Built Stock Explorer

– an interactive platform for data-driven energy planning

RUSLAN ZHURAVCHAK
Norwegian University of Science and Technology
Department of Energy and Process Engineering
Kolbjørn Hejes v 1B, 7491 Trondheim, Norway
ruslan.zhuravchak@ntnu.no

NATASA NORD
Norwegian University of Science and Technology
Department of Energy and Process Engineering
Kolbjørn Hejes v 1B, 7491 Trondheim, Norway
natasa.nord@ntnu.no

HELGE BRATTEBØ
Norwegian University of Science and Technology
Department of Energy and Process Engineering
Kolbjørn Hejes v 1B, 7491 Trondheim, Norway
helge.brattebo@ntnu.no

Keywords: energy performance, built stock, data visualization, stochastic forecasting, uncertainty analysis

1. Introduction

The achievement of environmental and energy-related targets for nations and communities relies strongly on understating the energy performance of buildings that comprise the existing built stock. Such information is essential for developing more accurate models, aimed at forecasting the expected stock changes [1] and evaluating the feasibility of energy efficiency measures [2]. A variety of modelling and assessment methods [3], based on such information, enable to carry out "what-if" analysis and to define the feasible strategic energy pathways to follow. Such pathways may aim at the increase of energy performance, achieving economic benefits, environmental impact reduction or a combination of these.

The accuracy of forecasting and screening techniques, however, is hindered by uncertainties behind the complex and dynamic structure of built stock. Large-scale energy modelling of any kind must account for these uncertainties, summarized and informed by the appropriate statistical means.

This article aims at familiarizing the reader with Build Stock Explorer (https://buildingstockexplorer.indecol.no/) – a web platform developed for interactive analysis and visualization of key variables associated with Norwegian built stock. The user has a possibility to extract relevant statistical information, which can be used further for stochastic simulation, uncertainty and/or sensitivity analysis. An application is based on the Norwegian Energy Performance Certificates (EPC) dataset, has three visual elements and several features that are important for developing built stock models. These are elaborated further in the article.

2. Data and data management

EPC scheme is considered as a powerful mechanism to engage market forces to support the progress towards high energy performance in communities [4, 5]. Hence, the Norwegian EPC scheme has been in place since 2010 intended to ensure Nation's compliance with the Directive 2002/91/EC, improving the building energy awareness and the promotion of

Table 1. Descriptive statistics on the dataset - central tendency, dispersion and shape of the distribution.

	Construction year	Heated floor area (m^2)	Energy use ($kWh \cdot y^{-1}$)	Energy intensity ($kWh \cdot m^{-2} \cdot y^{-1}$)
Count		72169		
Mean	1981	686.43	144 224.00	156.34
STD	35	2 513.49	608 568.60	79.34
Min	1800	12.00	500.00	0.34
25% percentile	1960	65.00	7 999.00	105.02
50% percentile (median)	1989	91.00	12 596.00	134.56
75% percentile	2012	168.00	23 436.00	188.93
Max	2018	66 255.00	26 118 006.00	1 354.18

low-energy use. During six years of operation, more than 670 000 certificates were issued (20% of all dwellings). Certificates are obtained by the building owners through on-line registration of related data which is followed by automated validity checks of data entry. The scheme applies to both residential and non-residential buildings that are being sold or rented out [6]. Some users voluntarily reported the real total annual energy use for the building/dwelling. Only a subset of the EPC dataset where such information has been specified was used further in the application. A postprocessed dataset contains 72 169 records corresponding to 25 largest Norwegian cities.

Among the key variables that the applicants reported are four continuous, summarized in **Table 1**, and three of categorical type, namely city, building category and building type. Building category, 13 unique values, corresponds to the primary use purpose of the building whereas building type contains architectural or component-related information (25 unique values).

To enable intuitive and flexible operations on the dataset at the interface level, yet preserving computational efficiency and robustness, data management follows the scheme illustrated in **Figure 1**. "City", "Category" and "Type" components are the dropdown lists that support multiple items selection and therefore, enable comparative analysis. "Construction year", "Heated area" and "Total energy use" are the range sliders for slicing the dataset according to the user's interest.

