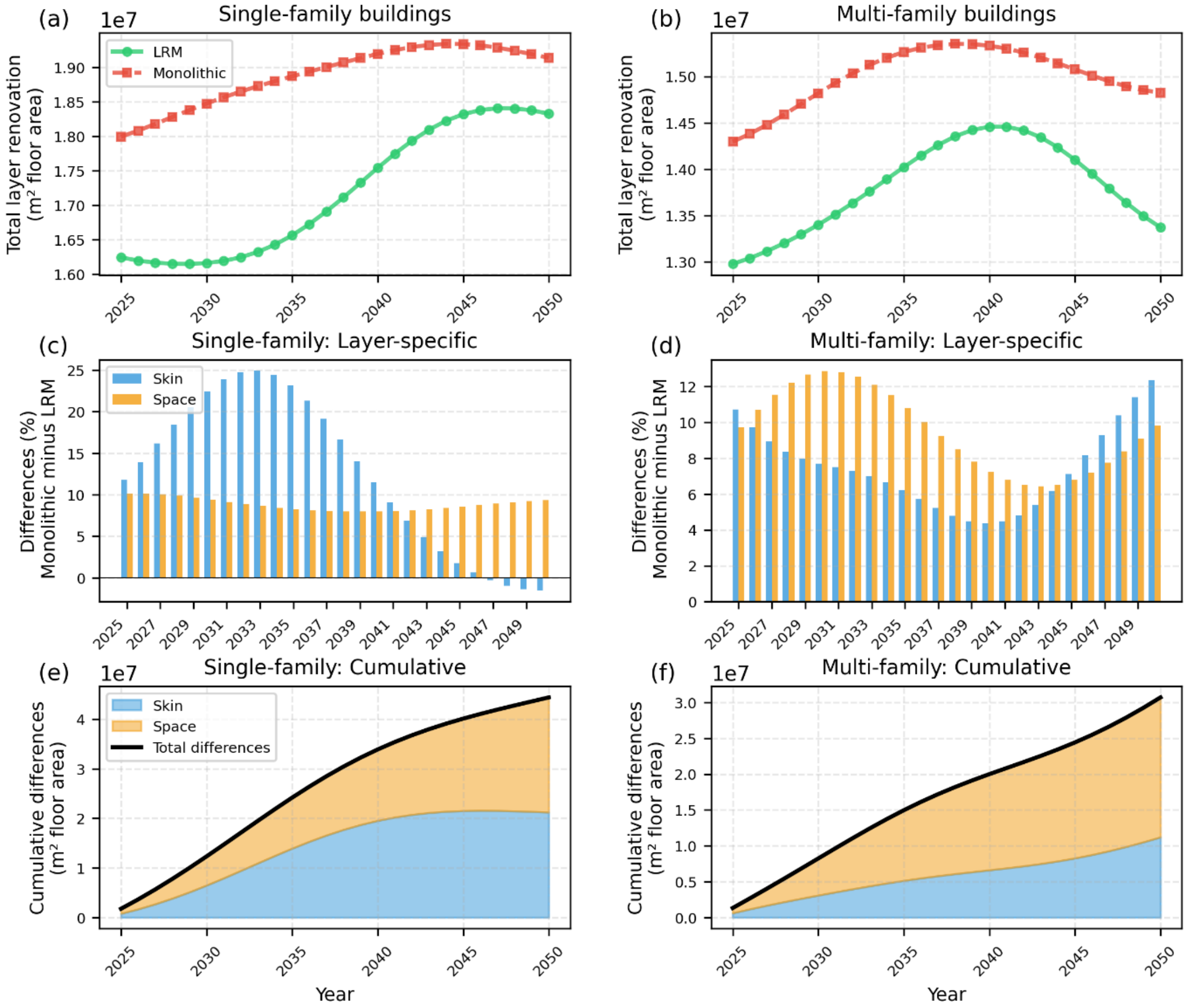

Figure 17. Comparison of the inflows of floor area from renovation activities between the monolithic and LRM (a and b). Shown are the percentages differences of renovations from the monolithic model for each layer (c and d), and the cumulative differences of renovated floor area (e and f). Source: **Paper III**.

The total amounts of renovation (in m^2 of GFA) from Year 2025 to Year 2050 for SF and MF buildings are shown in **Figure 17 (a and b)**.

For SF buildings, it is clear from Figure 6a that the results from the layered model follow the renovation cycles, while the results from the monolithic model do not. In Year 2050, the difference in GFA renovated is around 800,000 m^2, which represents a 4.2% difference. This rather small difference is most likely due to the relatively small proportion of SF buildings that are close to demolition, and thus the layered model is not renovating these buildings.

The trend in the results is more similar between the monolithic model and the layered model for the MF buildings. The difference in GFA renovated in Year 2050 for MF buildings is 1.3 million m^2, which is only an 8% difference. This difference is most likely due to MF buildings on average being built later and thus being further in time from demolition.

The differences in results from the monolithic model for each layer are presented in **Figure 17 (c and d)**. For SF buildings, on average over the modeled period, the monolithic model shows higher renovated GFA of 12.6% and 8.8% for the skin layer and space layer, respectively. For MF buildings, on average, the monolithic model shows higher renovated GFA of 7.4% and 9.5% for the skin layer and space layer, respectively. Lastly, the cumulative differences in renovated GFA from the monolithic model are presented in **Figure 17, e and f**. The differences in renovated GFA values for SF and MF buildings in Year 2050 are 44 million m^2 and 30 million m^2, respectively. To place these numbers in context, the differences represent 14% and 15% of the total GFA for SF and MF buildings, respectively.

CHAPTER 7

Carbon emissions and biogenic carbon storage

This section presents the result on carbon emissions and biogenic carbon storage answering **RQ IV**. The national level embodied results are shown first followed by results from **Paper IV** that are at an urban level for the city of Uddevalla. The road embodied emissions are taken from **Paper I,** while further analysis based on **Paper III** is carried out for residential buildings in the thesis. In **Paper III**, no new construction or stock growth is assumed to better demonstrate the LRM. However, it is important to put the embodied emissions from renovations into context by comparing the emissions with embodied emissions from new construction. Lastly, the operational emissions and biogenic carbon storage results related to **RQ V** are shown.

7.1 National level emissions

The comparison between the embodied emissions from renovation/maintenance compared to new construction is presented in **Figure 18**. To model the embodied emissions from new construction of residential buildings, a relative ambitious new construction scenario is used. The new construction for buildings is calculated by multiplying an assumed floor space per capita with projected population growth every year for SF and MF buildings for each municipality in Sweden. The floor space per capita is calculated using the historical average from 2019 to 2023 using data from Statistics Sweden [113]. Similarly, the population projections are also retrieved from Statistics Sweden to ensure consistency. This scenario can be considered as ambitious because the underlying assumption is that every net population addition will correspond to a newly built unit (the average floor space per capita in Sweden in 2025 is 42m^2), which might not be realistic but represents the higher range of expected new construction. The new construction for roads is calculated in a similar approach and is cited from a previous Swedish case study [114]. The material intensity for newly constructed buildings is modeled using the same material intensity as buildings built in 2020.

