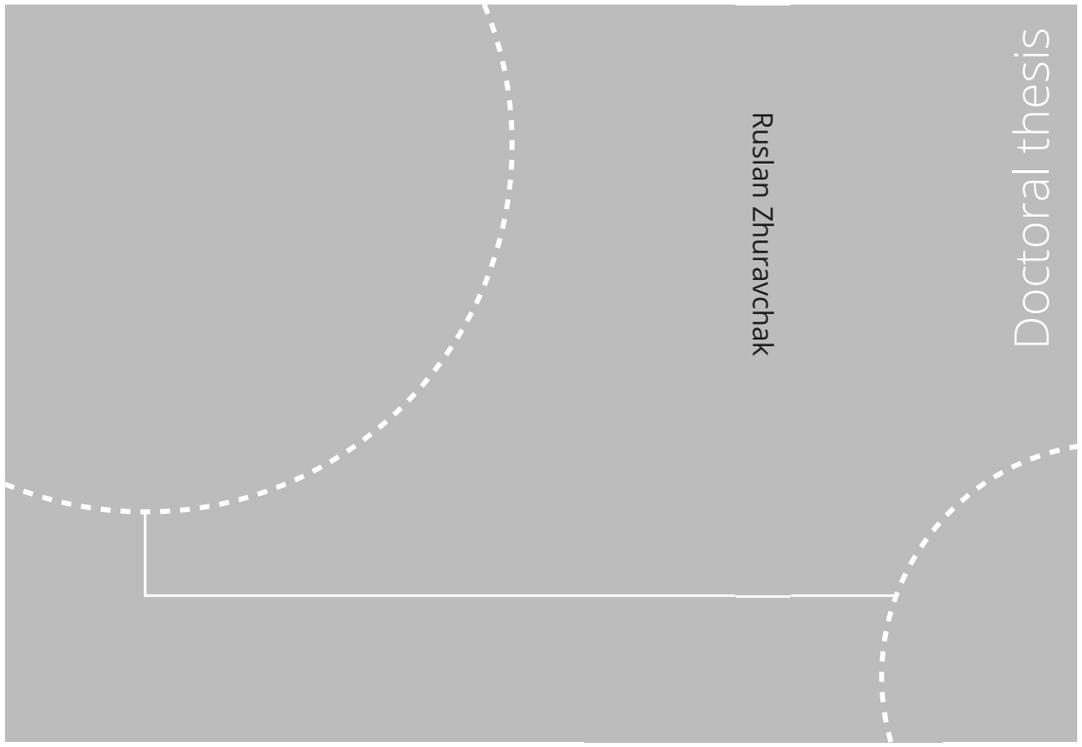
ISBN 978-82-326-5662-2 (printed ver.)
ISBN 978-82-326-5556-4 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (electronic ver.)



Doctoral theses at NTNU, 2022:239

Ruslan Zhuravchak

Variability of building energy performance at a scale:

conformity of predictive and synthesis with explanatory modelling practices

NTNU
Norwegian University of
Science and Technology
Thesis for the degree of
Philosophiae Doctor
Faculty of Engineering
Department of Energy and Process Engineering

NTNU
Norwegian University of
Science and Technology

Doctoral theses at NTNU, 2022:239

NTNU

NTNU
Norwegian University of
Science and Technology

Ruslan Zhuravchak

Variability of building energy performance at a scale:

conformity of predictive and synthesis with explanatory modelling practices

Thesis for the degree of Philosophiae Doctor

Trondheim, December 2022

Norwegian University of Science and Technology
Faculty of Engineering
Department of Energy and Process Engineering

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Engineering
Department of Energy and Process Engineering

ISBN 978-82-326-5662-2 (printed ver.)
ISBN 978-82-326-5556-4 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (electronic ver.)

Doctoral theses at NTNU, 2022:239

PRINTED IN NORWAY
NO - 1598

NORDIC SWAN ECOLABEL
Printed matter
2041 0731

Printed by Skipnes Kommunikasjon AS

Abstract

Throughout its existence, humankind invented countless means and practices to design, construct and equip buildings. The number of ways to use these buildings, with or without strictly following their initially intended purpose, is even more significant. The historical and the anticipated future evolution of buildings at a varying pace further amplifies their diversity. An already complex phenomenon of building energy use is hence further entangled. Substantial variations also stem from you, my reader, whose lifestyle and occupancy patterns often cause a logical nightmare for the energy analysts like myself. These are some of the challenges in large-scale building energy research. This discipline intends to mediate the transition to a more sustainable built environment with the associated energy supply systems. Since the discipline's inception, the inherent modelling practices follow either a bottom-up or top-down approach. These two seemingly incompatible paradigms not only address the subject matter in a radically distinct manner but differ substantially in their accuracy, sensitivity, transferability, versatility, computability and usability. So far, both approaches have been used independently, concerned with leveraging their advantages and, generally, overlooking the limitations of one or another. It was, however, expected that the best interests of practice and policymaking necessitate a synergy or a combination of approaches rather than their application individually.

Seeking ways to complement bottom-up and top-down approaches laid the foundation for the thesis you are holding. The analysis of modelling purposes, targeted system's complexities, model's characteristics and the associated uncertainties were expected to provide meaningful answers. It was also understood that, under the discipline's quest for accurate prediction, explanatory modelling had been largely overlooked. Formulating and testing the causal theories can improve the understanding of building energy performance, the means to mediate it and aid with developing better predictive models. Therefore, examining the interplay of explanatory and predictive modelling, bottom-up and top-down, is another objective of this thesis. It does so through: i) four research papers that attempt to answer why and to which extent the phenomenon varies beyond its best estimates; ii) a case study that exemplifies and examines the conformity of the modelling results obtained with bottom-up and top-down reasoning. These involve a collection of instruments from detailed building energy performance simulation, known as white-box methods, and their cousins of somewhat darker shades. The latter, in

this work, consists of the methods of probabilistic programming that involve statistical hypothesis testing, univariate density estimation and Monte-Carlo simulation. The methods of combinatorial analysis and numerical optimisation are applied when necessary. The modelling principles consider numerous building types, design characteristics, energy supply solutions, occupancy-related tendencies and geopolitical contexts. The findings are based on and supported by experimental data, which is, together with the essential analytical instruments, made available in Built Stock Explorer (`https://buildingstockexplorer.indecol.no/`). To a large extent, this research software enables reproducing/replicating our results, should you be curious about doing that. The Explorer is written in Python, a lingua franca of today's scientific computing, and evolves to facilitate an interactive built stock energy analysis and the relevant (statistical) modelling.

It is shown in this study that numerically similar built stock energy model results are achievable with either bottom-up or top-down model design. Mutual verification of the model performance in such a way can elevate the confidence of the decision making based on them. Furthermore, to prevent misleading urban developments suggested by poorly performing models. Given the importance of mediating building energy efficiency at all levels of governance, mutual verification of built stock energy modes is advocated as the means for more effective and timely achievement of energy and environmental targets. The complexities, diversity, scale and dynamics associated with building energy use at the built stock level motivate model parsimony. Explanatory modelling may inform more rational and better performing predictive model design. Also, explanatory modelling practices are expected to find applications in discovering new causal relationships, empirically validating the existing knowledge and monitoring the evolution of building energy performance at a scale. Better awareness about the phenomenon and better performance of the models at predicting it further elevate the demand for empirical data quantity and quality. This thesis advocates robust study design, data accessibility, and transparency to address the latter.

Therefore, this academic work contributes to the body of knowledge available in large-scale building energy research by focusing on and articulating the need for synthesis between various modelling practices instead of further diverging them. To the best interest of the discipline and the objectives it pursues.

Contents

Figures

Symbols and acronyms

$\Downarrow$ Modelling scope: top-down reasoning. v, vi, 8–11, 14, 15, 20–22, 27, 30, 32, 36, 41, 43, 47–50
$\Uparrow$ Modelling scope: bottom-up reasoning. v, vi, 8, 9, 11, 14, 15, 22, 24, 25, 30, 35, 40, 43, 47–49
$\Leftrightarrow$ Modelling scope: single building. v, vi, 8, 9, 11, 13–15, 24, 33, 34, 47, 48
$\Leftrightarrow$ Modelling scope: built stock level. v, vi, 8–11, 14, 15, 20–22, 28, 29, 38, 41, 43, 46–48, 50

c.r.v. continuous random variable. 15
CDF Cumulative Distribution Function. vii, 15, 16, 18, 20

d.r.v. discrete random variable. 15

ECDF Empirical Cumulative Distribution Function. vii, 39, 40, 45, 50
EPC Energy Performance Certificate. vii, 11, 21–23, 36, 40, 41, 52

IQR Interquartile Range. 17, 21

KS Kolmogorov-Smirnov (test). 11, 20, 21, 27, 38, 39, 45, 46

MC Monte Carlo (simulation). 11, 24, 27, 32, 44
MLE Maximum Likelihood Estimation (method). 19, 20, 27, 34, 41

PDF Probability Density Function. vii, 15–19, 27, 41, 42

r.v. random variable. 15

SARIMA Seasonal Autoregressive Integrated Moving Average (model). 4, 11, 24
STD Standard Deviation. 17, 21, 36, 41

Chapter 1

Introduction

1.1 Buildings and building energy use

Recent estimates suggest that the final energy use of the building sector globally exceeds 128 *EJ* per annum. This number grew by 8.5% within the last decade and is expected to keep growing in the foreseeable future (International Energy Agency 2020b). Such a high final energy use is the driver of direct and indirect environmental impacts that, together with the other life cycle stages of the buildings, account for the estimated 40% of global greenhouse gases emissions (International Energy Agency 2020a). Hence, decarbonisation of the building sector often takes precedence over the other climate change mitigation plans at various levels of governance (Economidou et al. 2020).

An effective course of action towards improving the energy and emissions performance requires a high degree of understanding, cooperation and information exchange across the entire value chain of the buildings, from planning, through construction and operating stages to the demolition. These actions entail the necessity of engaging with the manufacturers and distributors of construction materials, architectural/ engineering/ maintenance firms, energy suppliers, real estate developers, investment and banking institutions, and building owners and occupants. The initiative must be taken at all levels, from national legislative bodies (Brøgger and Wittchen 2018; Economidou et al. 2020) through municipal authorities to the occupants. Facilitating this dialogue requires quantifying the current and the anticipated future state of buildings, their energy use, energy-related implications of short- and long-term strategic developments, and the associated environmental, economic and societal impacts. In the absence of accurate information describing either or both the status quo and the future state of buildings, large-scale building energy modelling is the only way to support these tasks. This subject area contributes primarily to building design and touches upon the other five (technology design; urban climate; systems design; policy assessment; land use and transportation) key areas that form a larger body of knowledge referred to as urban energy system modelling (Keirstead et al. (2012)).