With the structure shown in **Figure 1**, the user has the freedom to make changes at any component and see the updated results immediately, without following a predefined sequence of steps. This is achieved by linking every node with the final subset directly. For more intuitive use, categories of buildings available for selection are only those matching the criteria specified at a higher level in this hierarchy. The same applies to building type, which is the child node of "Category" component.

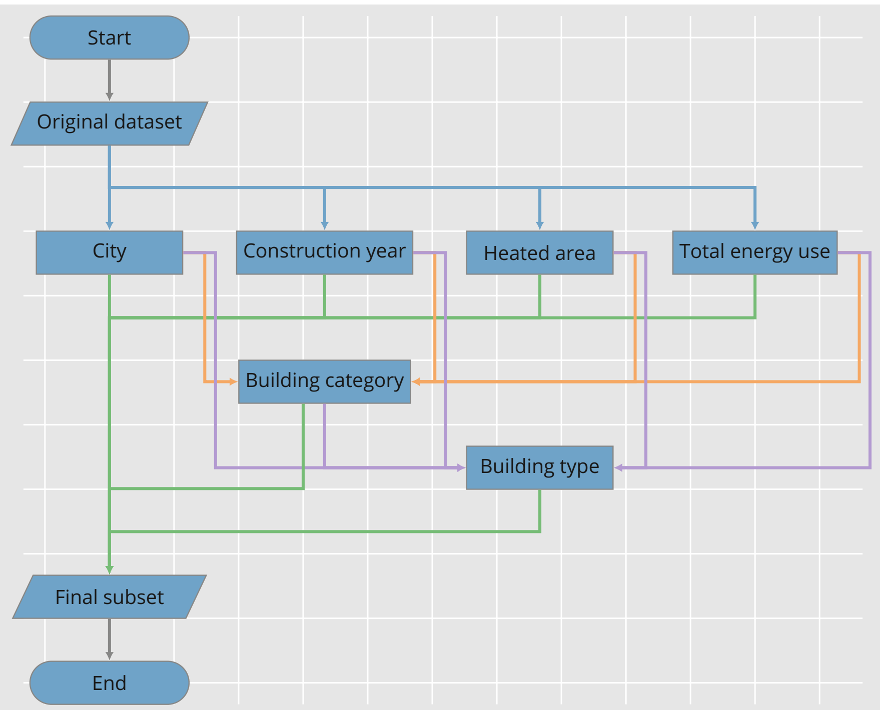

Figure 1. Data management hierarchy.

3. Key features

Given that the user selected and adjusted the dataset to their interest, Built Stock Explorer offers a comprehensive statistical toolset and the visualization of the dataset in three views:

- 3-dimensional scatter plot
- Univariate distribution plot
- Pairwise correlations plot

To display the data of various orders of magnitude, a logarithmic scale can be used optionally.

Three-dimensional scatterplot shown in **Figure 2** illustrates the three essential continuous variables: construction year, heated floor area (m^2), and total energy use ($kWh \cdot y^{-1}$) per building type selected. This view gives a general outline of how many records are selected, what are the magnitudes and trends between these variables.

Univariate distribution plot shown in **Figure 3** communicates the information on built stock variability and computes the essential statistical parameters related to central tendency and variance for the selected subset. The user can also experiment with matching a specific type of theoretical probability density function (PDF). An extra feature enables to find the best distribution function automatically. The algorithm iterates over 25 common distribution types, estimates their parameters, evaluates the goodness-of-fit and eventually selects the best one. The metrics for selecting best fit is the minimal Sum of Squared Residuals (SSR):

$$SSR = \sum_{i=0}^{n} (y_i - f_i)^2$$

where y_i is the value of the data-specific probability density and f_i - the value of theoretical PDF accordingly over 100 intervals.