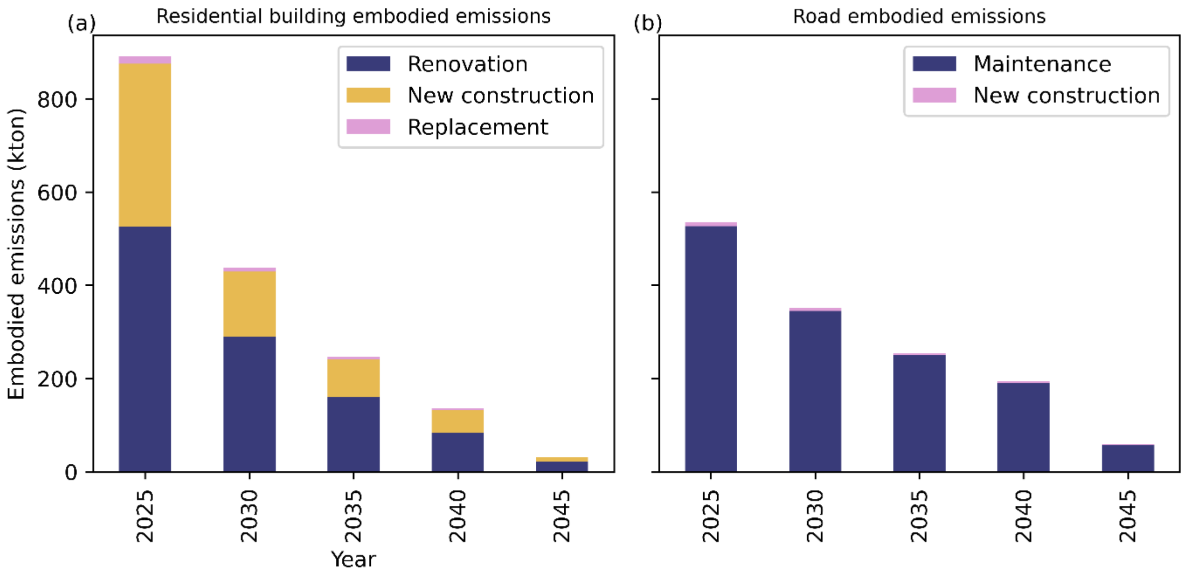

Figure 18. Embodied emissions from renovation/maintenance activities and new construction for (a) residential buildings and (b) roads. The replacement represents the replenishment of demolished building stock. Plotted specifically for the thesis.

The reduction scenario emission factors are used to demonstrate the potential for the embodied emissions to reduce overtime. The total embodied emissions for residential buildings in 2025 are approximately 830 kt, of which 59% or 490 kt is from renovation while new construction only accounts for 39% or 324 kt of emissions. Under the reduction scenario, the share of embodied emissions from renovation increases to 70% by 2045 while the share of embodied emissions from new construction decreases to 27% by 2045. The reason for this change is that the materials mainly used in new constructions such as concrete and steel are assumed to be reducing emissions at a faster rate than materials mainly used in renovation such as insulation materials. Unsurprisingly, embodied emissions from maintenance dominate the total emissions from roads. On average, the embodied emissions from new construction represent only 2% of the total emissions.

To put the numbers into perspective, in 2025 the Swedish economy emitted 51.2 million tons of CO_2 equivalents [113], which means that the total embodied emissions from residential buildings and roads corresponds to approximately 2.6% of the total emissions. Embodied emissions from renovation and maintenance on the other hand represents about 2% of the total emissions from the Swedish economy. The findings highlight the importance of developing or using existing policy instruments to tackle embodied emissions from renovations and maintenance of the built environment.

7.2 City level emissions

The analysis conducted in **Paper IV** narrows the geographical scope to urban level to estimate the emissions reduction potential by modeling different types of insulation materials in energy efficiency renovations to estimate both the operational and embodied emissions from renovation and the interplay between them. The reason for narrowing the scope to the city level is due to the high computational and data requirements (mainly height data) for the building energy demand modeling. The city of Uddevalla is chosen as case study due to the energy demand modeling framework is validated based on data from the city.

7.2.1 Operational vs embodied emissions

Energy efficiency renovation is important for reducing energy consumption in buildings and thus reducing the operational emissions. However, insulation materials that are required to conduct energy efficiency renovation also induces embodied emissions that could significantly hamper the operational emissions reduction especially in a low grid emissions environment. It is therefore important to understand

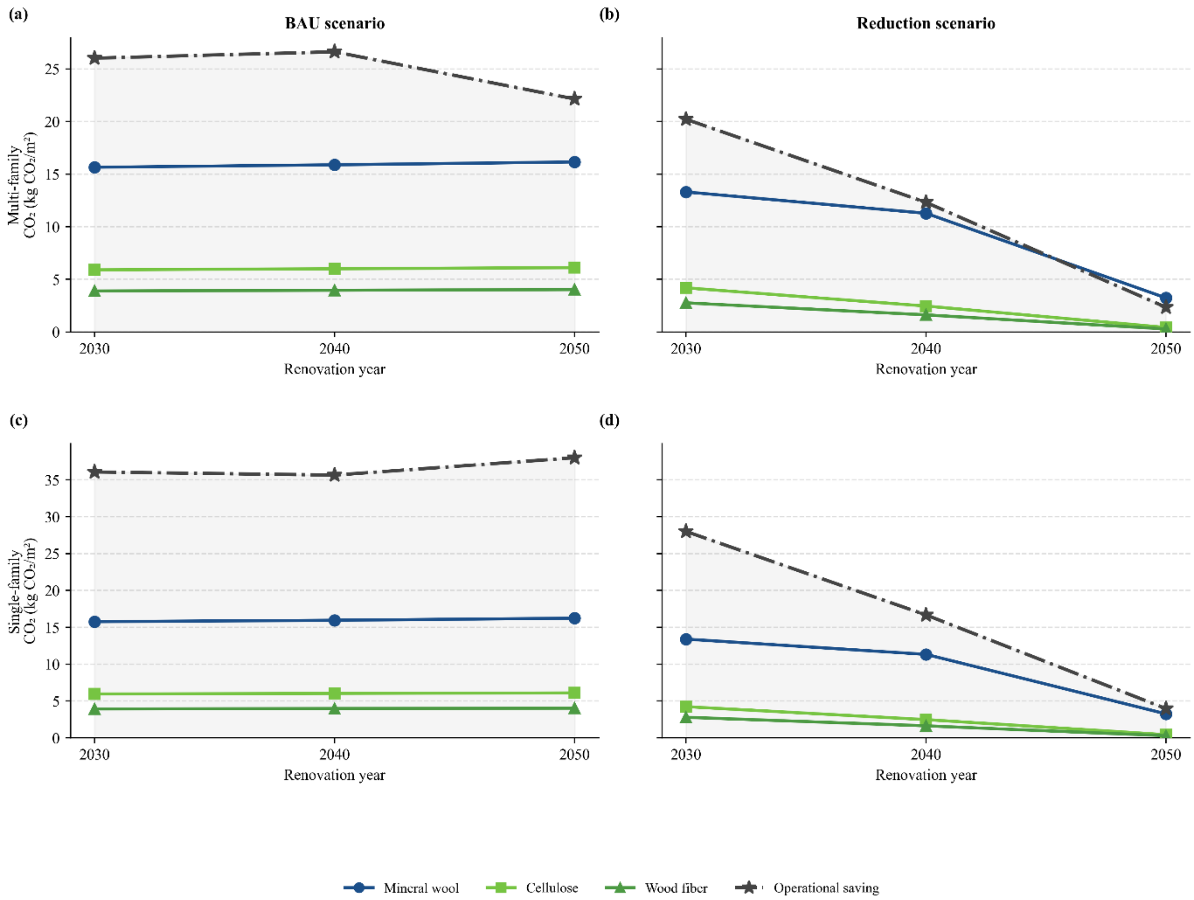

Figure 19. Comparison between the operational emissions savings and embodied emissions for (a) and (b) the business-as-usual scenario (BAU) and (c) and (d) the reduction scenario, from 2030 to 2050 for the City of Uddevalla. Operational emissions savings are calculated by multiplying the heat demand reduction per year multiplied by the emission factors for electricity and district heating (DH). Source: **Paper IV**.

The comparison between the operational emissions savings from energy efficiency renovation compared to the embodied emissions from the insulation materials are shown in **Figure 19**. The operational emissions savings are calculated by multiplying the heat demand reduction per year multiplied by the emission factors for electricity and district heating (DH). The reason for plotting the savings instead of absolute operational emissions is to isolate and compare the gains (operational savings) with the costs (embodied emissions) from energy efficiency renovation.

In the BAU scenario for MF buildings, the operational emissions savings increase slightly from 26 kg.CO_2eq/m^2 in 2030 to 26.6 kg.CO_2eq/m^2 in 2040 and decline to 22.1 kg.CO2eq /m^2 in 2050, which corresponds to a 17% decrease. The difference in operational savings is driven by the building stock renovation dynamics. A slightly different trend can be observed for the operational emissions savings in the BAU scenario for SF buildings. The operational emissions savings stay roughly constant at 36 kg.CO2eq/m^2 in 2030 and 2040, then increasing to 38 kgCO2eq/m^2 in 2050. The same goes for the embodied emissions, which stay roughly constant for all three materials, with the average embodied emissions being 16, 6, and 4 kg CO2eq/m^2 respectively for mineral wool, cellulose and wood fiber insulations for both SF and MF buildings. The result indicates that wood fiber insulation is the least emission intensive due to the low embodied emissions, and the embodied emission for mineral wool is approximately four times higher.