1.2 Large-scale building energy modelling

Although the exact objectives of built stock energy modelling vary, they are in one or another way concerned with energy use within certain spatiotemporal boundaries. The early endeavours to model this phenomenon date back to the emergence of the first building

energy performance standards, the plans for the development of energy infrastructure and the other pragmatic actions. These were addressed using various resources and modelling techniques. The taxonomy of such models, however, was unavailable until the publication by Swan and Ugursal (2009). The authors considered several aspects that affect how the modelling exercise is approached, e.g. purpose, methods and data needs. But a primary distinction is made by the reasoning from known to unknown geospatial scope. Sometimes the information is available for a few buildings, and the modeller assumes that a larger number of buildings have similar properties. Sometimes, however, the available information describes a large scope, and then the modeller attempts to determine the properties of fewer buildings, often by using exogenous information. These two paradigms are referred to as bottom-up and top-down approaches accordingly (Fig. 1.1).

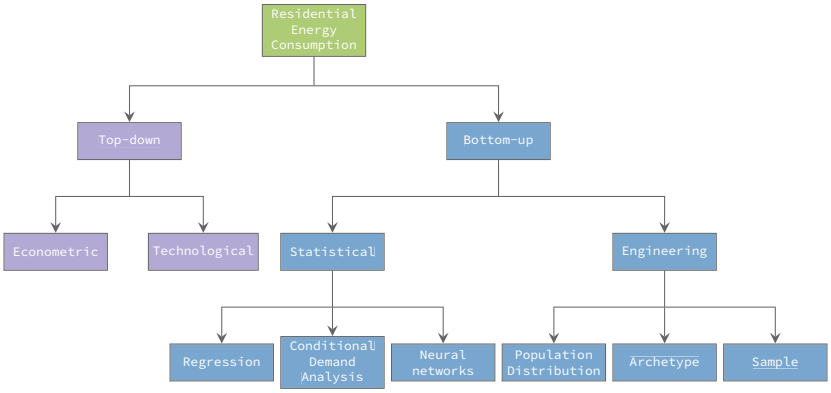


Figure 1.1: A hierarchy of built stock energy models by Swan and Ugursal (2009)

With the built stock model classification of Swan and Ugursal (ibid.) in Fig. 1.1 being the essential reference for over a decade, more recent modelling practices were further diversified in terms of numerous relevant aspects that were not necessarily accommodated by this hierarchical structure. A revised classification system proposed by Langevin et al. (2020) was aimed at resolving this challenge, partially through added flexibility and scalability. The distinction between bottom-up and top-down approaches was supplemented by the underlying technique (black-box or white-box) behind the model and grouping of factors addressed by the model into Energy, People, Built stock and Environment (Fig. 1.2). Additional aspects proposed to determine the model's taxonomic affiliation were system boundaries, spatial resolution, temporal dynamics and the procedures for handling the uncertainties. Such a quadrant-based structure reserved a separate place for distinct methods and approaches synthesised to handle various modelling aspects into "hybrid models".

The domain literature advocates bottom-up engineering- (or physics-)based simulation as capable of accounting for numerous architectural, technical and occupant-related characteristics of the buildings. The question of whether it does so accurately remains unanswered, primarily because of i) the underlying complexities and ii) the importance, and often the dominance, of non-physical aspects in determining the phenomenon. The transfer of heat through walls and fenestration components, for example, is accurately governed by U-values which can be modelled with most of the energy performance sim-

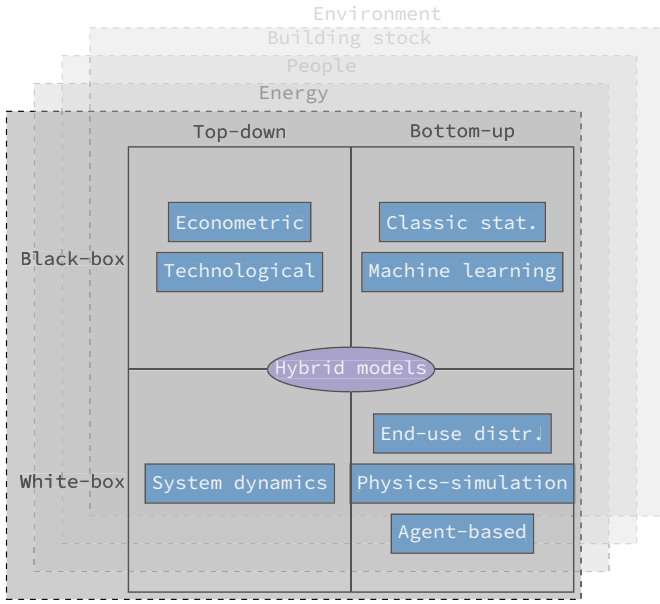


Figure 1.2: Classification of built stock energy models by Langevin et al. (2020)

ulation software and supported by laboratory experiments. However, beyond the laboratory or simulation environments, many aspects related to the construction or the use of buildings significantly affect the actual energy performance, overlooking which leads to the performance gap. Accounting for these details and their influence, individually and jointly, is often irrationally demanding for the intended model's purpose. Alternatively, bottom-up approaches that rely on black-box methods with the empirical data may address some of these complexities by generalising (approximating) the relationships. The use of statistical and machine learning methods (the boundary between the two being somewhat fuzzy) is associated with the risks stemming from an indirect account for the underlying physical mechanisms. Likewise, through empirical data available at a stock level, top-down approaches may (to some extent) explain the phenomena and predict future developments. Their use, however, is associated with a common opinion that the achievable level of technical details and end-uses is not sufficient for some practical applications (Swan and Ugursal 2009; Kavgic et al. 2010; Österbring et al. 2016; Reinhart and Davila 2016; Soto and Jentsch 2016; Moghadam et al. 2017; Brøgger and Wittchen 2018; Langevin et al. 2020). The latter is likely to stem from the absence of studies that suggest the opposite.

In addition to illustrating the diversity of modelling techniques, the available classification systems are meant to support their choice, given the intended application and the available resources. This objective is undermined by several problems, a vague understanding of the predictive capabilities of bottom-up versus top-down modelling being one of them. The others often involve a limited instrumental and methodological basis for accommodating the phenomenon's complexity. Thus, for example, Swan and Ugursal

(2009) distinguish between bottom-up and top-down approaches based on geospatial scope propagated by the model from known to unknown. Langevin et al. (2020) used broader terms "aggregated system" and its "constituent subsystems". It is debatable what the latter convention could refer to i) energy sectors, e.g. district heating system and the power grid; ii) building types, e.g. residential, commercial, educational; iii) urban system, buildings, transport, infrastructural components; iv) end-uses, e.g. heating, lighting, appliances; or any other definition of system's components. Because of the numerous specificities that these subsystems may be subject to, modelling each of these may require a distinct approach or a unique collection of techniques. It is often of practical interest to predict short-term future development of the phenomenon based on the previous observation of the same phenomenon, which is the problem dealt with in time series forecasting. This type of modelling does not imply changes in the spatial levels and is therefore incompatible with these classification systems.

Semantical challenges are likewise common. Hussain et al. (2016), for example, relies on the Seasonal Autoregressive Integrated Moving Average (model) (SARIMA) model and referred by Langevin et al. (2020) as top-down black-box econometric modelling. SARIMA, however, is a time series forecasting model with an origin in applied mathematics and used in numerous domains, micro-and macro-economics being only some of them. This model is more naturally classified as statistical. What is referred to as econometric and technological modelling techniques have their roots in the domain of applied mathematics/statistics. A model classification system proposed by Li et al. (2017), also based on the one of Swan and Ugursal (2009), partially resolves this inconsistency by adding a "statistical analysis" as a subclass in the top-down branch.

The distinction between black- and white-box modelling, although it may be seen as rather radical in other disciplines, has substantial semantical and methodological shortcomings in building energy research. To be classified as a white-box model, each element of the built stock needs to be modelled explicitly and in detail. Despite explicitly modelling some of its components, all the modelling practices that involve approximations, assumptions, and simplifications are not white-box models anymore. This problem is especially relevant to those models involving archetype/typical/representative buildings (Brøgger and Wittchen 2018), expected occupancy schedules, weather profiles and others. A reverse statement also holds when a principally black-box model acquires some components addressed by the established knowledge in physics. Practical gains of combining bottom-up physics-based and bottom-up data-driven models were elaborated by Kavgic et al. (2010), Zhao and Magoulès (2012) and Wei et al. (2018), who defined such modelling practices as "hybrid" (Kavgic et al. 2010) and "grey" (Zhao and Magoulès 2012; Wei et al. 2018) accordingly. Purely white-box modelling of a stochastic system, such as built stock, is not only irrational but practically impossible.

Modelling practices in building energy research are compromised by such dilemmas, and the attempts to classify/diverge them entangle the subject matter even further. Moreover, the excessive focus on one technique or approach and disregarding the other elevates the risks of developing irrelevant theory, obtaining unreliable outputs, and making questionable or misleading conclusions. In applied mathematics, such a situation is referred to as "the error of the third kind"[1], i.e. arriving at an accurate answer to a poorly formulated question. Both the original (Fig. 1.1) and the updated (Fig. 1.2) classifications advance the understanding of various modelling techniques, their capabilities and shortcomings, data sources and computational resources required. However, it becomes

[1]Type I (rejecting the true hypothesis) and Type II (accepting the false hypothesis) of errors in statistical hypothesis testing are discussed further in Chapter 2.

evident that the phenomenon of such complexity, scale and dynamics cannot be fully explained and accurately modelled using a single technique. The tendency for comparison across models further shifts the attention to the mechanics of modelling, away from its two fundamental purposes. Namely, i) practical benefit of the predictive accuracy and ii) knowledge gained about the nature of the phenomena in question.

The purposes and the phenomena that the models are focused on are likewise debatable. A holistic research concept of Energy Epidemiology proposed by Hamilton et al. (2013) emphasises which components of the system should be modelled, with the exact modelling techniques being of secondary importance. The authors propose to consider energy use at the population level as a composite of i) (physical) energy processes and systems, ii) energy practices (interactions between socio-cultural aspects and the physical system) and iii) energy context (external factors affecting the structure of these systems). In health sciences, these components are addressed through bio-medical, socio-behavioural, and environmental/exposure models. The article has suggested seeking their equivalents in building energy research as a promising way ahead. Alternatively, Limpens et al. (2019) suggests distinguishing the models by their pragmatic purposes, which is either simulation or optimisation of the urban energy system. Finally, Ahmad et al. (2018) suggested classifying data-driven built stock energy models purpose-wisely as: benchmarking, energy mapping, energy forecasting and energy profiling.

