UNCOVERING THE CIRCULAR POTENTIAL: ESTIMATING MATERIAL FLOWS FOR BUILDING SYSTEMS COMPONENTS REUSE IN THE SWISS BUILT ENVIRONMENT

SHUYAN XIONG[1], EDWIN ZEA ESCAMILLA[2], GUILLAUME HABERT[3]
[1,2,3]*Chair of Sustainable Construction, ETH Zurich.*
[1]*xiong@ibi.baug.ethz.ch, 0000-0002-4563-730X*
[2]*zea@ibi.baug.ethz.ch, 0000-0002-8677-161X*
[3]*habert@ibi.baug.ethz.ch, 0000-0003-3533-7896*

Abstract. The construction industry plays a critical role in global resource consumption and greenhouse gas emissions, highlighting the urgent need for sustainable development practices. However, a key challenge in this area is the lack of effective models for resource use that align with circular economy principles. This gap hinders efforts to achieve sustainable resource management, especially in the face of increasing urbanization and material demand. To address this issue, our study presents a Parametric Predictive Model (PPM) to improve resource efficiency, specifically targeting the often-underestimated building systems. The model takes a bottom-up approach, utilizing local databases to accurately assess material stocks of building systems, thereby improving the granularity of data on material composition. Using advanced machine learning algorithms, the model processes both categorical and non-categorical data. The output, an enriched comprehensive database can support more informed decision making in sustainable resource recovery and allocation, but also contribute to the broader goals of reducing waste and promoting resource efficiency in the built environment.

Keywords. Building Systems, Building Stock Modelling, Predictive Model, Circular Economy, Parametric Model.

1. Introduction

The construction industry significantly impacts global resource consumption, energy use, and greenhouse gas (GHG) emissions and faces challenges due to increasing material demand and the escalating urbanization dilemma. The United Nations and the Intergovernmental Panel on Climate Change (IPCC) (Hamilton et al., 2020; Rogelj et al., 2018) emphasize the urgency of sustainable urban development and transitions in construction practices to mitigate climate change effects. The circular economy (CE) is defined as a regenerative approach aiming to maximize resource utility and value (ARUP, 2016).

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 545-554.

Building Stock Modelling (BSM) serves as the core to fulfill the CE concept and allows for material management (Pasichnyi et al., 2019). It predicts future availability by measuring the materials in a system and facilitates effective management. The top-down approach using generalized archetypes, is fast and scalable but lacks traceability of specific building components and precision and may not suit all cultural and economic contexts. In contrast, the bottom-up 'building-by-building' method provides a detailed analysis of each building's material flows, offering greater precision and potential for future monitoring, though it is more costly and complex (D'Alonzo et al., 2020). Moreover, current study focuses mainly on heavy materials like concrete, often overlooking building systems which are lightweight but play an important role in total building emissions (Hoxha et al., 2021). While repurposing them might not always be economical, the easy disassembly nature of metallic parts offers practical, environmentally beneficial solutions.

The implementation of CE strategies in the building sector is hampered by the lack of comprehensive material inventory data and the lack of modern digital models for buildings facing demolition (Leao et al., 2001). The application of machine learning (ML) for assessing old buildings and facilitating material reuse is impeded by data scarcity. For instance, in Switzerland, although numerous public databases exist, obtaining detailed building data for material estimation and identifying reusable materials remains a challenge. Manual data collection is time-consuming, inefficient, and often impractical(Verellen & Allacker, 2020).

Sustainable building solutions involve integrating ML-based predictors with parametric modelling for decision making (Murphy, 2012). However, data from old buildings should be carefully considered to avoid losing valuable insights. Expanding databases with highly accurate ML predictions will improve material estimation and building stock modelling. Parametric modelling, which relies on parameters and equations (Davis, 2013), is meanwhile built by incorporating expert knowledge of building system, and can provide design flexibility, simplification of the design process, and accuracy by complying with the design rules.