In the reduction scenario, the operational emissions savings are reduced from 20.2 to 12.3 and then to 2.3 kg.CO2eq/m^2 from 2030 to 2040 and then to 2050 for MF buildings. This reduction corresponds to an 88% decline compared to the baseline. The reduction is due to the declining emission factors for DH and electricity. For SF buildings, the reduction in operational emissions savings ranges from 28 to 16.7 and then to 4 kg.CO2eq/m^2 from 2030 to 2040 and then to 2050. Operational emissions savings from energy efficiency renovation decreases as the grid decarbonizes. Embodied emissions are also declining, with mineral wool's emissions reducing from 13.3 kg/m^2 in 2030, to 11.3 kg CO2eq/m^2 in 2040, and lastly 3.2 kg CO2eq /m^2 in 2050 for SF and MF buildings. For the biobased insulation materials, the embodied emissions of cellulose are reduced from 4,2 kg CO2eq /m^2 to 2.5 kg CO2eq /m^2 and then to 0.4 kg CO2eq /m^2, while the embodied emissions of wood fiber are reduced from 2.7 kg/m^2 to 1.6 kg CO2eq /m^2 and eventually to 0.3 kg CO2eq /m^2 in 2050 for both MF and SF buildings. The main finding of the results is that in the reduction scenario the embodied emission from mineral wool is almost the same as the operational emissions savings from energy efficiency renovation for SF buildings and exceeds the operational emissions savings for MF buildings in 2050. The implication of the result is that under the reduction scenario mineral wool should be phased out as a choice of insulation for energy efficiency renovations by 2050.

7.2.2 Carbon storage potential

Lastly, the biogenic carbon storage potential of biobased insulation materials is estimated using the dynamic GWPbio method. The results of the cumulative embodied emissions for all insulation materials and the biogenic carbon storage are presented in **Figure 20**.

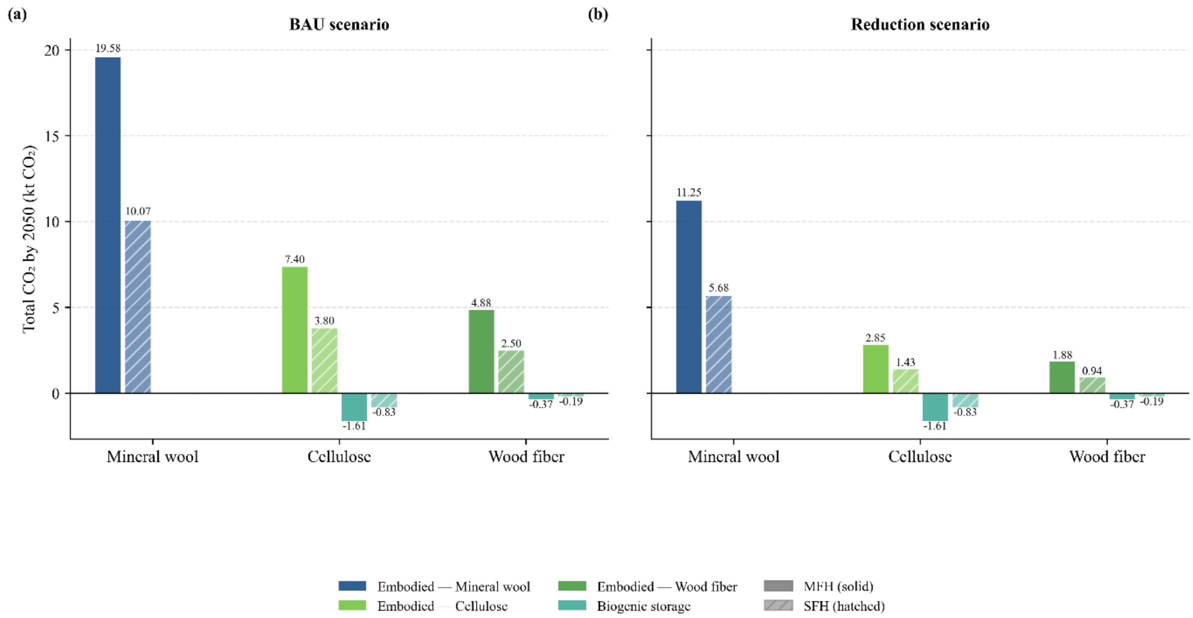

Figure 20. Cumulative embodied from 2025 to 2045 carbon for all insulation materials for the (a) BAU and (b) reduction scenarios. The negative values represent the cumulative biogenic carbon storage for biobased insulation materials. The solid bars represent multi-family buildings, and the hatched bars represent single family buildings. Source: **Paper IV**.

The total embodied emissions for mineral, cellulose, and wood fiber are 19.6 kt CO2eq, 7.4 kt CO2eq, and 4.9 kt CO2eq, respectively for MF buildings. For SF buildings, the total embodied emissions are lower on average at 10 kt CO2eq, 3.8 kt CO2eq, and 2.5 kt CO2eq for mineral wool, cellulose, and wood fiber. On the other hand, in the reduction scenario the total embodied carbon for mineral wool, cellulose, and wood fiber is 11.3 kt CO2eq, 2.9 kt CO2eq, and 1.9 kt CO2eq respectively for MF buildings and 5.7 kt CO2eq, 1.4 kt CO2eq, and 0.9 kt CO2eq for SF buildings.

The carbon storage potential is not affected by the scenarios as the scenarios only affect operational and embodied emissions, and the total storage is 2.4 kt CO2eq for cellulose and 0.5 kt CO2eq for wood fiber insulation. The embodied emissions are thus offset by 22% in the BAU scenario and 57% in the reduction scenario for cellulose and 8% in the BAU scenario and 20% in the reduction scenario for wood fiber insulation. The main result is that in the reduction scenario, cellulose insulations have a lower overall emission at 1.8 kt CO2eq compared to the 2.3 kt CO2eq overall emissions of wood fiber. The result highlights that even when using biobased insulation materials and dynamically accounting and attributing for biogenic carbon storage, embodied emissions remain a significant challenge that needs to be added.

CHAPTER 8

Discussion and future work

This section provides a discussion around the method development and results presented in the thesis as well as an outlook on potential future research for this work.

8.1 Reflection on AI models and coding

The modeling framework was developed over the course of nearly five years, and one of the main new technology developments that has had a profound impact on research has been the rise of Large Language Models (LLMs) such as ChatGPT (therein referred to as AI). The machine learning (ML) research done in this work (**Paper I** and **II**) was carried out before these models became widely available, and this subsection discusses some takeaways from learning to develop machine learning models without the help of AI and how AI impacts the usability of ML for Industrial Ecology researchers.

In the beginning of the ML development process, the biggest challenge was the learning curve. With the current capabilities of AI models, it is entirely possible to give an AI model a dataset and ask it to come up with a full ML pipeline. The ease of use of AI models could make ML much more assessable as a tool, but it will also be difficult to validate the results. There is also a risk of not understanding the dataset distribution. A key takeaway learning from the model development process carried out in this work is that the most important step in any stock modeling and/or machine learning study is to explore and understand the dataset, and this step should not be entirely outsourced to any AI model. AI could be useful in providing quick summary statistics, but the preprocessing steps such as deciding what features to keep or drop or what resampling method to use should be done manually.

At the time of writing this thesis, AI models in general should be used as an assistant but not as an independent agent when it comes to coding. As discussed in Donati et al. [115], AI models should be used in a way that is transparent and trustworthy, and for the researcher community to be able to trust the result from a study, the author needs to be able to trust the model. This trust can only come from understanding the model we are working with. On the flipside, AI can be incredibly useful as a guide by asking the AI how to do something instead of letting the AI do the coding directly. Instead of reading through documentation, the AI assistant could answer a certain research question directly with precision. The overall reflection on AI usage for research and coding in general is that AI models are highly productive tools if used in an appropriate way, and the research community should decide as a collective what is appropriate without hindering productivity gain.