Thus, for the models to facilitate overcoming the challenges of sustainable future urban systems, the definition of i) modelling purpose and ii) system components and their relationships must take precedence. Then there are several criteria for the models to cope with. In addition to the accuracy, sensitivity and transferability(reproducibility) of the models discussed by Soto and Jentsch (2016), three other essential properties that affect the appropriateness of built stock energy models were identified by Sousa et al. (2017): versatility, computability, and usability. A rational decision on modelling approaches/ techniques/ methods/ resources addresses the tradeoffs between these criteria. However, numerous sources of uncertainty associated with modelling undermine each of them.

1.3 Uncertainties in large-scale building energy modelling

The engineering and physics-based foundations of the discipline traditionally focus on best estimates of the phenomenon. Ideally, the deviation from these estimates are neglegeable either in magnitude and/or in frequency of their occurence. Shifting the attention from individual building to a large, heterogeneous, complex, dynamic, stochastic system such as built stock, elevates the magnitude of potential errors. The extra complexities imply simplifications and assumptions in model design, and often affected by the lack of information about the true behaviour of the phenomena. All of these may lead to the risks of biased model outputs which may ultimately compromise the effectiveness of the decision making based on such models. Dealing with uncertainties is therefore a recurring topic in the discipline.

1.3.1 Model uncertainties

By adapting some modelling practices used in medicine (similarly to Hamilton et al. (2013)), Booth et al. (2012) defined three sources of uncertainties encountered in build-

ing energy research: i) aleatory uncertainty, ii) heterogeneity, and iii) epistemic uncertainty, all of which are discussed further.

Natural stochasticity (or aleatory uncertainty)

The energy performance of the buildings is often affected by factors that lack either order or pattern or coherence or a combination of these and are therefore considered random. As a result, they may cause substantial variabilities in energy performance even across identical buildings. Two large groups of such factors are occupancy-related and local climate-related. The individual occupancy schedules, choice of appliances, cooking habits, indoor environmental quality preferences and building maintenance practices represent the former—the latter stems from factors related to the outdoor temperature, precipitation, irradiance, wind and others. They vary across occupants, locations and buildings, but they also fluctuate on a daily, seasonal and annual basis.

Heterogeneity and non-representativeness

Heterogeneity implies that the buildings with a particular attribute in common are distinctive in terms of the other attributes. The archetype or representative building is defined based on architectural characteristics, age, use purpose and other factors common for the group of buildings. However, the properties of the envelope, energy supply technologies and appliances may vary within the archetype. This uncertainty can be potentially eliminated by a more detailed definition of what attributes define the group. Considering the diversity of attributes and factors, the number of their distinct combinations grows to the extent that makes it unpractical to account for all of them individually. Distinguishing between the attributes that do from those that do not affect the phenomenon at a stock level is therefore essential.

The missing knowledge (or epistemic uncertainty)

Whereas aleatory uncertainty and heterogeneity accommodate the properties of the phenomenon, epistemic uncertainty is concerned with how the model represents this phenomenon. It involves: i) simplifications and assumptions behind the choice of the model structure and model parameters; ii) simplifications and assumptions about the future developments, e.g. scenarios; iii) assumptions about the applicability of the model in a different geospatial scope, temporal horizon and other.

1.3.2 Data uncertainties

In addition to the sources of uncertainty mentioned above, modelling practices involving empirical data need to address several other uncertainties. The studies typically elaborate on those associated with data collection (sampling bias, noise filtering, intentional or unintentional misreporting) and the downstream data wrangling (imputation of missing data, outlier and duplicated records detection/removal).

1.3.3 Quantification of uncertainties

Recent modelling practices tend to account for some of them either as a part of a formal uncertainty analysis procedure or as an intrinsic component of the model. A comprehensive review of methods and approaches used to address the uncertainties in the built

stock energy research was carried out by Tian et al. (2018) where the authors distinguish between forward and inverse uncertainty propagation methods. The former, commonly represented by sampling-based methods, implies systematically altering model inputs or parameters to obtain a set of likely outputs. The latter is concerned with relating already observed variability in empirical data to the model parameters and inputs. These two approaches are considered mutually exclusive and applicable to distinct modelling approaches. Their synergy toward more accurate modelling is not present in the literature.

1.4 Explanatory modelling

Whereas the general philosophy of science defines two distinct scientific goals (Dubin 1969; Shmueli and Koppius 2009), namely i) prediction and ii) understanding (explaining) the phenomena, the literature on large scale building energy research (and the models accordingly) is strongly dominated by the former. The latter is excluded from the model classification systems of both Swan and Ugursal (2009) and Langevin et al. (2020). However, explanatory modelling is essential for testing the causal theories/hypotheses that have numerous implications for policy and practice.

1.5 Research outline

1.5.1 Research motivation and research questions

There is a noticeably growing interest in models for addressing various open questions about the energy performance of the built stock for research and practical applications. Instead of using a single method/technique, a more sophisticated model design appears by combining these to minimise the uncertainties and maintain a rational balance between accuracy, sensitivity, reproducibility, versatility, computability and usability. Numerous examples are available in the literature where hybrid bottom-up models were used for predictive purposes. Nevertheless, the attempts to harmonise bottom-up and top-down modelling are absent in the domain to the best awareness of the authors. Hence, a primary research objective of this work is to examine the conformity of bottom-up and top-down model results. The conformity may enable mutually verifying and synthesising the models, mitigating the uncertainties and better addressing the model appropriateness criteria. Substantial attention is given to reflect upon the parsimony and interpretability of such models in light of the "Occam's razor"[2] with their implications on these criteria. These findings are expected to contribute to the ongoing scientific debate on modelling techniques most applicable for strategic energy planning for cities and communities.

To achieve the main objective, this research examines the capability of the models to predict under the acute variability of the phenomenon. Quantifying the variability is, as shown later, enabled by explanatory modelling. As a secondary objective of this work, we exemplify such modelling and outline the necessity of its use in the discipline to develop better predictive models and build better theoretical foundations for building energy performance at a scale.

These objectives are addressed by answering a set of research questions:

[2]In predictive modelling and knowledge discovery, "Occam's razor" is the term used to emphasise the preference for a simpler model or theory amongst the competing alternatives, which often undermines the accuracy of the analytical conclusions. See Domingos (1999) for more details.

1. How can energy modelling assist with addressing the sustainability of the built stock?
2. How to quantify the variability of building energy performance and how to accommodate these variabilities in predictive modelling?
3. How, given these variabilities, to make robust conclusions about what affects the phenomenon at a stock level and what is not?
4. What is the role of explanatory modelling in i) better understanding the nature of the phenomena and ii) improving the predictive modelling practices?
5. Do the bottom-up and top-down reasoning lead to conforming model outputs, and why should we care if they do not?
6. What are the demands from experimental data to enable the above?

1.5.2 Research design and publications

The answers to the research questions are provided by examining and synthesising four publications that constitute this thesis. Each of these represents either explanatory or predictive modelling practice, distinct design principles and uncertainty propagation methods, as illustrated in Fig. 1.3.

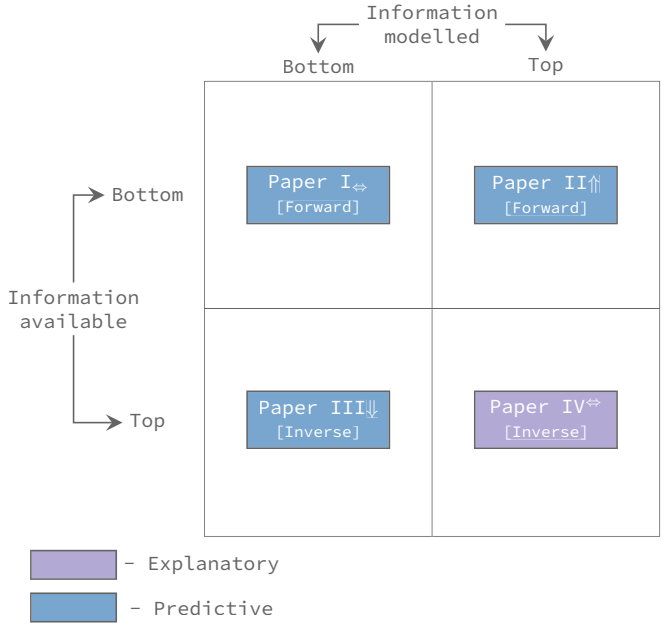


Figure 1.3: Relevance of papers by scope, approach and uncertainty propagation

Fig. 1.3 highlights several design aspects of this research. The spatial levels of the initially available and the modelled information are either bottom (an individual or a few buildings) or top (a larger geopolitical zone, e.g. urban, regional or national scale). A list of four symbols [$_{\Leftrightarrow}$, $\Uparrow$, $\Downarrow$, $^{\Leftrightarrow}$] is used through the thesis to denote the scope of the

modelling approach that the paper represents: i) Single building; ii) Bottom-up model; iii) Top-down model and iv) Built stock level accordingly. Bottom-up and top-down reasoning imply changing the spatial level from bottom to top and from top to bottom accordingly. Papers II⇑ and III⇓ exemplify these techniques. Papers $I_{\Leftrightarrow}$ and $IV^{\Leftrightarrow}$ represent the modelling practices not accommodated by the available classification systems. They are used at distinct spatial levels but do not involve the change.

The modelling objectives are likewise distinct, whereas papers $I_{\Leftrightarrow}$, II⇑ and III⇓ are focused on predicting the phenomenon with uni- (I⇑) and multi-variate techniques (II⇑, III⇓), paper $IV^{\Leftrightarrow}$ represents the modelling of explanatory kind, i.e. establishes the causal relationships between the phenomenon and the explanatory variables.

Two distinct uncertainty propagation methods are involved in these models. Forward propagation is an intrinsic component of the time series forecasting model in paper $I_{\Leftrightarrow}$, where the best estimates of the future (and the estimates of uncertainty) are based on the past observations. In paper II⇑, the uncertain element is the extent to which various technologies will be deployed in the future, likewise propagates forward. Inverse propagation is carried out in papers III⇓ and $IV^{\Leftrightarrow}$, where the dispersion of the univariate distribution, as a measure of uncertainty, is reduced stepwise by adding the explanatory variables.

The papers quantify the variability of energy performance or building attributes either at a building level (Paper $I_{\Leftrightarrow}$) or at a stock-level (papers II⇑, III⇓, $IV^{\Leftrightarrow}$). The rationale for doing this however varies for each of these: to obtain best estimates of the phenomenon in time series forecasting (paper $I_{\Leftrightarrow}$); to determine the most frequent, and therefore representative, parameter of the building (paper II⇑) and to infer the effect of building attribute on the energy performance and to obtain a statistical model for it (papers III⇓ and $IV^{\Leftrightarrow}$).