This research was supported and funded by the Future Cities Laboratory (FCL) and is part of a work package focusing specifically on Switzerland case study. In this paper this was addressed by combining comprehensive datasets with advanced modelling, particularly in complementing BSM at the component level. This enabled accurate material estimates and efficient planning for material recovery, and optimizing resource use (Röck et al., 2018). The proposed methodology reduces data dependency, allowing it to be widely used with minimal dataset requirements and to predict larger quantities of materials at lower computational cost. The adaptability of the parametric model makes it a versatile open-source platform capable of integrating a wide range of building components and evolving to meet future needs. The results of the study will be further validated with a smaller sample dataset and a reuse potential framework will be developed to provide insight into material reusability, marketability, and comprehensive reuse guidelines.

2. Methodology

Parametric Predictive Modelling (PPM) development in the Python environment consists of two main components. First, the predictive model generation begin with compiling the existing federal building database (Die Schweiz in 3D, n.d.; Home / GEAK, n.d.; Swiss Geoportal, n.d.; Office, 2023), and the missing data is divided into non-categorical and categorical subsets. For non-categorical data such as building dimensions and year of renovation, a linear regression algorithm from the Scikit Learn (Pedregosa et al., 2011) toolkit is used. For categorization of building types and energy systems, a neural network implemented in PyTorch is used (Paszke et al., 2019). Secondly, the parametric model integrates key parameters from the database and generates material quantities by following building system design principles and expert knowledge. Figure 1 illustrates the general workflow and the relevant information exchanged between each step.

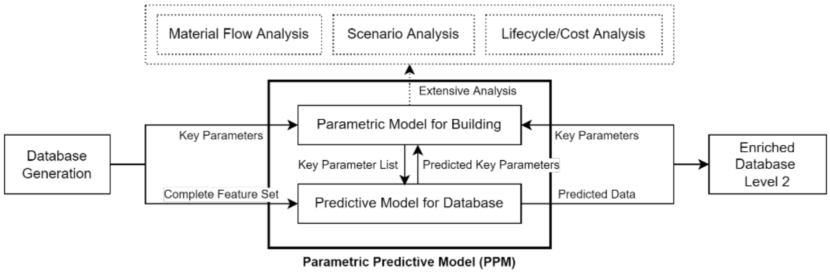

Figure 1. Workflow Diagram of Parametric Predictive Model (PPM) Development

The main categories considered were HVAC, electrical and plumbing systems. PPM initially features selected equipment such as radiators, boilers, ventilation ducts, plumbing, and electrical cables, which are chosen based on reuse experience and metal content, with high reuse potential that could provide significant carbon savings opportunities through reuse.

2.1. PARAMETRIC MODEL DEVELOPMENT

The parametric model development is demonstrated in this subsection, taking radiators as an example. The model assumes its quantity in a heated room is based on the heating system type and presence. The total weight of radiators in a building is calculated by:

$$W_{total\ radiator} = N_{radiator} \times W_{radiator}$$

The number of radiators is estimated using the heated area, number of rooms, or building's heating energy consumption, with a factor indicating one radiator per room (Bornhoft, 2024). The final count of radiators compares the number calculated by the number of rooms with that calculated by the size of the building or heating demand to ensure accuracy.

$$N_{radiator,unit} = Max(N_{radiator,area}, N_{radiator,\#room})$$

$$N_{radiator,area} = A_{unit}/(RC \times HLC)$$

$$N_{radiator,\#room} = \alpha N_{room} + N_{kitchen}$$

$$N_{radiator,building} = Q_{heating,building}/(T_{heating} \times RC)$$

$$N_{radiator} = Max(\sum N_{radiator,unit}, N_{radiator,building})$$

Where:
N_{room}: total number of rooms excluding kitchen
$W_{radiator}$: radiator unit weight
$N_{kitchen}$: total number of kitchens
A_{unit}: area of the living space of apartment
RC: radiator capacity in kW
α: number of radiators per room
HLC: heat loss coefficient in W/m^2
$Q_{heating,building}$: heating demand of the building in kWh
$T_{heating}$: heating period in hours

This approach balances detail with practicality, aiming to closely reflect real-world scenarios. To improve the accuracy of the model, connection matrices for different radiator models and typical installation years can be incorporated into the model, thereby improving the ability to replicate the real world. The approach is also applicable to other building components, thereby improving robustness and generalizability.