8.2 Methodological framework and generalizability

The main contribution of this work is to develop a methodological framework to estimate the emissions from road and building renovations. This subsection discusses the generalizability of the model, which is important for both the research community in general and to support evidence-based decision making [116].

The datasets collected and generated is a large part of this work, and the dataset is easily generalizable for other research projects focusing on Sweden. The goal of the research work is to create a national level dataset for residential buildings and to a lesser extent roads that can be used for other research projects. For example, the building stock inventory dataset generated in **Paper II** can be used for building energy demand modeling as demonstrated in **Paper IV**. The dataset can also be used for other types of geospatially explicit analysis such as estimating the potential for rooftop solar PV panels or modeling low-voltage grid capacity [117].

The ML models developed in this work to predict building attributes are not directly transferable, as each input dataset to a ML model is unique and thus a specific training and testing process is required. A potential solution to the lack of transferability is a method called transfer learning, which can be used to apply an existing ML pipeline to a similar dataset [118]. Although the exact model pipeline and parameters cannot be directly transferred to another area of study, the use of urban form features is easily testable and transferable. The predictive power of urban form features should be more widely tested, especially where reliable building height data are not available. For the road width prediction ML model, the transferability is higher as road datasets tend to be more homogeneous and could be easily tested on datasets such as OpenStreetMap.

The layer renovation model is developed as a transferable framework, but a key requirement for applying the LRM is a set of BUD-MI (refer to Chapter 2) or a similar type of material intensity data that can be disaggregated across building layers. The collection process of such MI dataset is labor intensive and is not currently widely available.

Lastly, the building energy demand and renovation modeling framework can be easily transferred to other study areas. The two core models (LRM and EPSM) are both designed to be used in different contexts, and the key data inputs such as TABULA and the climate database are both opensource.

8.3 Implications for policy makers

The findings of this study can inform policy development and decision-making at both the national and municipal levels. The findings also highlight the importance of embodied emissions from renovation activities for roads and especially for residential buildings. At the time of writing, the Energy Performance of Buildings Directive (EPBD) introduced by the European commission does not enforce emissions requirements for renovations [20]. In the Swedish context, the Swedish National Board of Housing, Building and Planning (Boverket) acknowledges the importance of emissions from renovations, but the limit values for emissions currently does not include lifecycle stages related to renovation [21]. The results from this work add to the existing evidence that embodied emissions from renovations and especially related to energy efficiency renovations need to be more emphasized. In addition, this work challenges the conclusion by Boverket that A1 – A5 stages 'covers the vast majority of climate impact that is taking place today' [21], which might be true at an individual level but not at the overall national building stock level.

In addition, limit value could be a useful policy instrument to promote the decarbonization of insulation materials such as mineral wool, similar to the effects of a procurement requirement [119]. A limit value for insulation materials could also incentivize the deployment of biobased insulation materials, which has potential to store carbon and thus provide even greater emission reductions potential. Another policy implication related to energy efficiency renovation highlighted by the results is the need to consider embodied emissions related to energy efficiency renovations. In a low grid emission country such as Sweden, enforcing a high energy efficiency renovation rate across the building stock might not be the most effective way to decarbonize the building stock. The high upfront embodied emissions make the payback period considerably long for traditional insulation materials such as mineral wool while not significantly reducing operational emissions.

For the Swedish Transport Administration, the same implication can be concluded which is that more stringent procurement requirements should be used to accelerate the decarbonization of the transport infrastructures in Sweden as well as key industries such as asphalt and steel.

8.4 Implications for research and future work

The findings from this work also have implications for the Industrial Ecology research community in general, and particularly for researchers studying material stocks and flows, by highlighting that data availability remains a critical bottleneck despite rapid advances in data science methods and AI models. Machine learning has emerged as a useful tool to predict missing data, but collecting and obtaining primary data is the main barrier for stock and flow research. For example, primary data on building inventory data is very limited for non-residential buildings in Sweden. The ML framework developed in this work could easily be applied to non-residential buildings if similar building registry data are available. Another aspect to consider is whether data set predicted by machine learning models should be used as training input data due to issues such as error propagation.

The research community should develop a streamlined template or guideline on how to train ML models similar to guidelines related to AI usage. In addition, it would be beneficial to develop a standard template on what to report for ML workflows so that the models are transparent and easily comparable across different works.

For building stock modeling specifically, more effort should be directed towards detailed material intensity data collection such as BUD-MI to facilitate more detailed material stock and flow modeling. The challenge remains that the process is highly labor-intensive, and each country or region requires different approaches that are hard to automate. Machine learning or AI models could be developed and tested to extract MI data at scale, enabling faster data collection and analysis.

In terms of future work, the first potential research direction is to incorporate circularity potential assessment into the existing work, especially to the layer renovation model. The BUD-MI could be disaggregated into building component levels to investigate the potential to reuse or recycle components. Circularity potentials are important both from the resource efficiency perspective as well as reducing embodied emissions, and the interplay between circularity measures and emissions have been shown to be significant [19].

The next potential addition to the work is the inclusion of costs to make the results more comprehensive for decision making purposes. Cost is a key perspective in renovation projects, especially related to which materials are the most cost-effective. The building energy efficiency renovation model can further develop into an optimization model with costs for different insulation materials and renovation measures.

Another future direction for the research would be to model the material stock and flow of non-residential buildings. Non-residential buildings remain a key knowledge gap, mainly due to the lack of both inventory and material intensity data. By narrowing the scope of the study to an urban level combined with the use of methods such as street view image recognition could be one way to model non-residential building stock without official data. It is important to understand what orders of magnitude the material stock and embodied emissions are from constructing non-residential buildings to fully decarbonize the Swedish building stock.

Finally, a future research direction is to look into more demand side measures such as sufficiency and degrowth [120], [121]. The work conducted in this thesis focuses on supply side measures to reduce emissions while assuming either no growth or linear growth for the stock, which can be questioned if it is the most compatible with the vision of a sustainable future. The highly detailed building stock inventory could be used to investigate questions such as converting existing buildings instead of new construction, or be used to model decent living standards at a high geographical resolution using the Lorenz curve suggested by Pauliuk [122].

8.5 Limitations and uncertainties

This subsection discusses the limitations and uncertainties from each appended paper. The main limitations of **Paper I** are that dissipative flow and ground reinforcements are not modeled. Dissipative flow from roads are the results of wear and tear and especially from studded tires in winter [123], but no such data is openly available and thus dissipative flow were not included in the analysis. The inclusion of dissipative flows would increase the need for maintenance and thus embodied emissions. Ground reinforcements are for example concrete pillars that are sometimes used in the capping layer of the road, but the need for ground reinforcements are dependent on geological conditions which are not available. The exclusion of ground reinforcements means that the embodied emissions from new construction could be underestimated, but the exclusion does not affect maintenance.

The limitation of **Paper II** as previously discussed is that the ML models are not directly transferable to other geographical areas without retraining. Previously described transfer learning could be one way to test the transferability of the models.

The main limitation of **Paper III** is that the LRM does not manage to model the extension of structural lifetime through skin and space layer renovation. The goal of the LRM is to ensure that the renovation probabilities are coupled with demolition probability by modeling that renovation only happens to the surviving stock. The result of this assumption is that demolition has already happened before any renovation takes place in a year. In another sense, the LRM is attempting to capture the coupling between the physical structural lifetime of a building and renovation but not the economic lifetime of a building and renovation. Renovation in reality might extend the economic lifetime of a building by making it more desirable to live in, but if the building is reaching its physical end-of-life and needs to be demolished due to engineering and safety reasons, renovations are unlikely to extend the lifetime of the building. Modeling renovations in a way that extends lifetime is much more complicated, but could be more useful in studies regarding, for example energy efficiency renovations and the LRM is not able to capture that dynamic.

Lastly, the main limitation of **Paper IV** is the lack of renovation state data for the building stock. This lack of data limits the ability for the model to make building specific recommendations and policy conclusions. In the future, when more complete building energy performance certificate data becomes available, such data should be used as the renovation status in the model. In addition, the renovation status data would ideally also indicate the type of renovation measure that was completed.