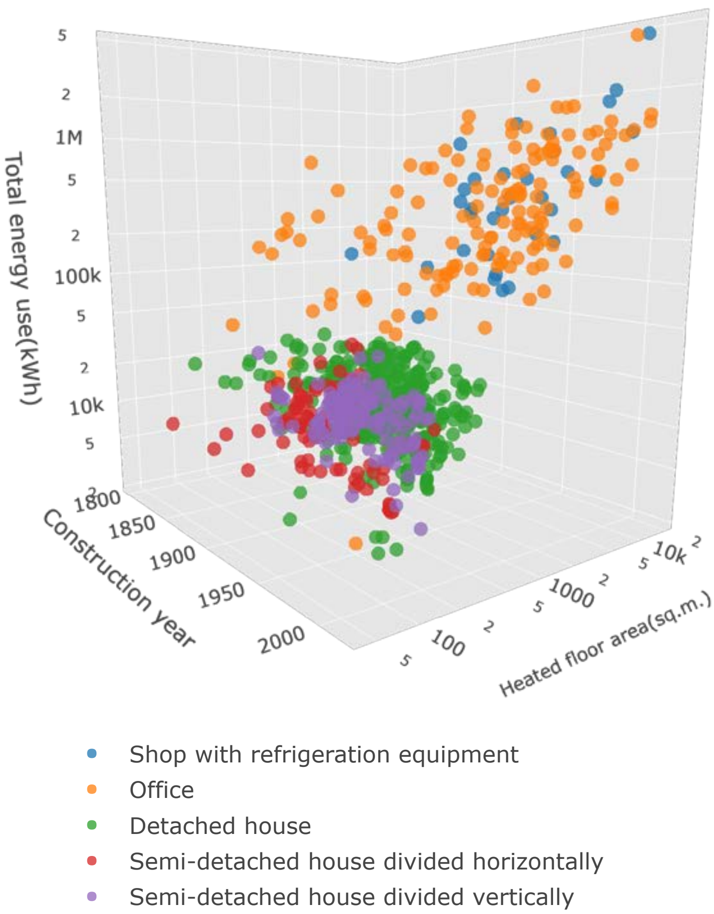

Figure 2. 3-dimensional scatter plot of sample subset.