2.2. PREDICTIVE MODEL DEVELOPMENT

For the development of the predictive model, it is necessary to create a comprehensive database that integrates various sources with key parameters to ensure flexibility and reliability for practical applications. Key databases used include: *Gebäude- und Wohnungsregister (GWR), Gebäude- und Wohnungsstatistik (GWS), Der Gebäudeenergieausweis der Kantone (GEAK) and 3D Data of Swiss Topography (Swisstopo)* (Die Schweiz in 3D, n.d.; Home / GEAK, n.d.; Swiss Geoportal, n.d.; Office, 2023). The databases cover a wide range of buildings, with available data spanning from about 0.13 to 5 million buildings, encompassing diverse aspects such as building geometry, energy sources, and registered information.

2.2.1. Preparation of Data

To manage the extensive data in this study, various sources are consolidated into a unified database by merging the databases on unique building identifier numbers (EGID), as demonstrated in Figure 2. From the parametric model development section above, essential parameters are identified to estimate equipment quantities and materials. The data is refined and standardized for predictive model development, and harmonized by excluding non-essential details, like the recording year, to focus on relevant features. Included parameters cover aspects such as building geometries, number of floors, residents, apartments, and details about construction, demolition, renovation, building categories, and heating systems.

The data is converted back to the original format from a specialized coding system. In addition, the complexity of the building systems and building categories makes it necessary to summarize the various types of heating systems and buildings by "map" function. The simplified categories allow the study to focus on residential and office buildings, consistent with the primary interest of the study.

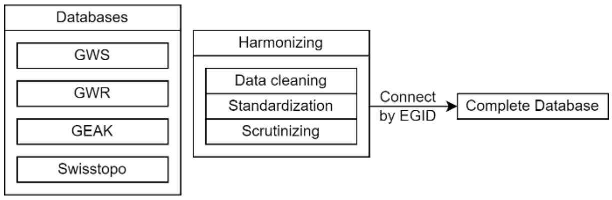

Figure 2. Data Preparation Workflow Diagram

2.2.2. Filling Data Gap

The main challenge in analysing building data lies in the large discrepancies and gaps in existing databases. Switzerland lacks comprehensive data on all buildings, with the GEAK database covering less than 5% of buildings. Even in the more inclusive GWR and GWS databases, data inconsistencies and gaps are evident, which hinders the construction of PPM and the estimation of equipment and material quantities.

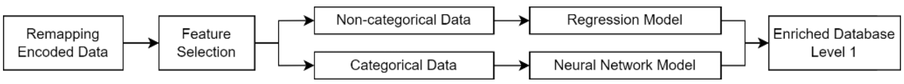

Figure 3. Workflow Diagram for Predictive Model Development and Database Enrichment

To address these gaps, especially for predicting material quantities necessary for further analysis, ML-based predictors are utilized instead of archetype-based clustering. This ML approach allows for the creation of detailed BSM at the building level and is adaptable for exploring specific categories or entire building ranges, thus bridging data gaps, and enhancing predictions for sustainable practices. As shown in Figure 3, building data is categorized into two types for prediction: categorical and non-categorical, each processed using appropriate predictors.

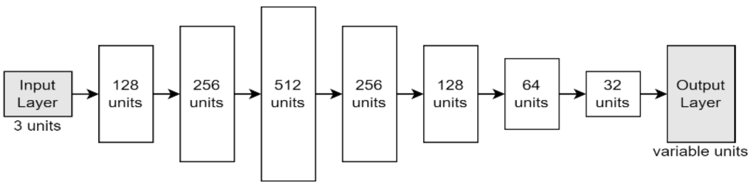

Figure 4. Neural Network Architecture for Predictor of Classification Problem

The data preprocessing step is detailed as follows. Rows with missing values are first removed to maintain the integrity of the dataset. The data is then split into features and target variables. Feature selection is performed using a cross-correlation matrix to identify the top three correlation parameters as key features. These features are then validated by expert knowledge to ensure logical predictions.

Regression methods are suited for continuous or numeric data like unique values per household, including building geometrical and occupancy details. For these non-categorical data, XGBoost regression is employed for data imputation. An XGBoost regression model customized with specific hyperparameters is used to predict and fill in missing values. The validity of the model is measured by the R-squared metric, which assesses how well the model explains the variance of the target variable.