A key source of uncertainty across the dMFA modeling in **Papers III** and **IV** is the choice of lifetime data. The lifetime distributions for skin and space layers are taken from literature [124], while empirical data for residential building lifetimes, especially at the building layer level is scarce. The choice of lifetime has a direct impact on the results, especially for embodied emissions. In addition, the lifetime that is calculated based on demolition rate of residential buildings are drastically longer than what is commonly assumed in literature [102], with the exception of historical city center in a case study for Padua [125]. The building stock research community could benefit from conducting future lifetime comparison studies.

CHAPTER 9

Conclusion

This thesis develops a bottom-up modeling framework to assess the material flows and embodied emissions from maintenance and renovation of Swedish roads and residential buildings with limited data. Machine learning models are trained and applied to impute missing road width, building usable floor area and construction year. Subsequently, existing dynamic material stock and flow methods to model building renovation are developed to incorporate the shearing layer concept as to more accurately capture renovation dynamics. The application of the layer renovation model is to estimate the potential to reduce embodied emissions from renovations. Lastly, the model framework is expanded to investigate the change in energy demand from energy efficiency renovations and the climate benefits of using biobased insulation materials.

From a method development perspective, the results show that machine learning models can be effectively utilized to predict missing inventory data, even in absence of key features such as height for buildings. However, more effort is required to improve the generalizability of machine learning models for material stock modeling. Furthermore, the results from the layer renovation model theorizes that existing monolithic models might be overestimating the number of renovations as the monolithic model does not consider the lifetime extension effects of renovations. In terms of embodied emissions from renovations, biobased insulation materials show promising results compared to mineral wool, and the result also highlights that the main benefit of using biobased insulation materials is the low embodied emissions and not the biogenic carbon storage potential in the Swedish context.

To conclude, the main takeaway of this thesis is that the decarbonization of the Swedish building and construction sector remains a challenge. Despite the progress in research and policy making to limit emissions from new constructions, maintenance and renovation remains an area that needs to be tackled. In particular, countries with low operational emissions such a Sweden should pay attention to the trade-offs between embodied and operational emissions for building renovations.

References:

[1] Building fast. Falling short. As climate risks rise and cities grow, we must rethink how we build to create better lives for all - Global Status Report for Buildings and Construction 2025/26. United Nations Environment Programme, 2026. doi: 10.59117/20.500.11822/49531.

[2] UNEP, "2022 Building Construction Global Status Report," U. N. Environ. Programme, p. v, 2022.

[3] C. E. Caballero-Güereca et al., "Towards net zero: sectoral ambitions and global trends in building decarbonisation," Build. Cities, vol. 7, no. 1, pp. 396–424, Mar. 2026, doi: 10.5334/bc.724.

[4] M. Röck, E. Baldereschi, E. Verellen, A. Passer, S. Sala, and K. Allacker, "Environmental Modelling of Building Stocks – An Integrated Review of Life Cycle-Based Assessment Models to Support EU Policy Making," Renew. Sustain. Energy Rev., vol. 151, p. 111550, 2021, doi: 10.1016/j.rser.2021.111550.

[5] D. B. Müller, "Stock dynamics for forecasting material flows—Case study for housing in The Netherlands," Ecol. Econ., vol. 59, no. 1, pp. 142–156, 2006, doi: 10.1016/j.ecolecon.2005.09.025.

[6] V. Augiseau and S. Barles, "Studying Construction Materials Flows and Stock: A Review," Resour. Conserv. Recycl., vol. 123, pp. 153–164, 2017, doi: 10.1016/j.resconrec.2016.09.002.

[7] H. Tanikawa, T. Fishman, K. Okuoka, and K. Sugimoto, "The Weight of Society Over Time and Space: A Comprehensive Account of the Construction Material Stock of Japan, 1945-2010: The Construction Material Stock of Japan," J. Ind. Ecol., vol. 19, no. 5, pp. 778–791, 2015, doi: 10.1111/jiec.12284.

[8] N. Alaux et al., "Whole-life greenhouse gas emission reduction and removal strategies for buildings: Impacts and diffusion potentials across EU Member States," J. Environ. Manage., vol. 370, no. September, 2024, doi: 10.1016/j.jenvman.2024.122915.

[9] I. Karlsson, J. Rootzén, F. Johnsson, and M. Erlandsson, "Achieving Net-Zero Carbon Emissions in Construction Supply Chains – A Multidimensional Analysis of Residential Building Systems," Dev. Built Environ., vol. 8, p. 100059, 2021, doi: 10.1016/j.dibe.2021.100059.

[10] N. Alaux, B. Steubing, G. Habert, M. Ruschi Mendes Saade, and A. Passer, "Future Life-Cycle Greenhouse Gas Emission Scenarios for the Austrian Building Stock: A Systematic Approach," Environ. Sci. Technol., 2025, doi: 10.1021/acs.est.4c12138.

[11] G. Habert et al., "Carbon budgets for buildings: harmonising temporal, spatial and sectoral dimensions," Build. Cities, vol. 1, no. 1, pp. 429–452, 2020, doi: 10.5334/bc.47.

[12] S. O. S. E. zu Ermgassen et al., "A home for all within planetary boundaries: Pathways for meeting England's housing needs without transgressing national climate and biodiversity goals," Ecol. Econ., vol. 201, no. August, 2022, doi: 10.1016/j.ecolecon.2022.107562.

[13] Y. Liu, W.-Q. Chen, T. Lin, and L. Gao, "How Spatial Analysis Can Help Enhance Material Stocks and Flows Analysis?," Resources, vol. 8, no. 1, p. 46, Mar. 2019, doi: 10.3390/resources8010046.

[14] C. Fu, Y. Zhang, T. Deng, and I. Daigo, "The Evolution of Material Stock Research: From Exploring to Rising to Hot Studies," J. Ind. Ecol., p. jiec.13195, 2021, doi: 10.1111/jiec.13195.

[15] D. Rajaratnam, R. A. Stewart, T. Liu, and A. S. Vieira, "Building stock mining for a circular economy: A systematic review on application of GIS and remote sensing," Resour. Conserv. Recycl. Adv., vol. 18, no. March, p. 200144, 2023, doi: 10.1016/j.rcradv.2023.200144.

[16] K. Skillington, R. H. Crawford, G. Warren-Myers, and K. Davidson, "A review of existing policy for reducing embodied energy and greenhouse gas emissions of buildings," Energy Policy, vol. 168, p. 112920, Sep. 2022, doi: 10.1016/j.enpol.2022.112920.

[17] T. Ibn-Mohammed, R. Greenough, S. Taylor, L. Ozawa-Meida, and A. Acquaye, "Operational vs. embodied emissions in buildings—A review of current trends," Energy Build., vol. 66, pp. 232–245, 2013, doi: 10.1016/j.enbuild.2013.07.026.

[18] M. Röck et al., "Embodied GHG emissions of buildings – The hidden challenge for effective climate change mitigation," Appl. Energy, vol. 258, p. 114107, 2020, doi: 10.1016/j.apenergy.2019.114107.