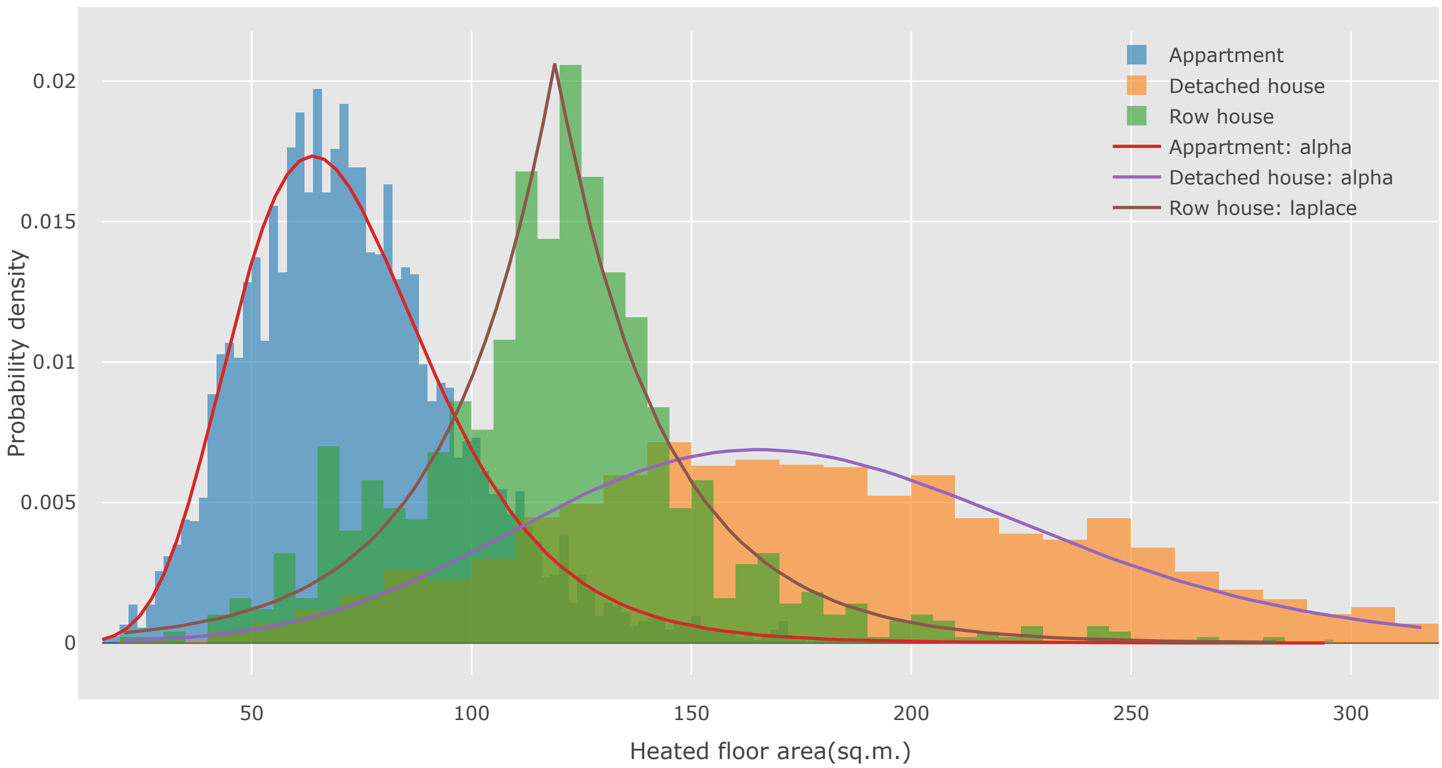

Figure 3. Univariate distribution plot of sample dataset and the corresponding PDF fitted.