In contrast, neural networks (NN) are highly effective in classification tasks with limited choices, such as categorizing building types and heating systems, due to their proficiency in handling predefined categories. To leverage this strength, a deep feedforward neural network is constructed using the PyTorch framework, specifically tailored for categorized data. The network is composed of several fully connected layers. Each layer incorporates batch normalization and the ReLU activation function for enhanced performance. The architecture's core features 8 hidden layers and dropout regularization is applied to prevent overfitting. Regarding the network's parameters, the batch size is set to 12,800 across 100 epochs. The output feature number of the last layer is set based on the unique category count in the target column. For the optimization process, the Adam algorithm is employed with a learning rate of 1e-3. Additionally, the ReduceLROnPlateau scheduler is utilized for optimizing learning rate adjustments.

3. Results

3.1. PREDICTIVE MODEL AND ENRICHED DATABSE

The enrichment score indicates the percentage of newly filled data fields in the database, which initially had over 40% missing values, especially in non-categorical data, due to sparse Swiss building coverage in the GEAK database from null values and limited size. Post predictive model application, database completeness significantly improved, potentially enriching data up to 95%, especially in previously underrepresented areas.

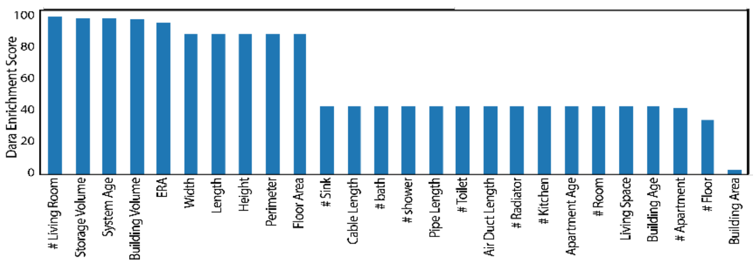

Figure 5. Enrichment Score of the Predictive Model for Non-Categorical Datasets

3.2. PARAMETRIC MODEL

Parametric model is designed to estimate the availability of the building components based on key parameters. Connection matrices are included for addressing equipment availability and variability by year, building categories etc.

3.2.1. Connectivity Matrix

Central to this model is the connectivity matrix, which enriches the database by including assumptions on inter-relationships of energy sources, building categories, and construction periods. For instance, regarding the relationship between building systems and building categories, the prevalence of Swiss residential buildings typically has no HVAC systems. To facilitate understanding, buildings categories are simplified into residential, office, and others, integrating binary availability assumptions from expert knowledge, as sampled in Table 1. Same approach is also applied to analyse relationships of equipment and energy sources, equipment (material) and installation years. This integration enhances the model with detailed layers and predicts the quantities of available equipment and related material amounts, offering flexibility for additional matrices as required.

Table 1. Sample Connectivity Matrix for Building Systems Across Different Categories

Building Systems	Component	Building Category		
		Residential	Office	Other
Heating	Radiator	1	1	0
HVAC	Air Duct	0	1	0
Plumbing	Water Pipe	1	1	1
Electrical	Electrical Cable	1	1	1

3.3. UNCERTAINTY QUANTIFICATION

3.3.1. Parametric Model

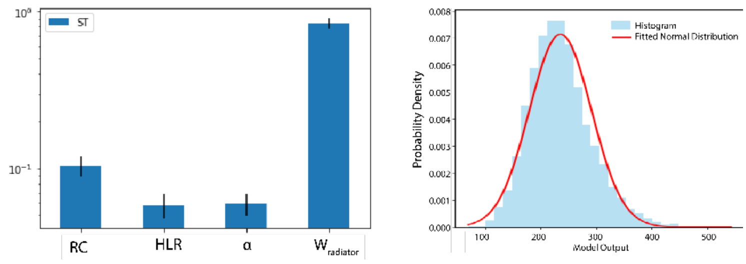

Figure 6. Evaluation of the Parametric Model for Radiator Quantity Estimation. (a) Sensitivity Analysis. (b) Probabilistic Distribution of Model Results.