[19] C. Zhang et al., "Revealing the Interplay between Decarbonisation, Circularity, and Cost-effectiveness in Building Energy Renovation," Nat. Commun., 2025.
[20] European Commission, "Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the energy performance of buildings (recast) (Text with EEA relevance)," 2025, [Online]. Available: https://energy.ec.europa.eu/topics/energy-efficiency/energy-performance-buildings/energy-performance-buildings-directive_en
[21] Boverket., Limit values for climate impact from buildings and an expanded climate declaration. 2023.
[22] M. Lanau et al., "Taking Stock of Built Environment Stock Studies: Progress and Prospects," Env. Sci Technol, vol. 53, no. 15, pp. 8499–8515, 2019, doi: 10.1021/acs.est.8b06652.
[23] C. Fu, Y. Zhang, T. Deng, and I. Daigo, "The Evolution of Material Stock Research: From Exploring to Rising to Hot Studies," J. Ind. Ecol., p. jiec.13195, 2021, doi: 10.1111/jiec.13195.
[24] M. Kavgic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic, "A review of bottom-up building stock models for energy consumption in the residential sector," Build. Environ., vol. 45, no. 7, pp. 1683–1697, 2010, doi: 10.1016/j.buildenv.2010.01.021.
[25] H. Tanikawa, T. Fishman, K. Okuoka, and K. Sugimoto, "The Weight of Society Over Time and Space: A Comprehensive Account of the Construction Material Stock of Japan, 1945-2010: The Construction Material Stock of Japan," J. Ind. Ecol., vol. 19, no. 5, pp. 778–791, 2015, doi: 10.1111/jiec.12284.
[26] E. Müller, L. M. Hilty, R. Widmer, M. Schluep, and M. Faulstich, "Modeling Metal Stocks and Flows: A Review of Dynamic Material Flow Analysis Methods," Env. Sci Technol, vol. 48, no. 4, pp. 2102–2113, 2014, doi: 10.1021/es403506a.
[27] N. Heeren and T. Fishman, "A Database Seed for a Community-Driven Material Intensity Research Platform," Sci Data, vol. 6, no. 1, p. 23, 2019, doi: 10.1038/s41597-019-0021-x.
[28] S. Marinova, "Global construction materials database and stock analysis of residential buildings between 1970-2050," J. Clean. Prod., p. 13, 2020.
[29] T. Fishman, A. Mastrucci, Y. Peled, S. Saxe, and B. Van Ruijven, "RASMI: Global ranges of building material intensities differentiated by region, structure, and function," Sci. Data, vol. 11, no. 1, p. 418, Apr. 2024, doi: 10.1038/s41597-024-03190-7.
[30] P. Gontia, C. Nägeli, L. Rosado, Y. Kalmykova, and M. Österbring, "Material-Intensity Database of Residential Buildings: A Case-Study of Sweden in the International Context," Resour. Conserv. Recycl., vol. 130, pp. 228–239, 2018, doi: 10.1016/j.resconrec.2017.11.022.
[31] M. Lanau et al., "BUD-MI: a template that optimizes material intensity data collection and utilization," J. Ind. Ecol., Mar. 2026, doi: 10.1007/s44498-026-00044-w.
[32] D. Grossegger, K. MacAskill, and A. Al-Tabbaa, "A critical review of road network material stocks and flows: Current progress and what we can learn from it," Resour. Conserv. Recycl., vol. 205, no. August 2023, p. 107584, 2024, doi: 10.1016/j.resconrec.2024.107584.
[33] Trafikverket, "Klimatkalkyl – infrastrukturens klimatpåverkan och energianvändning i ett livscykelperspektiv," Trafikverket. [Online]. Available: https://www.trafikverket.se/for-dig-i-branschen/miljo---for-dig-i-branschen/energi-och-klimat/Klimatkalkyl/
[34] M. Ma et al., "Building floorspace and stock measurement: A review of global efforts, knowledge gaps, and research priorities," Nexus, vol. 2, no. 3, p. 100075, Sep. 2025, doi: 10.1016/j.ynexs.2025.100075.
[35] N. Milojevic-Dupon et al., "OPEN EUBUCCO v0.1: European building Data Descriptor stock characteristics in a common and open database for 200+ million individual buildings," 2023.
[36] European Environment Agency, "CORDA." [Online]. Available: https://corda.eea.europa.eu/_layouts/15/CustomLoginPageFBA/CustomLogin.aspx?ReturnUrl=%2F_layouts%2F15%2FAuthenticate.aspx%3FSource%3D%252Fsitepages%252Fhome%252Easpx&Source=%2Fsitepages%2Fhome.aspx
[37] European Comission, "EU Building Stock Observatory." [Online]. Available: https://ec.europa.eu/energy/eu-buildings-database_en
[38] Microsoft, "GlobalMLBuildingFootprints." [Online]. Available: https://github.com/microsoft/GlobalMLBuildingFootprints

[39] L. S. A. Rousseau, B. Kloostra, H. AzariJafari, S. Saxe, J. Gregory, and E. G. Hertwich, “Material Stock and Embodied Greenhouse Gas Emissions of Global and Urban Road Pavement,” Environ. Sci. Technol., vol. 56, no. 24, pp. 18050–18059, 2022, doi: 10.1021/acs.est.2c05255.
[40] Q. Liu, J. Rootzén, and F. Johnsson, “Development of a machine learning model to improve estimates of material stock and embodied emissions of roads,” Clean. Environ. Syst., vol. 14, no. June, 2024, doi: 10.1016/j.cesys.2024.100211.
[41] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, The elements of statistical learning: data mining, inference, and prediction, vol. 2. Springer, 2009.
[42] G. Majeau-Bettez, J. Frayret, A. Ramaswami, Y. Li, and N. Heeren, “Data innovation in industrial ecology,” J. Ind. Ecol., vol. 26, no. 1, pp. 6–11, Feb. 2022, doi: 10.1111/jiec.13256.
[43] Y. Gong et al., “The evolution of research at the intersection of industrial ecology and artificial intelligence,” J. Ind. Ecol., vol. 29, no. 2, pp. 440–457, Apr. 2025, doi: 10.1111/jiec.13612.
[44] N. Milojevic-Dupont et al., “Learning from urban form to predict building heights,” PLOS ONE, vol. 15, no. 12, p. e0242010, 2020, doi: 10.1371/journal.pone.0242010.
[45] J. H. Arehart, F. Pomponi, B. D’Amico, and W. V. Srubar, “A New Estimate of Building Floor Space in North America,” Env. Sci Technol, vol. 55, no. 8, pp. 5161–5170, 2021, doi: 10.1021/acs.est.0c05081.
[46] F. Nachtigall, N. Milojevic-Dupont, F. Wagner, and F. Creutzig, “Predicting building age from urban form at large scale,” Comput. Environ. Urban Syst., vol. 105, no. May, p. 102010, 2023, doi: 10.1016/j.compenvurbsys.2023.102010.
[47] T. R. Tooke, N. C. Coops, and J. Webster, “Predicting building ages from LiDAR data with random forests for building energy modeling,” Energy Build., vol. 68, no. PARTA, pp. 603–610, 2014, doi: 10.1016/j.enbuild.2013.10.004.
[48] E. M. Farella, E. Özdemir, and F. Remondino, “4D building reconstruction with machine learning and historical maps,” Appl. Sci. Switz., vol. 11, no. 4, pp. 1–24, 2021, doi: 10.3390/app11041445.
[49] F. Biljecki and M. Sindram, “ESTIMATING BUILDING AGE WITH 3D GIS,” ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci., vol. IV-4/W5, pp. 17–24, 2017, doi: 10.5194/isprs-annals-IV-4-W5-17-2017.
[50] J. F. Rosser, D. S. Boyd, G. Long, S. Zakhary, Y. Mao, and D. Robinson, “Predicting residential building age from map data,” Comput. Environ. Urban Syst., vol. 73, pp. 56–67, 2019, doi: 10.1016/j.compenvurbsys.2018.08.004.
[51] T. Mao, Y. Liu, W.-Q. Chen, N. Li, N. Dong, and Y. Shi, “Quantifying spatiotemporal dynamics of urban building and material metabolism by combining a random forest model and GIS-based material flow analysis,” Front. Earth Sci., p. 13, 2022.
[52] O. M. Garbasevschi et al., “Spatial factors influencing building age prediction and implications for urban residential energy modelling,” Comput. Environ. Urban Syst., vol. 88, p. 101637, 2021, doi: 10.1016/j.compenvurbsys.2021.101637.
[53] F. Biljecki, H. Ledoux, and J. Stoter, “Generating 3D city models without elevation data,” Comput. Environ. Urban Syst., vol. 64, pp. 1–18, 2017, doi: 10.1016/j.compenvurbsys.2017.01.001.
[54] F. Biljecki, J. Stoter, H. Ledoux, S. Zlatanova, and A. Çöltekin, “Applications of 3D city models: State of the art review,” ISPRS Int. J. Geo-Inf., vol. 4, no. 4, pp. 2842–2889, 2015, doi: 10.3390/ijgi4042842.
[55] T. Esch et al., “World Settlement Footprint 3D - A first three-dimensional survey of the global building stock,” Remote Sens. Environ., vol. 270, no. January, p. 112877, 2022, doi: 10.1016/j.rse.2021.112877.
[56] Y. Che et al., “3D-GloBFP : the first global three-dimensional building footprint dataset,” Earth Syst. Sci. Data, vol. 11319912, no. June, pp. 1–28, 2024.
[57] X. X. Zhu, S. Chen, F. Zhang, Y. Shi, and Y. Wang, “GlobalBuildingAtlas: an open global and complete dataset of building polygons, heights and LoD1 3D models,” Earth Syst. Sci. Data, vol. 17, no. 12, pp. 6647–6668, Dec. 2025, doi: 10.5194/essd-17-6647-2025.