These distinct design choices and methodological and instrumental characteristics are intended to address several modelling purposes discussed below.

Paper $I_{\Leftrightarrow}$

Ruslan Zhuravchak, Natasa Nord and Helge Brattebø (2019a). 'Control strategy for battery - supported photovoltaic systems aimed at peak load reduction'. In: *E3S Web Conf.* 111, p. 05027. DOI: 10.1051/e3sconf/201911105027

R. Z: Conceptualisation, Methodology, Software, Formal analysis, Writing - Original Draft, Visualisation **N. N:** Validation, Writing - Review & Editing, Supervision **H. B:** Validation, Writing - Review & Editing, Supervision.

The paper objectifies developing the means for more rational scheduling of electric battery operation to minimise peak loads and maximise the photovoltaic energy self-consumption. Such active demand-side management is intended to reduce peak grid load and peak PV feed-in, thus preventing the otherwise necessary grid reinforcement and infrastructure expansion, with the associated costs for energy consumers, distributors and producers. The advocated instrumental basis for achieving this objective are: i) short-term (day-ahead) univariate energy use forecasting model, ii) day-ahead PV output forecasting and iii) load matching. The key focus of the paper is on designing/parameterising the model for predicting and matching the forecasted electric energy use to the forecasted PV generation. This modelling is carried out at a single building level. The paper also elaborates on uncertainty in forecasting as the critical aspect affecting the benefits of the proposed scheduling approach.

Paper II⇑

Ruslan Zhuravchak, Natasa Nord and Helge Brattebø (Oct. 2019b). 'Influence of emerging technologies deployment in residential built stock on electric energy cost and grid load'. In: *IOP Conference Series: Earth and Environmental Science* 352, p. 012038. DOI:

10.1088/1755-1315/352/1/012038

R. Z: Conceptualisation, Methodology, Software, Formal analysis, Writing - Original Draft, Visualisation **N. N:** Conceptualization, Software, Formal analysis, Writing - Original Draft, Supervision **H. B:** Validation, Writing - Review & Editing, Supervision.

The paper examines the likely long-term, large-scale penetration of several technologies related to residential building energy use and the associated changes in energy costs and grid interaction patterns. The proposed methodological and instrumental means are intended to: i) inform strategic energy planning of the power infrastructure with its capital budgeting and ii) mediate the acceptance of such emerging technologies in the best interest of a sustainable built environment. This modelling purpose requires: i) detailed (hourly or sub-hourly) power demand and grid feed-in profiles associated with each combination of technologies under consideration; iii) energy cost profile with the same frequency and iii) the means to accommodate the uncertain future developments concerning the acceptance of new technology in the population. The modelling principles proposed by the paper synthesise these components and address the problem through bottom-up reasoning where white-box energy performance simulation results obtained for a single building are further upscaled to the stock level.

Paper III⇓

Ruslan Zhuravchak, Raquel Alonso Pedrero, Pedro Crespo del Granado, Natasa Nord and Helge Brattebø (2021). 'Top-down spatially-explicit probabilistic estimation of building energy performance at a scale'. In: *Energy and Buildings* 238, p. 110786. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2021.110786

R. Z: Conceptualisation, Methodology, Software, Formal analysis, Writing - Original Draft, Visualisation **R. A. P:** Conceptualisation, Software, Formal analysis, Writing - Original Draft, Visualisation **P. C. del G:** Validation, Writing - Review & Editing **N. N:** Validation, Writing - Review & Editing, Supervision **H. B:** Validation, Writing - Review & Editing, Supervision.

This study emphasises the need for spatial mapping of the energy use hotspots within the built environment for more rational planning of the energy system and urban developments. Such modelling must address the extreme heterogeneity of the built environment and the modelled phenomenon's complexity. Under such circumstances, with the need for an accurate yet parsimonious model, the paper advocates top-down reasoning based on experimental data available at a higher spacial level. Although the modelling objective is to predict the phenomenon, some elements of explanatory modelling are utilised to inform the predictive model's design and elevate its predictive performance. Natural handling of uncertainties is considered one of the main strengths of such modelling.

Paper IV⇔

Ruslan Zhuravchak, Natasa Nord and Helge Brattebø (2022). 'The effect of building attributes on the energy performance at a scale: an inferential analysis'. In: *Building Research & Information* 0.0, pp. 1–19. DOI: 10.1080/09613218.2022.2038537. eprint: https://doi.org/10.1080/09613218.2022.2038537

R. Z: Conceptualization, Methodology, Software, Formal analysis, Writing - Original Draft, Visualization **N. N:** Conceptualization, Validation, Writing - Review & Editing, Supervision **H. B:** Conceptualization, Validation, Writing - Review & Editing, Supervision.

This study elaborates on the need for better understanding of building energy use at the stock level and its causal relationships, direct and indirect, with the building attributes. The obvious benefit of such information is the possibility of mediating the phenomenon through socio-economic and regulatory mechanisms, namely by screening the options

and identifying the opportunities within the strategic urban development programs. The paper also advocates the use of explanatory modelling results to arrive at a better predictive model, hence partially based on methods and instruments discussed in paper III⇓. The discussion on the latter is more detailed compared to paper III⇓. Such explanatory modelling is carried out at the stock level and does not involve the change of the spatial scope. However, a statistical model that approximates the variability and mimics the data generation process may have, amongst the others, a predictive purpose.

Author's contribution

Paper I$_{\Leftrightarrow}$

Lead authorship; case study design; descriptive analysis of time series dataset; designing and programming a sequence of: i) fitting the SARIMA model, ii) k—fold Cross-Validation, iii) Grid Search over the parameters space, iv) quantification of the goodness-of-fit, v) model selection, vi) sending the data requests and processing the response from Solcast application programming interface (API); computational implementation of the Dichotomous (Binary) Search algorithm for load matching; carrying out the simulation; analysis of simulation results; presenting the results at the 13th REHVA World Congress CLIMA 2019, Bucharest, 2019-05-26 - 2019-05-29.

Paper II⇑

Lead authorship; case study design; descriptive analysis of the Energy Performance Certificate (EPC) dataset; design and programming the sequence of: i) a nested Monte Carlo (simulation) (MC) procedure, ii) load data aggregation and iii) cost data aggregation; carrying out the simulation; analysis of simulation results; presenting the results at the 1st Nordic Conference on Zero Emission and Plus Energy Buildings 2019, Trondheim, 2019-11-06 - 2019-11-07.

Paper III⇓

Lead authorship; case study design; descriptive analysis of the EPC dataset and the Cadastral System's dataset; design and programming of the estimation and simulation components of the study; carrying out the simulation; analysis of simulation results.

Paper IV$^{\Leftrightarrow}$

Lead authorship; case study design; descriptive analysis of the EPC dataset; computational implementation of the Cartesian product between building attributes; programming the procedures for i) pairwise Kolmogorov-Smirnov (test) (KS) between samples per combination of attributes, ii) density estimation and goodness-of-fit metrics extraction per sample; carrying out the simulation; analysis of simulation results.

Chapter 2

Methods and materials

This chapter describes the methodological basis of this doctoral work. Section 2.1 elaborates on the overall course of actions taken with their relevance to the research questions and the factors that substantially affected the academic output. The conceptual basis of this work, to a large extent, utilises the applicable methods of probability theory discussed in Section 2.2 and the empirical dataset described in Section 2.3. These methods and data are the components of the research methodology followed in the four articles (Section 2.4) and the case study (Section 2.5). The latter is intended to be an illustrative example that facilitates answering the relevant research questions. Within the research accessibility, reproducibility and replicability enhancement initiative taken by the author and the collaborative partners, some of the data (Section 2.3) and the computational implementation of some of the methods (Section 2.2) are made available in Built Stock Explorer: `https://buildingstockexplorer.indecol.no`.

2.1 Research process

This doctoral work and the associated research process spanned from late 2017 to early 2021. The initial steps built the awareness of the need for modelling and understanding the kinds of models necessary for various applications related to current practices and policymaking. Modelling customs in the discipline appeared as case-specific, following diverse design choices and numerous unique combinations of methods and instruments, in the best interest of their targeted purpose. Although this tendency may be seen as the consequence of a broad range of modelling objectives they pursue, the underlying phenomenon is the same and therefore, numerical outputs are expected to: i) adequately represent the reality, despite the variability, and ii) conform across the diverse modelling approaches. Arriving at such outputs from modelling exercises with different reasoning, design and instrumental basis was set as an objective of this research. Further analysis of these results against the model appropriateness criteria was expected to provide valuable insights into the research questions set by this thesis.

The first modelling exercise was focused on short-term energy forecasting of the simplest component of the built stock - an individual building and resulted in paper $I_{\Leftrightarrow}$. This study examined the temporal variations of energy use, proposed a predictive model that reflected upon these variations and assessed the potential for peak load reduction by utilising photovoltaic energy in an alternative and more rational manner. However, projecting the analytical conclusions obtained for an individual building to the stock level,

no matter the objective, needs to consider the heterogeneity of the buildings.

Hence, the energy-related technological variability of a building with its implications on the stock-level modelling results was examined in paper II⇑. This study attempted to predict some of the implications of the expected future stock-wide adoption of technologies considered novel today. It is concluded that the idealisation of all variabilities except those modelled explicitly is a substantial drawback of the bottom-up reasoning with white-box modelling elements.

To address such challenges, further work consisted in examining the alternative modelling techniques, those having the reasoning from experimental data that reflects all sources of variability in the phenomenon targeted, i.e. top-down approaches. Despite substantial advantages, these techniques were underrepresented in the discipline. This paradigm of top-down modelling is elaborated and advocated in paper III⇓ as capable of parsimonious and uncertainty-aware modelling. Within and beyond the illustrated objective of building energy mapping. It is, however, emphasised in paper III⇓ that leveraging the advantage of this approach necessitates a rational choice of the explanatory variables. These variables, often reflecting specific architectural, technical, geopolitical and occupancy-related considerations, must have empirical evidence of a causal relationship with the phenomenon at a stock level.

Finding and validating the causal effects is the subject of explanatory modelling, which was likewise used seldom in the discipline. Methodological and instrumental means for doing that are discussed in paper IV⇔, together with the illustrative examples provided where applicable.