To calculate radiator quantity with a parametric model, basic data like living area and room count are used, but certain values are based on assumptions, introducing uncertainty. A three-stage sensitivity analysis is conducted to mitigate this. Firstly, assumption-based parameters are refined through iterative analysis, reducing reliance on assumptions. Secondly, parameter ranges reflecting real-life scenarios are chosen, with radiator unit weight being a key factor (Figure 6(a)). Lastly, uncertainties are quantified using Monte Carlo simulation within the SciPy package(Virtanen et al., 2020), resulting in a 20% relative standard deviation. This comprehensive approach enhances the model's reliability by critically evaluating and narrowing down essential parameters and their ranges.

3.3.2. Performance of Predictor

The performance of the predictors, including both regression and neural network models, is evaluated as shown in Figure 7. The regression model's accuracy is determined using the R-squared (R^2) score, which measures the variability of model. The R^2 score for cooking equipment is 0.912 which indicates 9% variance.

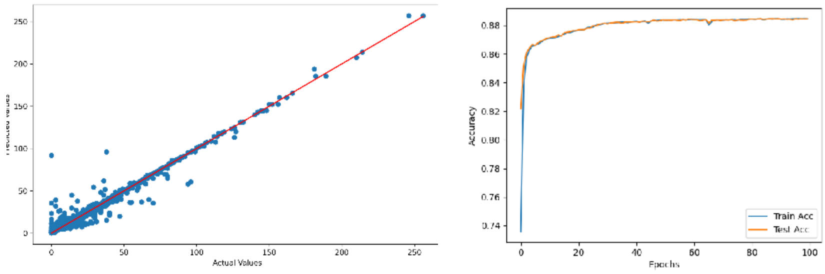

Figure 7. Performance of Predictors. (a) Scatter Plot: Actual vs. Predicted Values for the Regression Model on Cooking Facilities. (b) Evolution of Accuracy for the NN Model on Heat Generator.

For the NN-based classification model, predictor performance is assessed based on the accuracy achieved after all training epochs are completed. The accuracy of the predictor, as shown in Figure 7(b), can reach nearly 90% for determining the heat generator used in each building, crucial for understanding the availability of related heating equipment on a building-by-building basis.

4. Discussion

To accurately estimate resources, the probability of different processing streams and equipment lifetimes must be considered. Model validation involving additional data and expert knowledge is also critical. Through cross-validation and real-world data, model parameters can be adjusted, creating a feedback loop for continuous improvement.

However, the effectiveness of the method depends on the quality of the database used. The comprehensive database includes diverse characteristics of the building stock across different regions and time periods, and how the variation may affect the predictive accuracy of the model further investigated. Switzerland is unique in that it has very detailed public resources, including extensive federal register information and detailed data on building performance and energy systems. While this level of detail may not be replicable in other countries, the robustness of the PPM has been proven. Even with minimal requirements, the model improves the completeness of data characterization with a high degree of accuracy. In addition, PPM is adaptable to a wide range of situations and can be customized based on local expert knowledge and region-specific building system design rules. For example, in Singapore, where HVAC systems are prevalent, PPM can be adapted to focus on the components associated with these systems.

5. Conclusion

Current BSM approaches, primarily archetype-based, often lack detailed information on individual buildings, limiting their utility in assessing reusable materials on a local scale. This gap affects the development of decentralized markets for the circular economy. BSM has traditionally focused on structural materials and building skins, which are challenging to reuse due to labour intensity and limited environmental savings potential. However, building system materials, mainly metals, offer a longer service life and high reuse potential, yet their environmental impact and reuse potential are often overlooked.

This study highlights the importance of data-rich integrated databases in developing parametric models for the building sector. By combining various data sources and applying machine learning tools, a data-driven predictive model is developed, capable of accurately predicting building systems and materials embedded in the current building stock. While our focus is on leveraging these robust models for filling the missing values, other variants may exhibit comparable performance, which opens avenues for future research.

Despite inherent data gaps and uncertainties, this research lays the groundwork for data-driven decision-making in construction. Efforts are made to reduce uncertainty, such as minimizing assumption-based parameters and enhancing accuracy through machine learning predictions. Future potential lies in refining these models with real project data, offering flexibility to extend to various building components and aligning with design principles. In summary, this research contributes significantly to the construction industry's sustainability by enabling informed decisions in reuse, recycling, and renovation, thus promoting resource efficiency and sustainable building practices.