[58] R. Mohammadiziazi and M. M. Bilec, "Building material stock analysis is critical for effective circular economy strategies: a comprehensive review," Environ. Res. Infrastruct. Sustain., vol. 2, no. 3, p. 032001, Sep. 2022, doi: 10.1088/2634-4505/ac6d08.

[59] A. Stephan and A. Athanassiadis, "Quantifying and Mapping Embodied Environmental Requirements of Urban Building Stocks," Build. Environ., vol. 114, pp. 187–202, 2017, doi: 10.1016/j.buildenv.2016.11.043.

[60] A. Stephan and A. Athanassiadis, "Towards a more circular construction sector: Estimating and spatialising current and future non-structural material replacement flows to maintain urban building stocks," Resour. Conserv. Recycl., vol. 129, no. April 2017, pp. 248–262, 2018, doi: 10.1016/j.resconrec.2017.09.022.

[61] I. Bertin, R. Mesnil, J.-M. Jaeger, A. Feraille, and R. Le Roy, "A BIM-Based Framework and Databank for Reusing Load-Bearing Structural Elements," Sustainability, vol. 12, no. 8, p. 3147, Apr. 2020, doi: 10.3390/su12083147.

[62] É. Gobbo, E. M. Nia, A. Straub, and A. Stephan, "Exploring the effective reuse rate of materials and elements in the construction sector," J. Build. Eng., vol. 98, p. 111344, Dec. 2024, doi: 10.1016/j.jobe.2024.111344.

[63] R. Andersen and K. Negendahl, "Lifespan prediction of existing building typologies," J. Build. Eng., vol. 65, no. July 2022, p. 105696, 2023, doi: 10.1016/j.jobe.2022.105696.

[64] F. Aguilar Lopez, R. G. Billy, and D. B. Müller, "A product–component framework for modeling stock dynamics and its application for electric vehicles and lithium-ion batteries," J. Ind. Ecol., vol. 26, no. 5, pp. 1605–1615, Oct. 2022, doi: 10.1111/jiec.13316.

[65] I. Daigo, K. Iwata, M. Oguchi, and Y. Goto, "Lifetime Distribution of Buildings Decided by Economic Situation at Demolition: D-based Lifetime Distribution," Procedia CIRP, vol. 61, pp. 146–151, 2017, doi: 10.1016/j.procir.2016.11.221.

[66] N. Džubur and D. Laner, "Evaluation of Modeling Approaches to Determine End-of-Life Flows Associated with Buildings: A Viennese Case Study on Wood and Contaminants," J. Ind. Ecol., vol. 22, no. 5, pp. 1156–1169, Oct. 2018, doi: 10.1111/jiec.12654.

[67] M. Roca-Puigròs, R. G. Billy, A. Gerber, P. Wäger, and D. B. Müller, "Pathways toward a carbon-neutral Swiss residential building stock," Build. Cities, vol. 1, no. 1, pp. 579–593, 2020, doi: 10.5334/bc.61.

[68] S. Brand, How buildings learn: What happens after they're built. Penguin, 1995.

[69] I. Sartori, N. H. Sandberg, and H. Brattebø, "Dynamic building stock modelling: General algorithm and exemplification for Norway," Energy Build., vol. 132, pp. 13–25, 2016, doi: 10.1016/j.enbuild.2016.05.098.

[70] Statistics Sweden, "Number of dwellings in demolished multi-dwelling buildings," Number of dwellings in demolished multi-dwelling buildings. [Online]. Available: https://www.scb.se/en/finding-statistics/statistics-by-subject-area/housing-construction-and-building/construction-and-conversion/statistics-on-building-permits-new-construction-and-conversion/pong/tables-and-graphs/conversion/number-of-dwellings-in-demolished-multi-dwelling-buildings/

[71] D. Stankovic et al., "Reconditioning and Reconstruction: A Second Wind for Serbian Kindergartens," Procedia Eng., vol. 117, pp. 751–765, 2015, doi: 10.1016/j.proeng.2015.08.218.

[72] Regeringskansliet, "Riksdagen antar historiskt klimatpolitiskt ramverk." [Online]. Available: http://www.regeringen.se/pressmeddelanden/2017/06/riksdagen-antar-historiskt-klimatpolitiskt-ramverk/

[73] Trafikverket, "Lastkajen – Sveriges väg- och järnvägsdata," Trafikverket. [Online]. Available: https://www.trafikverket.se/tjanster/data-kartor-och-geodatatjanster/hamta-var-oppna-data/lastkajen---sveriges-vag--och-jarnvagsdata/

[74] Lantmäteriet, "Real Property Register," Lantmateriet.se. [Online]. Available: https://www.lantmateriet.se/en/real-property/property-information-/real-property-register/

[75] Lantmäteriet, "Municipal surveying authorities." Accessed: Dec. 04, 2023. [Online]. Available: https://www.lantmateriet.se/en/about-lantmateriet/our-collaboration-with-others/municipal-surveying-authorities/

[76] M. Lanau and G. Liu, “Developing an Urban Resource Cadaster for Circular Economy: A Case of Odense, Denmark,” Env. Sci Technol, vol. 54, no. 7, pp. 4675–4685, 2020, doi: 10.1021/acs.est.9b07749.
[77] X. Bian, M. Berghauser Pont, J. Cohen, and M. Lanau, “Enhanced material stock accounting by predicting building structures with morphological indicators,” J. Ind. Ecol., May 2026, doi: 10.1007/s44498-026-00094-0.
[78] T. Loga, B. Stein, and N. Diefenbach, “TABULA building typologies in 20 European countries—Making energy-related features of residential building stocks comparable,” Energy Build., vol. 132, pp. 4–12, Nov. 2016, doi: 10.1016/j.enbuild.2016.06.094.
[79] S. Somanath, A. Hollberg, and Y. Yu, “EPSM - Energy Performance Simulation Manager,” EPSM - Energy Performance Simulation Manager.
[80] N. Japkowicz, “Learning from imbalanced data sets: a comparison of various strategies,” AAAI Workshop Learn. Imbalanced Data Sets, pp. 0–5, 2000.
[81] Boverket, “Miljonprogrammet.” [Online]. Available: https://www.boverket.se/sv/samhallsplanering/stadsutveckling/miljonprogrammet/
[82] G. Dong and H. Liu, Feature engineering for machine learning and data analytics. CRC press, 2018.
[83] A. Zheng and A. Casari, Feature engineering for machine learning: principles and techniques for data scientists. O’Reilly Media, Inc., 2018.
[84] M. Berghauser Pont et al., “The spatial distribution and frequency of street, plot and building types across five European cities,” Environ. Plan. B Urban Anal. City Sci., vol. 46, no. 7, pp. 1226–1242, 2019, doi: 10.1177/2399808319857450.
[85] M. Fleischmann, “momepy: Urban Morphology Measuring Toolkit,” J. Open Source Softw., vol. 4, no. 43, p. 1807, 2019, doi: 10.21105/joss.01807.
[86] P. Zhang, Y. Jia, and Y. Shang, “Research and application of XGBoost in imbalanced data,” Int. J. Distrib. Sens. Netw., vol. 18, no. 6, 2022, doi: 10.1177/15501329221106935.
[87] G. LemaÃŽtre, F. Nogueira, and C. K. Aridas, “Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning,” J. Mach. Learn. Res., vol. 18, no. 17, pp. 1–5, 2017.
[88] I. Mani and I. Zhang, “kNN approach to unbalanced data distributions: a case study involving information extraction,” in Proceedings of workshop on learning from imbalanced datasets, ICML, 2003, pp. 1–7.
[89] M. Kubat and S. Matwin, “Addressing the curse of imbalanced training sets: one-sided selection,” in Icml, Citeseer, 1997, p. 179.
[90] J. Laurikkala, “Improving identification of difficult small classes by balancing class distribution,” in Artificial Intelligence in Medicine: 8th Conference on Artificial Intelligence in Medicine in Europe, AIME 2001 Cascais, Portugal, July 1–4, 2001, Proceedings 8, Springer, 2001, pp. 63–66.
[91] N. Milojevic-Dupont et al., “Learning from urban form to predict building heights,” PLoS ONE, vol. 15, no. 12 December, pp. 1–22, 2020, doi: 10.1371/journal.pone.0242010.
[92] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
[93] L. Breiman, “Random forests,” Mach. Learn., vol. 45, pp. 5–32, 2001.
[94] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.
[95] G. Ke et al., “LightGBM: A highly efficient gradient boosting decision tree,” Adv. Neural Inf. Process. Syst., vol. 2017-Decem, no. Nips, pp. 3147–3155, 2017.
[96] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “Catboost: Unbiased boosting with categorical features,” Adv. Neural Inf. Process. Syst., vol. 2018-Decem, no. Section 4, pp. 6638–6648, 2018.
[97] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A Next-generation Hyperparameter Optimization Framework,” Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 2623–2631, 2019, doi: 10.1145/3292500.3330701.