Hence, papers I⇆, II⇑, III⇓ and IV⇔ contribute to answering the research question 1. Question 2 is concerned with the instrumental means discussed in papers I⇆, II⇑ and III⇓. Quesions 3 and 4 are partially answered in papers III⇓ and IV⇔. The case study (Section 2.5) adopts some of the instrumental basis of papers II⇑ and III⇓ and elaborates on the questions 3 and 4. Question 5 is fully addressed through a case study. The answer to question 6 reflects on the quality of experimental data used in papers I⇆, II⇑, III⇓, IV⇔ and in the case study.

Providing access to the growing amount of data and making relevant methods available within and beyond this research was a primary motive for developing Built Stock Explorer. This software is expected to facilitate a better understanding of energy-related phenomena in the built stock and foster more accurate modelling and more rational decisions.

This research progress and the output of it has been positively affected by the author's affiliation to international research and the professional community involved in the built stock energy modelling: IEA EBC Annex 70: "Building energy epidemiology". The scientific objectives, problems, and questions raised by this thesis were also examined within the collaborative workflow. The context of numerous specific problems and, subsequently, their solutions were developed in close collaboration with several industrial partners. The contextualization of the problems raised in paper I⇆ - by Trondheim Municipality, Sintef Building Energy Research and FME Zen. The problem addressed in paper II⇑ contextualized by Haugaland Kraft AS and Balcoo. Papers II⇑ and IV⇔ largely reflect on the needs for advanced analytics and data-driven decision making towards more sustainable built environment formulated by Enova and the Norwegian Water Resources and Energy Directorate (NVE). In particular, with relevance to their mission in promoting or discouraging some energy-related tendencies in the population. Enova also contextualised and supported several features currently available in the Built Stock Explorer.

2.2 Methods of probabilistic programming

2.2.1 Random variables

Probabilistic programming represents the phenomenon by means of random variable (r.v.) X, which is a function that generates real numbers $\mathbb{R}$ from either countable or uncountable sample space S:

$$X : S \to \mathbb{R} \tag{2.1}$$

If the sample space is uncountable, the r.v. X is said to be a continuous random variable (c.r.v.). Alternatively, for the countable S, r.v. X is a discrete random variable (d.r.v.). A single observable realisation (outcome) of either c.r.v. or d.r.v. X is the random variate x. Establishing the properties of c.r.v. or d.r.v. X given a set of random variates x observed in the data is the subject of statistical inference. Statistical inference is applied to the empirical distribution (Section 2.2.2) of x with certain parameters (Section 2.2.3) and consists of density estimation (Section 2.2.4) and statistical hypothesis testing (Section 2.2.5). The focus of this section is on c.r.v. since this is the type of r.v. that all the papers $I_{\Leftrightarrow}$, II$\Uparrow$, III$\Downarrow$, IV$^{\Leftrightarrow}$, as discussed in Section 2.4 and the case study (Section 2.5) deal with. The majority of metrics and methods discussed, however, are also applicable for the d.r.v..

2.2.2 Density and cumulative density

For the c.r.v. X, a Probability Density Function (PDF) $f_X(x)$ defines the relative likelihood of random variate x to occur through the entire sample space. PDF has two essential properties:

$$\begin{aligned} &1) : f_X(x) \geq 0 \quad \forall x \in \mathbb{R}; \\ &2) : \int_{-\infty}^{\infty} f_X(x)dx = 1. \end{aligned} \tag{2.2}$$

The integral of the PDF is the Cumulative Distribution Function (CDF) $F_X(x)$:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt. \tag{2.3}$$

The properties of the PDF in Eq. 2.2 define the limits of CDF: $F_X(x) \in [0, 1]$.

Given the available PDF or the CDF, the probability of c.r.v. X producing x in range $x \in [a \leq x \leq b]$ can be evaluated as follows:

$$P_X[a \leq x \leq b] = \int_a^b f_X(x)dx = F_X(b) - F_X(a). \tag{2.4}$$

Fig. 2.1 illustrates the PDF (top), CDF (bottom) and the meaning of probability that follows from Eq. 2.4.

2.2.3 Parameters of distribution

The distribution of empirical data is characterised by several parameters (summary statistics) related to central tendency, dispersion and shape. Each communicates important quantitative information measured on the sample and may further be projected to the population.

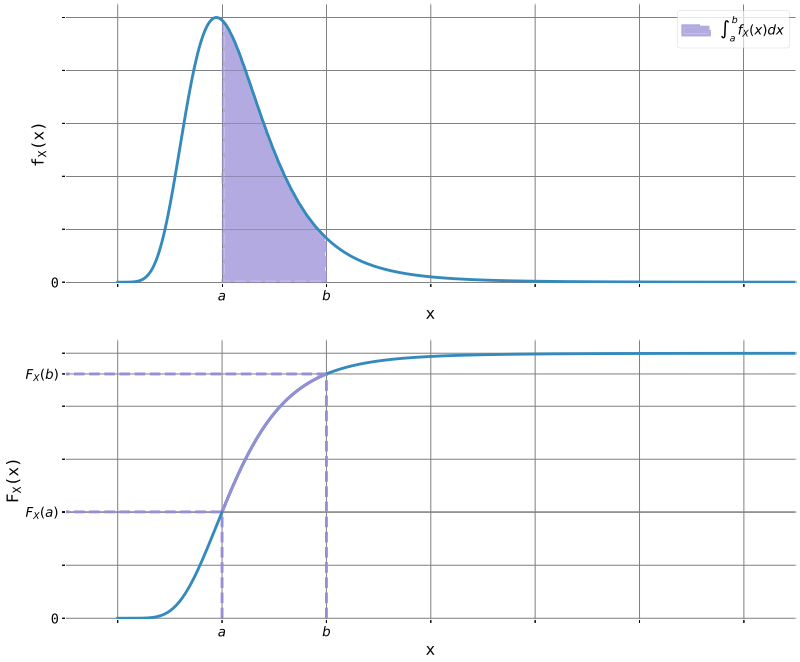


Figure 2.1: PDF, CDF and the meaning of probability

Central tendency

Central tendency suggests a particular point of the distribution where the random variates are most likely to occur and indicates the best estimates of the uncertain phenomenon. For random variates x with n observations, the central tendency is quantified by:

- arithmetic mean: $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
- geometric (normalised) mean: $g = \sqrt[n]{\prod_{i=1}^{n} x_i}$
- median - middle-most value in a sorted (parentheses (...) are used to denote sorting operation) list of values:

$$\widetilde{m} = \begin{cases} (x)[\frac{n}{2}] & if\ n\ is\ even; \\ \frac{(x)[\frac{n-1}{2}]+(x)[\frac{n+1}{2}]}{2} & if\ n\ is\ odd. \end{cases}$$

- mode - most frequent value (peak of density): $\widehat{m} = \underset{x \in S}{\operatorname{argmax}}\, f_X(x)$

Fig. 2.2 illustrates arithmetic and geometric mean, median and mode of the non-symmetric alpha-distributed random variates x.

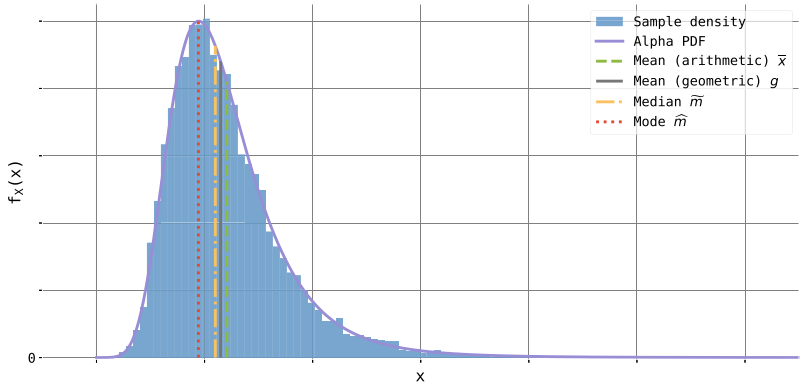


Figure 2.2: Central tendency of the distribution

Dispersion

Dispersion of the empirical distribution is a measure of a distance from some central reference towards the region with a smaller likelihood of its occurrence. Standard deviation, variance, percentiles, quartile and range are the parameters of dispersion:

- Standard Deviation (STD): $S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}$;
- variance: $S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$;
- i^{th} percentile - a value below which i percent of records can be found;
- quartile - [0^{th} (minimum), 25^{th}, 50^{th} (median $\widetilde{m}$), 75^{th}, 100^{th} (maximum)] percentiles denoted as [Q_0, Q_1, Q_2,Q_3, Q_4] accordingly;
- Interquartile Range (IQR): $IQR = Q_3 - Q_1$ - the range between 25^{th} and 75^{th} quartiles;
- range (peak-to-peak): $R = Q_4 - Q_0$ - the difference between maximum and minimum values.

Fig. 2.3 illustrates the STD, IQR and a percentile of the non-symetric alpha-distributed random variates x also shown in Fig. 2.2.

Shape

The shape of the distribution is characterised by location, scale, skewness and kurtosis. Location and scale measure the deviation from the standard form of distribution with zero and one for location and scale parameters accordingly. The location represents the horizontal shift of the distribution, whereas scale controls its width. Fig. 2.4 illustrates several variations of these parameters of normal distribution, for which the location and scale parameters correspond to the mean μ and STD σ. Normal distribution with $\mu = 0$ and $\sigma = 1$ in Fig. 2.4 is a standard normal distribution.