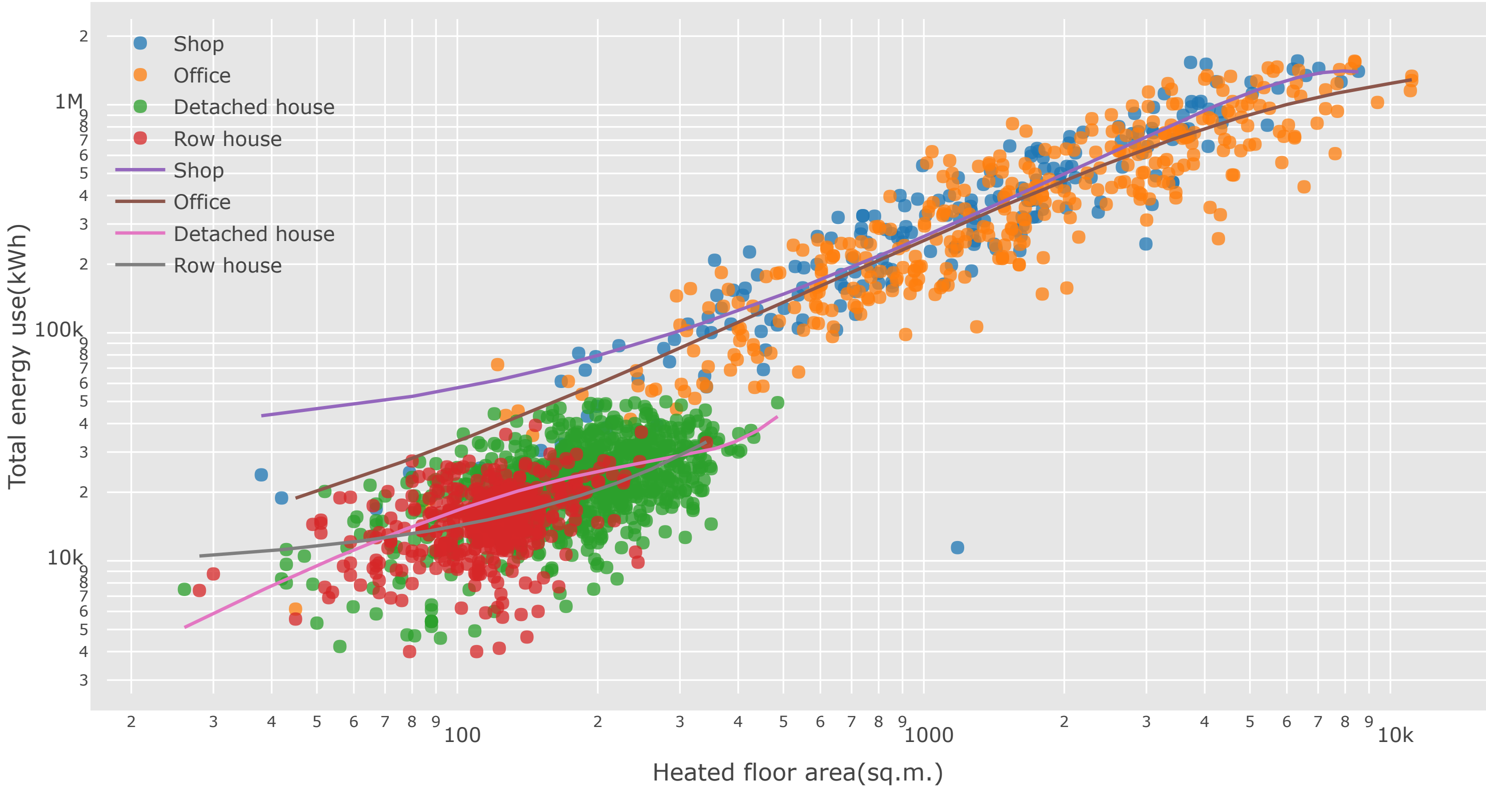

Figure 4. 2-dimensional scatter plot of sample subset and 3-rd order polynomial fit.

Probability density histograms and PDFs, specific to sample buildings types and variable ranges are illustrated in **Figure 3**.

The third view mode enables to explore pairwise correlations between the variables of interest with **two-dimensional scatter plot**, see **Figure 4**. It is possible to find both linear and non-linear (up to 5th order) relationships with interpolating polynomial fit of a form:

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0$$

where is x the variable under consideration, n is the order of the polynomial, and a are polynomial coefficients.

4. Conclusions

Heterogenous structure and dynamic development of built stock make it challenging to establish energy pathways for cities and urban districts to follow. With data-driven methods, however, uncertainties in building energy use and its key driving variables can be incorporated into modelling practices. At its current step development progress, Built Stock Explorer offers a statistical toolset for quantifying these uncertainties and the means for communicating them visually. ■

5. References

1. Sartori, I., N.H. Sandberg, and H. Brattebø, *Dynamic building stock modelling: General algorithm and exemplification for Norway.* Energy and Buildings, 2016. **132**: p. 13-25.
2. Nord, N. and S.F. Sjøthun, *Success factors of energy efficiency measures in buildings in Norway.* Energy and Buildings, 2014. **76**: p. 76-487.
3. Swan, L.G. and V.I. Ugursal, *Modeling of end-use energy consumption in the residential sector: A review of modeling techniques.* Renewable and Sustainable Energy Reviews, 2009. **13**(8): p. 1819-1835.
4. Li, Y., et al., Review of building energy performance certification schemes towards future improvement. Renewable and Sustainable Energy Reviews, 2019. **113**: p. 109244.
5. Prieler, M., M. Leeb, and T. Reiter, *Characteristics of a database for energy performance certificates.* Energy Procedia, 2017. **132**: p. 1000-1005.
6. Brekke, T., O.K. Isachsen, and M. Strand, *EPBD implementation in Norway. Status in December 2016.* 2018, Enova, Norwegian Water Resources and Energy Directorate (NVE), Norwegian Building Authority (DIBK).