[98] S. Pauliuk and N. Heeren, "ODYM—An open software framework for studying dynamic material systems: Principles, implementation, and data structures," J. Ind. Ecol., vol. 24, no. 3, pp. 446–458, 2020, doi: 10.1111/jiec.12952.
[99] X. Zhong et al., "Global greenhouse gas emissions from residential and commercial building materials and mitigation strategies to 2060," Nat. Commun., vol. 12, no. 1, p. 6126, 2021, doi: 10.1038/s41467-021-26212-z.
[100] S. Deetman, S. Marinova, E. van der Voet, D. P. van Vuuren, O. Edelenbosch, and R. Heijungs, "Modelling global material stocks and flows for residential and service sector buildings towards 2050," J. Clean. Prod., vol. 245, p. 118658, 2020, doi: 10.1016/j.jclepro.2019.118658.
[101] É. Gobbo, E. M. Nia, A. Straub, and A. Stephan, "Exploring the effective reuse rate of materials and elements in the construction sector," J. Build. Eng., vol. 98, p. 111344, Dec. 2024, doi: 10.1016/j.jobe.2024.111344.
[102] A. Miatto, H. Schandl, and H. Tanikawa, "How important are realistic building lifespan assumptions for material stock and demolition waste accounts?," Resour. Conserv. Recycl., vol. 122, pp. 143–154, Jul. 2017, doi: 10.1016/j.resconrec.2017.01.015.
[103] I. Sartori, N. H. Sandberg, and H. Brattebø, "Dynamic building stock modelling: General algorithm and exemplification for Norway," Energy Build., vol. 132, pp. 13–25, 2016, doi: 10.1016/j.enbuild.2016.05.098.
[104] D. B. Crawley et al., "EnergyPlus: creating a new-generation building energy simulation program," Energy Build., vol. 33, no. 4, pp. 319–331, Apr. 2001, doi: 10.1016/S0378-7788(00)00114-6.
[105] A. Logg, V. Naserentin, and D. Wästberg, "DTCC Builder: A mesh generator for automatic,efficient, and robust mesh generation for large-scale city modeling andsimulation," J. Open Source Softw., vol. 8, no. 86, p. 4928, Jun. 2023, doi: 10.21105/joss.04928.
[106] F. Cherubini and A. H. Strømman, "Life cycle assessment of bioenergy systems: State of the art and future challenges," Bioresour. Technol., vol. 102, no. 2, pp. 437–451, Jan. 2011, doi: 10.1016/j.biortech.2010.08.010.
[107] G. Guest, F. Cherubini, and A. H. Strømman, "Global Warming Potential of Carbon Dioxide Emissions from Biomass Stored in the Anthroposphere and Used for Bioenergy at End of Life," J. Ind. Ecol., vol. 17, no. 1, pp. 20–30, Feb. 2013, doi: 10.1111/j.1530-9290.2012.00507.x.
[108] Y. D. Priore, S. Delmenico, L. Rinquet, G. Habert, and T. Jusselme, "Potential contribution of biogenic materials in new and renovated buildings towards carbon budgets and storage in Switzerland," J. Clean. Prod., vol. 541, p. 147558, Jan. 2026, doi: 10.1016/j.jclepro.2026.147558.
[109] Boverket, "Climate database from Boverket (version 02.07.000).," 2026. Accessed: Jul. 05, 2026. [Online]. Available: https://klimatdatabasen.boverket.se/
[110] Z. Barjot and T. Malmqvist, "Limit values in LCA-based regulations for buildings – System boundaries and implications on practice," Build. Environ., vol. 259, p. 111658, Jul. 2024, doi: 10.1016/j.buildenv.2024.111658.
[111] T. Malmqvist, S. Borgström, J. Brismark, and M. Erlandsson, "Reference values for embodied carbon of Swedish building construction," 2023.
[112] K. Dabrock, J. Ulken, N. Pflugradt, J. M. Weinand, and D. Stolten, "Generating a nationwide residential building types dataset using machine learning," Build. Environ., vol. 274, no. October 2024, p. 112782, 2025, doi: 10.1016/j.buildenv.2025.112782.
[113] Statistikmyndigheten, "Statistics Sweden," Statistiska Centralbyrån. [Online]. Available: https://www.scb.se/en/
[114] C. Liljenström, S. Toller, J. Åkerman, and A. Björklund, "Annual Climate Impact and Primary Energy Use of Swedish Transport Infrastructure (Erratum Published in EJTIR 20(2), Pp. 36-40)," p. 40, 2019.
[115] F. Donati et al., "The future of artificial intelligence in the context of industrial ecology," J. Ind. Ecol., p. jiec.13313, 2022, doi: 10.1111/jiec.13313.
[116] J. Zhu, "Suggested use? On evidence-based decision-making in industrial ecology and beyond," J. Ind. Ecol., vol. 24, no. 5, pp. 943–950, Oct. 2020, doi: 10.1111/jiec.13009.

[117] T. Lundblad, M. Taljegard, N. Mattsson, E. Hartvigsson, and F. Johnsson, “An open data-based model for generating a synthetic low-voltage grid to estimate hosting capacity,” Sustain. Energy Grids Netw., vol. 39, p. 101483, Sep. 2024, doi: 10.1016/j.segan.2024.101483.
[118] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” J. Big Data, vol. 3, no. 1, p. 9, Dec. 2016, doi: 10.1186/s40537-016-0043-6.
[119] Å. Löfgren and J. Rootzén, “Brick by brick: Governing industry decarbonization in the face of uncertainty and risk,” Environ. Innov. Soc. Transit., vol. 40, pp. 189–202, 2021, doi: 10.1016/j.eist.2021.07.002.
[120] D. A. Ness, “Technological efficiency limitations to climate mitigation: why sufficiency is necessary,” Build. Cities, vol. 4, no. 1, pp. 139–157, 2023, doi: 10.5334/bc.297.
[121] A. Arceo et al., “Ten questions concerning housing sufficiency,” Build. Environ., vol. 277, no. March, p. 112941, 2025, doi: 10.1016/j.buildenv.2025.112941.
[122] S. Pauliuk, “Decent living standards, prosperity, and excessive consumption in the Lorenz curve,” Ecol. Econ., vol. 220, no. January, p. 108161, 2024, doi: 10.1016/j.ecolecon.2024.108161.
[123] B. Ebrahimi, L. Rosado, and H. Wallbaum, “Machine learning-based stocks and flows modeling of road infrastructure,” J. Ind. Ecol., vol. 26, no. 1, pp. 44–57, 2022.
[124] I. Sartori, N. H. Sandberg, and H. Brattebø, “Dynamic building stock modelling: General algorithm and exemplification for Norway,” Energy Build., vol. 132, pp. 13–25, 2016, doi: 10.1016/j.enbuild.2016.05.098.
[125] A. Miatto et al., “A spatial analysis of material stock accumulation and demolition waste potential of buildings: A case study of Padua,” Resour. Conserv. Recycl., vol. 142, pp. 245–256, 2019, doi: 10.1016/j.resconrec.2018.12.011.