Given the PDF $f(x)$ of a standard form, the location- (loc) and scale- ($scale$) adjusted

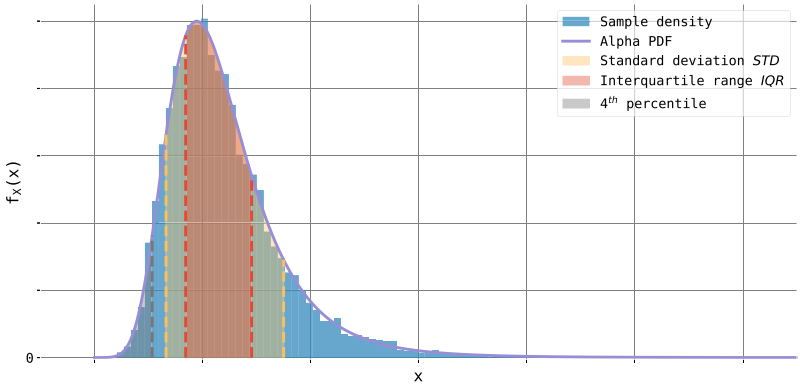


Figure 2.3: Dispersion of the distribution

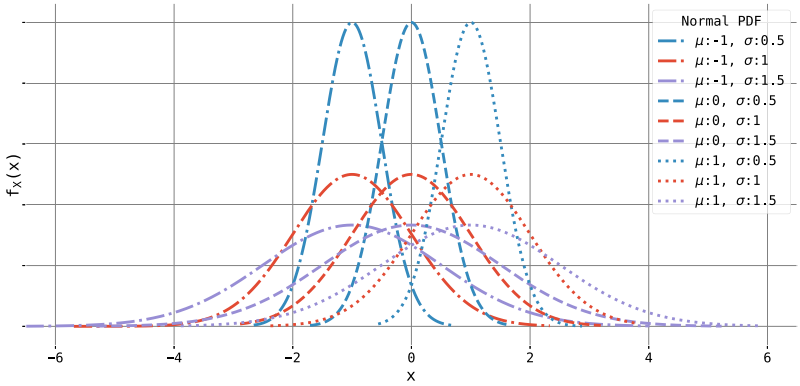


Figure 2.4: Location and scale of the distribution

PDF $\overline{f}$ is computed as:

$$\overline{f}(x, loc, scale) = \frac{1}{scale} f(\frac{x - loc}{scale}).$$

A location- and scale-adjusted CDF $\overline{F}$ is given as:

$$\overline{F}(x, loc, scale) = F\frac{(x - loc)}{scale}.$$

Skewnes characterises a non-symmetric tendency in the distribution and may be quantified through Pearson's median skewness coefficient:

$$g_p = \frac{3(\overline{x} - \widetilde{m})}{S}.$$

Positive skewness occurs if the peak of density is shifted to the right compared to symmetric (unskewed) distribution and vice-versa. The analysis of skewness, so as the comparison between mean $\overline{x}$ and median ($\widetilde{m}$), may reveal the presence of the outliers in the data. Given that the median is less affected by the outliers, skewness is negative if $\overline{x} < \widetilde{m}$. Fig. 2.5 illustrates a standard normal distribution ($\mu = 0, \sigma = 1$) with various skewness.

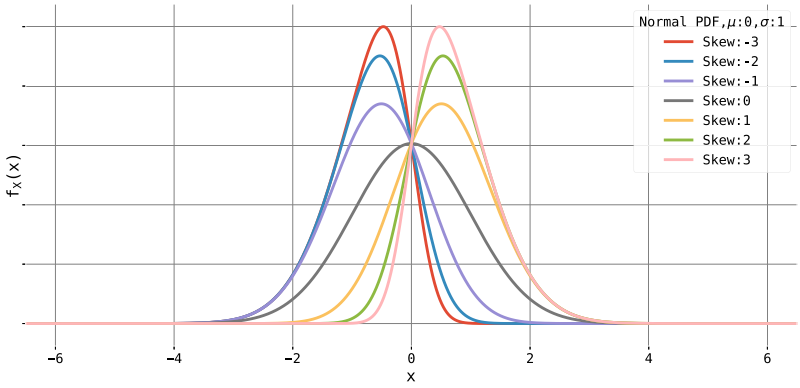


Figure 2.5: Skewness of the distribution

Kurtosis is a measure of sharpness or convexness associated with the peak of density. Using the Fisher's definition and method of moments, amongst the other ways, kurtosis may be computed as:

$$g = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2)^2} - 3.$$

Fig. 2.6 illustrates three standard ($loc = 0, scale = 1$) distributions with various kurtosis.

2.2.4 Density estimation

Density estimation is a procedure used to approximate the underlying data generation process by fitting the PDF to the empirical sample. One of the methods used for this purpose is Maximum Likelihood Estimation (method) (MLE). In MLE, the overall objective is set as follows: given the observed sample $x : [x_1, x_2, x_3, ..., x_n]$ and the specific type of distribution with PDF $f_X(x|\theta)$, find the vector θ of shape parameters that can generate such sample. This brings to the task of maximising the log-likelihood function of a form:

$$\begin{aligned} LL(\theta, x) = ln[f_X(x_1|\theta) \cdot f_X(x_2|\theta) \cdot f_X(x_3|\theta) \cdot ... \cdot f_X(x_n|\theta)] = \\ ln[f_X(x_1|\theta)] + ln[f_X(x_2|\theta)] + ln[f_X(x_3|\theta)] + ... + ln[f_X(x_n|\theta)] \end{aligned} \tag{2.5}$$

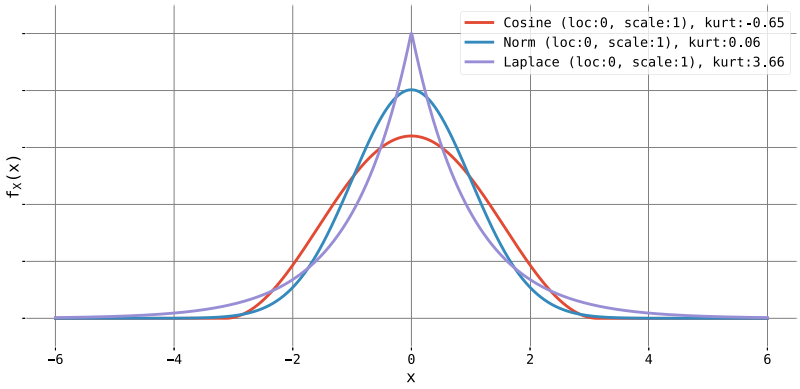


Figure 2.6: Kurtosis of the distribution

With the Eq. 2.5, the objective function of MLE is defined as:

$$f(\theta, x) = max_\theta \left\{ ln\left[\prod_i^n f_X(x_i|\theta) \right] \right\} = max_\theta \left\{ \sum_i^n ln[f_X(x_i|\theta)] \right\} \tag{2.6}$$

The solution for this objective function may be found through several non-linear optimisation techniques, e.g. Fisher scoring or gradient-based methods. Some of the most common scientific computing applications rely on downhill simplex (Nelder – Mead) (Nelder and Mead 1965) method for MLE.

2.2.5 Statistical hypothesis testing

The goal of statistical hypothesis testing, summarised by Downey (2014), is to answer: "Given a sample and an apparent effect, what is the probability of seeing such effect by chance?". The testing involves four steps:

1. selecting the test statistic that quantifies the apparent effect;
2. asserting the null- and the alternative hypotheses;
3. computing the p—value;
4. concluding about the statistical significance of the apparent effect given the available sample.

Test statistic represent the effect of interest that measures the difference between two groups. Papers III⇓ and IV$^{\Leftrightarrow}$, for example, make use of the Kolmogorov-Smirnov (test) (KS) statistic (D—value, illustrated in Fig. 2.7) which is the supremum of the absolute difference between two empirical CDFs:

$$D = \sup_x |F_{X_1(x)} - F_{X_2(x)}| \tag{2.7}$$

where:
$F_{X_1(x)}$ - CDF of sample being tested;
$F_{X_2(x)}$ - CDF of the sample against which the test is carried out.

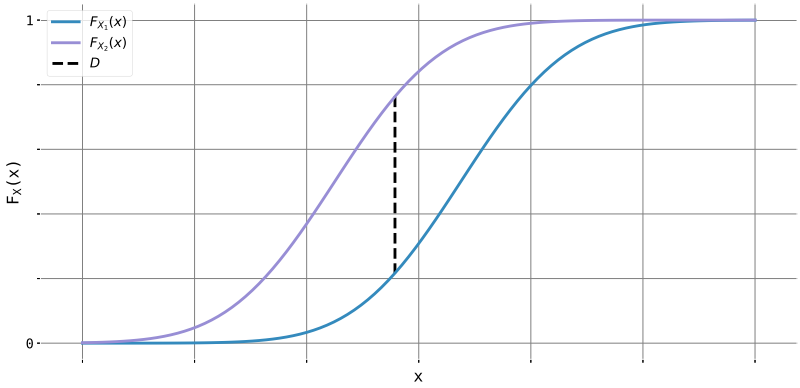


Figure 2.7: D—statistic

D—statistic highlights the non-conformity between two samples, no matter where in the range of possible values it occurs. Other test statistics may be used to quantify e.g.: i) the absolute or the relative difference between means, medians or modes; ii) the difference related to the dispersion metrics, e.g. STD and IQR; iii) correlation metrics, e.g. Pearson's or Spearman's coefficients; iv) differences in proportions and other.

The formulation of null- and alternative hypotheses asserts statements about the apparent effect at the population level. The null hypothesis typically asserts the absence of any effect measured by test statistics.

p—value measures the likelihood of obtaining the observed test statistic if the null hypothesis holds true. E.g. if the null hypothesis asserts the absence of the effect, how likely is it to observe the effect as large as the empirical samples suggest? Fig. 2.8 illustrates p—value as the probability of obtaining the observed or larger test-statistics x. Estimation of p—values involve the use of asymptotic distribution ($f_X(x)$), following the example in Fig. 2.8) of test statistic. In the case of KS statistic D (Fig. 2.7) used in papers III⇓ and IV$^{\Leftrightarrow}$, the asymptotic distribution derives from the method proposed by Marsaglia et al. (2003) implemented in `scipy.stats`.

The final step compares p—value with the established level of statistical significance α. If the p—value is small - the observed test statistic is highly unlikely under the asserted null hypothesis, the null hypothesis is rejected. On the other hand, a large p—value suggests that obtaining the observed test statistic is very likely, which leads to a failure to reject the null hypothesis at a level of significance α.

Statistical hypothesis testing deals with two kinds of potential errors and the probability of making them: Type I error rejects a valid hypothesis (false positive rate), and Type II error fails to reject a false hypothesis (false-negative rate).

2.3 Empirical data

Modelling techniques discussed in this study to various extents involve analysing experimental data. The Norwegian EPC registry is the data source that facilitated the findings

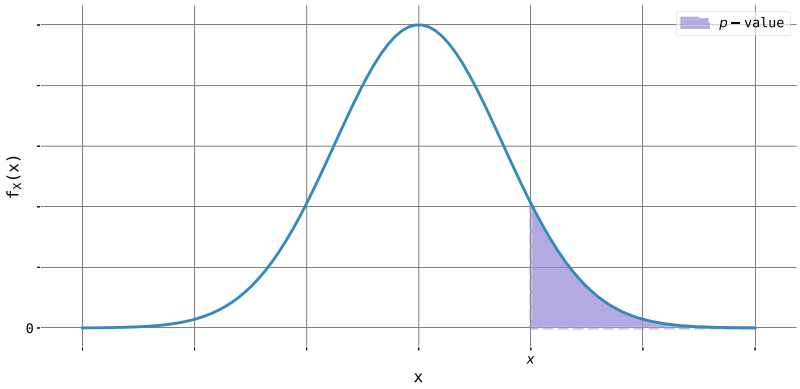


Figure 2.8: p—value

discussed further. This data is likewise essential for modelling procedures carried out in paper II⇑, paper III⇓ and paper IV⇔.