Acknowledgements

This research was conducted at the Future Cities Lab Global at ETH Zurich. Future Cities Lab Global is supported and funded by the National Research Foundation, Prime Minister’s Office, Singapore under its CREATE programme and ETHZ, with additional contributions from the NUS, NTU, SUTD. The processing of geometrical is conducted by ESD (ETHZ). Appreciation is extended to Huiying Zhang and Yunzhi Lin for supervision on ML and hardware support.

References

Bornhoft, A. (2024). The surprising world of swiss apartment life. Retrieved from https://www.wanderwego.com/home/14-things-about-swiss-apartments

D’Alonzo, V., Novelli, A., Vaccaro, R., Vettorato, D., Albatici, R., Diamantini, C., & Zambelli, P. (2020). A bottom-up spatially explicit methodology to estimate the space heating demand of the building stock at regional scale. Energy and Buildings, 206, 109581. https://doi.org/10.1016/j.enbuild.2019.109581

Davis, D. (2013, August 7). A History of Parametric. Daniel Davis. https://www.danieldavis.com/a-history-of-parametric/

Die Schweiz in 3D. (n.d.). Bundesamt für Landestopografie swisstopo. Retrieved Novem-ber 30, 2023, from https://www.swisstopo.admin.ch/de/karten-daten-online/karten-geodaten-online/3d-viewer.html

Hamilton, I., Kennard, H., Rapf, O., Kockat, J., Zuhaib, S., Abergel, T., Oppermann, M., Otto, M., Loran, S., Steurer, N., & others. (2020). Global Status Report for Buildings and Construction: To-wards a Zero-Emission. United Nations Environment Programme, Efficient and Resilient Buildings and Construction Sector: Nairobi, Kenya.

Home / GEAK. (n.d.). Retrieved October 23, 2023, from https://www.geak.ch/

Hoxha, E., Maierhofer, D., Saade, M. R. M., & Passer, A. (2021). Influence of technical and electrical equipment in life cycle assess-ments of buildings: Case of a laboratory and research building. Inter-national Journal of Life Cycle Assessment, 26(5), 852–863. https://doi.org/10.1007/s11367-021-01919-9

Leao, S., Bishop, I., & Evans, D. (2001). Assessing the demand of solid waste disposal in urban region by urban dynamics modelling in a GIS environment. Resources, Conservation and Recycling, 33(4), 289–313. https://doi.org/10.1016/S0921-3449(01)00090-8

Murphy, K. P. (2012). Machine learning: A probabilistic perspective.

Office, F. S. (2023, September 26). Gebäude- und Wohnungsstatistik (GWS), Geodaten 2022 | Publication. Federal Statistical Office. https://www.bfs.admin.ch/asset/en/27905171

Pasichnyi, O., Wallin, J., & Kordas, O. (2019). Data-driven building archetypes for urban building energy modelling. Energy, 181, 360–377. https://doi.org/10.1016/j.energy.2019.04.197

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Advances in Neural Information Processing Systems 32 (pp. 8024–8035). Curran Associates, Inc. http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

Röck, M., Hollberg, A., Habert, G., & Passer, A. (2018). LCA and BIM: Visualization of environmental potentials in building construc-tion at early design stages. Building and Environment, 140(December 2017), 153–161. https://doi.org/10.1016/j.buildenv.2018.05.006

Rogelj, J., Shindell, D., Jiang, K., Fifita, S., Forster, P., Ginzburg, V., Handa, C., Kheshgi, H., Kobayashi, S., Kriegler, E., & others. (2018). Mitigation pathways compatible with 1.5 C in the context of sustain-able development. In Global warming of 1.5 C (pp. 93–174). Inter-governmental Panel on Climate Change.

Swiss Geoportal. (n.d.). Geo.Admin.Ch. Retrieved October 23, 2023, from https://map.geo.admin.ch

Verellen, E., & Allacker, K. (2020). Using data-driven models to estimate the energy use of buildings based on a building stock model. IOP Conference Series: Earth and Environmental Science, 588(3), 032006. https://doi.org/10.1088/1755-1315/588/3/032006

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2