EPC dataset is one of the components of the Norwegian Energy Labelling System for Houses and Dwellings (Brekke et al. 2018), which facilitates the implementation of some requirements and recommendations originally set by the Energy Performance of Buildings Directive (EPBD) 2002/ 91/ EC - a mechanism for promoting high energy efficiency implies marketing the EPCs and economic benefit from acquiring them at the moment of selling, renting or deep renovation of the buildings. The EPC registration procedure involves reporting the geospatial location, type, age, heated floor area, primary envelope material, source of energy for space heating, ventilation system type and other building characteristics. A source-specific and total annual energy use ($kWh \cdot y^{-1}$) is subject to reporting either on a voluntarily (for residential units) or mandatory (non-residential) basis. Only those EPC records that have their annual energy use specified are taken into account in this study. Fig. 2.9 illustrates the number of EPC records (total of 73577) available per building type per 50 largest Norwegian municipalities as a colour-encoded pivot table.

Fig. 2.9 suggests which municipalities and building types are most often (light shades) certified in the Norwegian EPC registry. Thus, `Bergen, Bærum, Oslo, Stavanger` and `Trondheim` have the largest number of records. `RE. apartment` and `RE. house, detached` are the most commonly certified building types. Residential types significantly exceed non-residential ones by EPC records count.

The Norwegian EPC dataset enabled to:

- inform the expected age, the size (m^2) and the energy use intensity ($kWh \cdot y^{-1} \cdot m^{-2}$) of `RE. house, semi-detached V` located in Bergen to develop and to empirically validate a seed model discussed in paper II⇑;
- infer/parameterise univariate distributions of energy use intensity for all building typologies in Trondheim to enable random sampling in a top-down model proposed by paper III⇓.
- infer the causality between the combinations of attributes and energy use intensity of `RE. apartment` in Oslo elaborated in paper IV⇔, infer and parameterise uni-

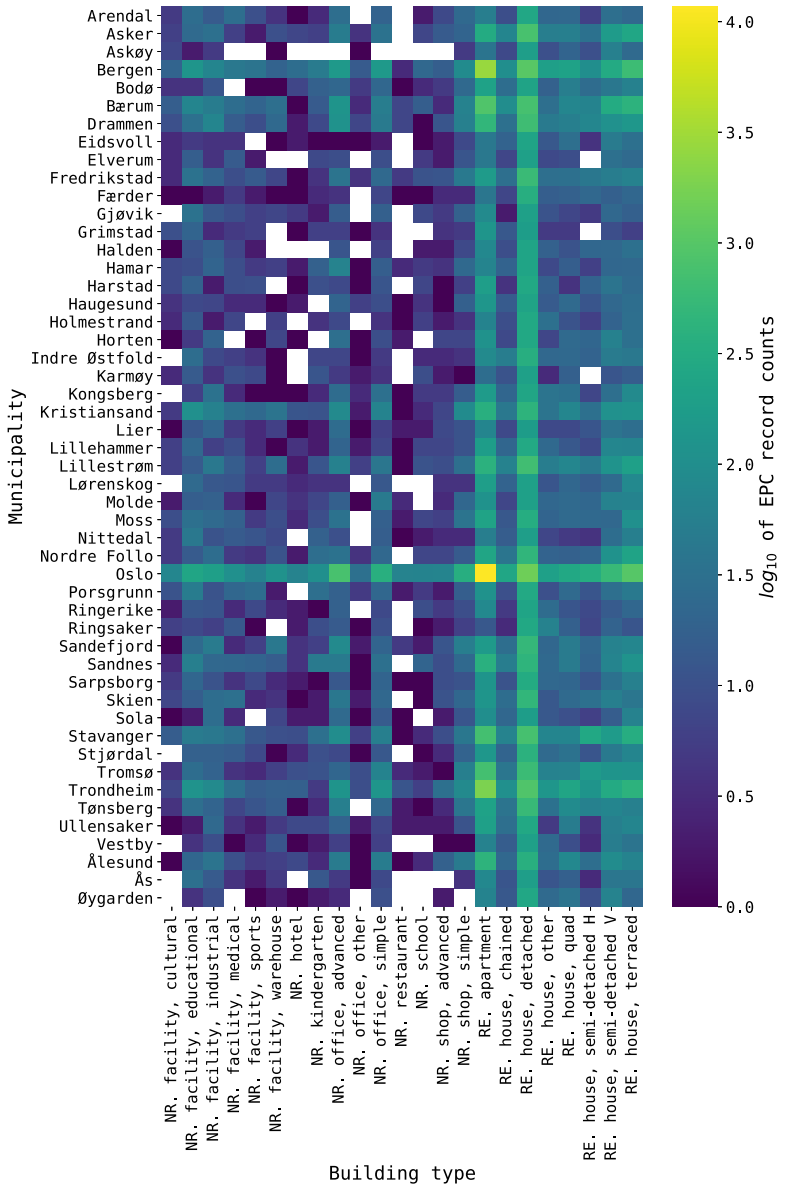


Figure 2.9: EPC records per municipality and type

variate distributions of energy use intensity for further applications in probabilistic programming;
- develop and operate Built Stock Explorer (Appendix A).

Another source of data used in paper I$_{\Leftrightarrow}$ is the historical time series on electric energy use (kWh) measured and reported by an automated metering system on an hourly basis. The case study is focused on one non-residential building managed by Trondheim Municipality, which collected and made available more than five years of historical data for the specified building.

2.4 Methods used in the articles

The methods discussed in Section 2.2 are the essential components of methodological procedures followed in each of the research papers. This section elaborates on these procedures per paper.

2.4.1 Paper I$_{\Leftrightarrow}$

This study examines the theoretical benefits of scheduling the operation of the electric batteries so that the critical grid loads are minimised. Such active demand-side management is also intended for maximising onsite self-consumption of photovoltaic energy. The approach involves three major components: i) short-term (day-ahead or week-ahead) forecasting of energy use in the building on an hourly basis, ii) forecasting of the photovoltaic system's output with the same time horizon and frequency, iii) allocating the expected photovoltaic energy to the upper part of the expected energy use profile through the day or the week. The second objective is achieved by automated periodical communication with the external forecasting service provider - Solcast 2022, and the third - by using a binary search method. The essential component with relevance to the thesis is the first, where SARIMA$(p,d,q)\times(P,D,Q)_s$ model is used for univariate time series forecasting of the phenomenon (Fig. 2.10).

Parameterising the model is based on the time series of a specific length with the most recent monitored energy use received from the monitoring system's application programming interface. The choice of the essential parameters, non-seasonal autoregressive (p) and moving-average (q) component, implies repetitive model fitting followed by performance assessment (a.k.a. cross-validation) for all combinations of reasonable values of the two parameters (a.k.a. grid search). The metrics used for performance assessment is the average for five cross-validation steps, root mean squared error (RMSE) between the model's prediction and the known values. Such a procedure enables to maximise the accuracy of model predictions. However, inaccuracies in predictions are inevitable in any practical application of time series forecasting. A better understanding of these inaccuracies (referred to in paper I$_{\Leftrightarrow}$ as residuals) may involve the analysis of their univariate distribution using the methods discussed in Section 2.2.

2.4.2 Paper II$\Uparrow$

This study suggests the methodological recipe to estimate, in the long-term, the stock-wide aggregated cost of energy and the changes in cumulative load profile caused by the acceptance of some of the emerging technologies. This approach is based on MC method where each building within the scope of modelling in every given year may accept some technology with the probability governed by $norm(\mu,\sigma)$ distribution. Once this random

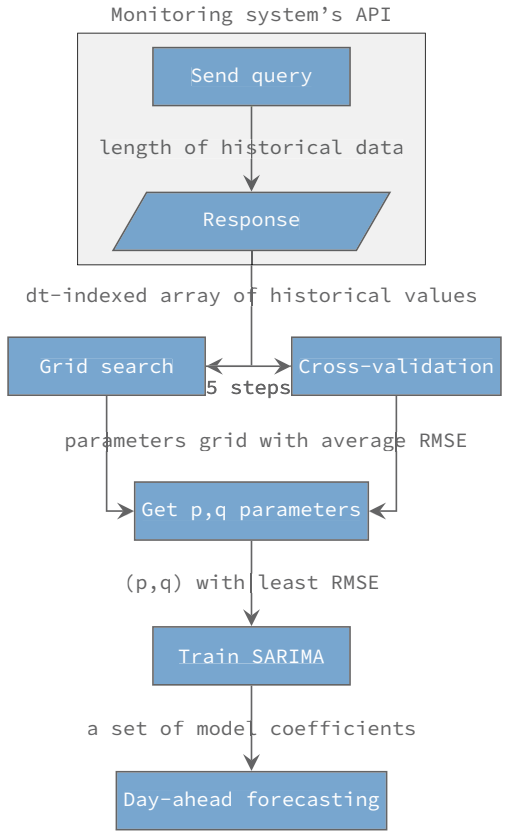


Figure 2.10: Short-term energy use forecasting (a process flow diagram)

variable draws acceptance, the exact technology is drawn from a uniformly-distributed list of technologies not yet accepted. This process is illustrated in Fig. 2.11 as a sample path for one building.

In Fig. 2.11, a sample building happens to accept all four technologies (domestic hot water tank (DHWT), solar water heating (SWH) system, electric vehicle (EV) and the photovoltaic (PV) system). With the probability of 0.07 to accept some technology in a given year, the first acceptance happens in 2022. At that moment, all four technologies are available and have an equal probability of being accepted (0.25). This happens to be DHWT. For several consecutive years after, no technologies are accepted until 2027. Three out of four technologies are available this time since DHWT is already equipped. The probability of acceptance of each of the three technologies is 0.3333, and the SWH system is drawn. The process continues until 2045, when the last available technology is accepted. Therefore, in 2050, the building will be equipped with all four.

Under the varying probability of acceptance (referred to in paper II⇑, as the techno-

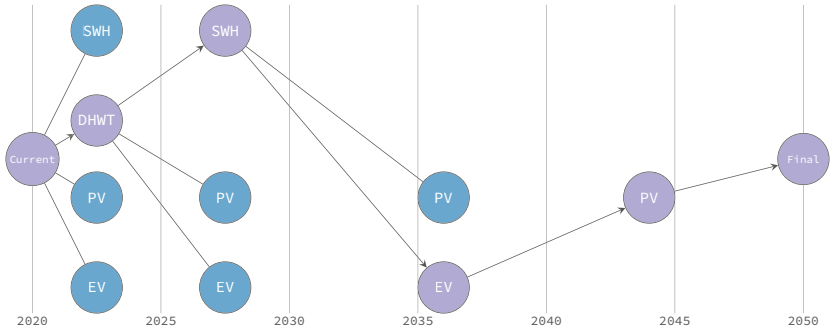


Figure 2.11: Long-term energy use prediction (a sample path)

logical acceptance rate) within each simulation trial, a sample path similar to the one in Fig. 2.11 is built for each building within the synthetic built stock composed of 1000 units. The final states of all possible paths given the four technologies correspond to one of 12 white box models of the building with a unique and technically rational combination of these technologies and the associated power/energy use profile. Thre are two outputs of each simulation trial:

- a cumululative (for 1000 units) load profile;
- a total (for 1000 units) annual energy cost (NOK) computed using the historical hourly price data for the year 2018 for each of six pricing methods.

Given many simulation trials, the former yields many likely cumulative load profiles, the latter - six alternative probability distributions examined further.

Although capable of answering some of the relevant questions, such a modelling approach has substantial shortcomings. These stem from i) idealizing the phenomenon itself, since there are substantial variations already at the individual building level and ii) the assumed homogeneity/ representativeness of the building having a specific list of properties and technologies. Fig. 2.12 illustrates the assumtption of homogeneity behind size of the building to model using the white-box principles.

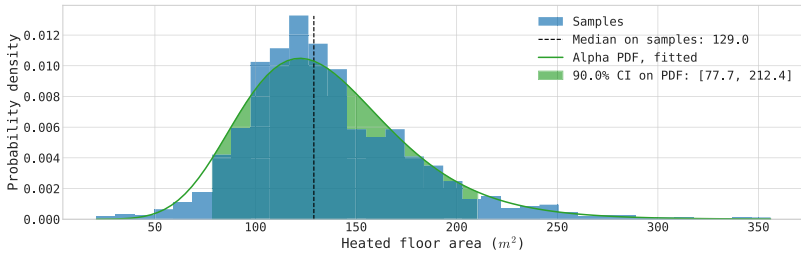


Figure 2.12: Heterogeneity building size and the homogenic abstraction from it

Fig. 2.12 shows a univariate empirical sample distribution of heated floor area (m^2)

for Semi-detached houses divided vertically and located in Bergen, Norway. The values are shown to vary in a wide range [25, - 350] m^2. A narrow range [77.7, 212.4] m^2 contains 90 % of records, making it the most representative. Should one be interested in a single value instead of a range, the parameters of central tendency (Section 2.2.3) are instrumental. Fig. 2.12, for example, illustrates a median value of 129 m^2 used in a white-box model of a building which is claimed, in the most modest manner, to have a representative size. Hence, a wide variation of heated floor area exhibited by the buildings is neglected by the model, which has negative implications on the soundness of modelling results.

Similarly to the heated floor area, such bottom-up modelling involves assumptions of homogeneity in virtually any other building attribute and adherence to patterns/standards. In contrast to the heated floor area, such assumptions are often made without the experimental data, i.e. lack the empirical basis. These problems further undermine the modelling accuracy. Additionally, such an approach requires substantial labour and computational resources. The need for methodological means to tackle these challenges in the discipline led to considering alternative model design principles. And hence, to the paper III⇓ discussed in Section 2.4.3.

2.4.3 Paper III⇓

Paper III⇓ suggests such an alternative model design. The elaborated top-down principles handle the heterogeneity of buildings and the aleatory uncertainty in their energy use. These sources of uncertainty are accommodated in the empirical dataset (Section 2.3), which is the input of the MC-based probabilistic simulation procedure. The underlying uncertainties propagate to and are reflected in the simulation results.

The accuracy of such modelling principles is mainly governed by the adequacy of the statistical model, i.e. parameterised PDF, that represents the phenomenon (energy use intensity, $kWh \cdot m^{-2} \cdot y^{-1}$) and is used to generate random numbers. Therefore, the study proposed a procedure for fitting the PDF, validating the fit and selecting the best amongst the alternatives to suit the purpose (Fig. 2.13).

The procedure illustrated in Fig. 2.13 iterates over the available building types and fetches an empirical sample with energy use intensity ($kWh \cdot m^{-2} \cdot y^{-1}$) of buildings of that type. For each of these types, each of 97 common theoretical distribution from `scipy.stats` are fitted using the MLE (Section 2.2.4) and tested using the KS (Section 2.2.5). Unless the KS yields a statistically significant difference ($p-$value ≤ 0.05), the theoretical distribution with the fitted parameters is considered a valid candidate for simulating the energy use intensity. In the case of multiple valid parameterised theoretical distributions, the one having the least test statistic (D, Section 2.2.5) is considered the best fit. The output of this procedure is the list of best-fit parameterised distributions that represent the data generation process for each of the available building types in the city.

Knowing the exact floor areas of the units of a particular type within certain geospatial boundaries and that the energy intensity that the units of this type in the city is well-approximated by the PDF found earlier, the following formula applies:

$$E_{zone\ tot} = \sum_{i=0}^{n}\sum_{j=0}^{m}(a_{i,j} \cdot r_{i,j}) = \sum_{i=0}^{n} A_i \cdot R_i^T \qquad (kWh \cdot y^{-1}) \tag{2.8}$$

where:

$E_{zone\ tot}$ - total (for all units) annual energy use of a geospatial zone;
$a_{i,j}$ - heated floor area (m^2) of j^{th} unit of the i^{th} type;

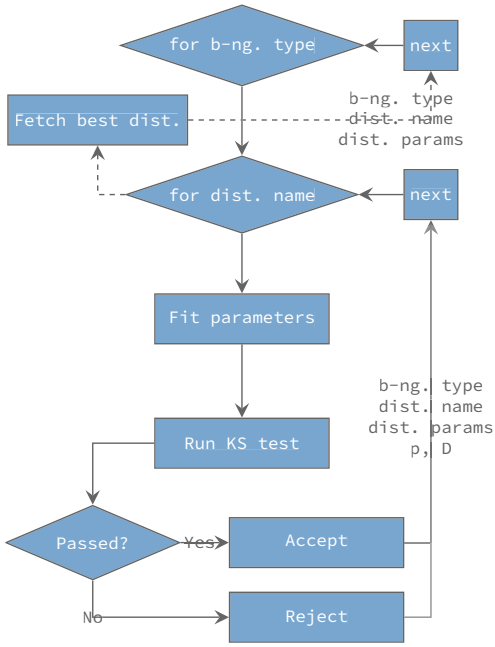


Figure 2.13: Density estimation procedure (a process flow diagram)

$r_{i,j}$ - energy use intensity ($kWh \cdot m^{-2} \cdot y^{-1}$) of j^{th} unit of the i^{th} type;

$A_i = \begin{bmatrix} a_{i,0} \dots a_{i,m} \end{bmatrix}$ - a row matrix containing the values of heated floor area (m^2) of all m units of the i^{th} type;

$R_i^T = \begin{bmatrix} r_{i,0} \dots r_{i,m} \end{bmatrix}^T = \begin{bmatrix} r_{i,0} \\ \dots \\ r_{i,m} \end{bmatrix}$ - a column matrix containing the values of energy use intensity ($kWh \cdot m^{-2} \cdot y^{-1}$) of all m units of i^{th} type, sampled randomly from the respective distribution.

A large number of runs, each of which draws a new matrix R_i^T yields a respective number of values of $E_{zone\ tot}$ distributed as $Norm(\mu, \sigma)$. These parameters μ and σ per geospatial zone are the ultimate outputs of the model.

The study articulates that the rational choice of explanatory variables is the key to effectively applying the proposed modelling approach. This rationale must reflect only the variables having a significant effect on the phenomenon at a population level. Therefore, finding and advocating the methodological means of establishing the causal effect based on empirical evidence is the objective of paper IV$^{\Leftrightarrow}$.

2.4.4 Paper IV⇔

Paper IV⇔ elaborates on the methods of inferential statistics for establishing the causal relationships between building attributes and the energy performance at a population level. These practices, referred to as explanatory modelling, are oriented towards a better understanding of the variability of the phenomenon and its relationship with building-related characteristics. The methods of density estimation (Section 2.2.4) and statistical hypothesis testing (Section 2.2.5) are therefore examined in this study.

It is also illustrated in this study that the built stock contains many distinct attributes, architectural, engineering-, occupancy-related and others. A systematic manner of examining the direct or indirect effect of these combinations on energy use for various applications of inferential analysis is necessary. The study suggests a hierarchical tree structure (Fig. 2.14) to assist such analysis and to elevate the robustness of the analytical conclusions.

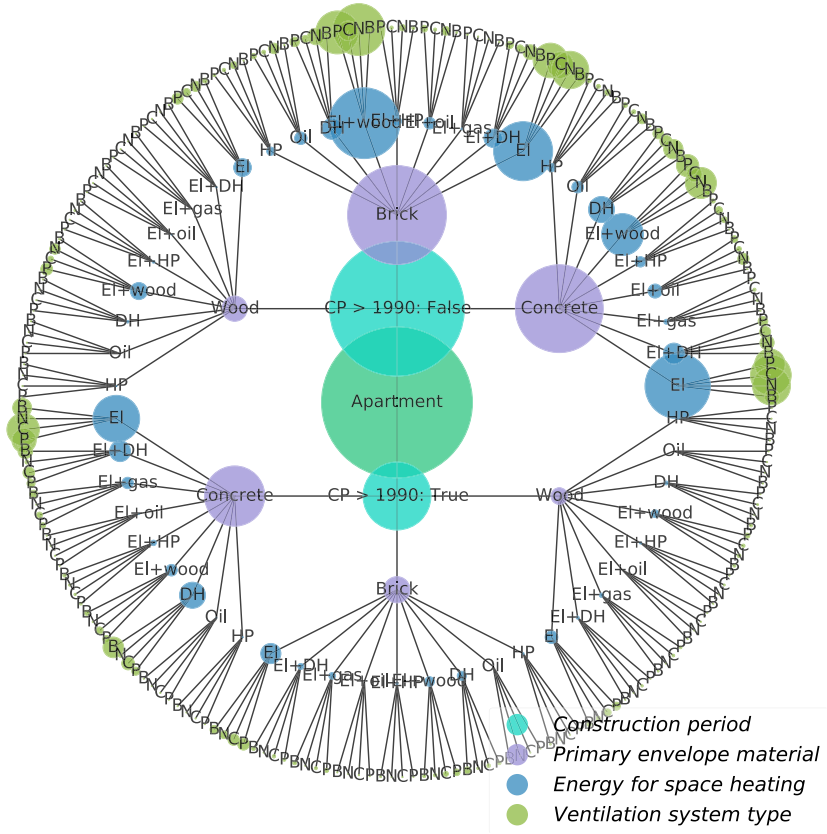


Figure 2.14: Hierarchical structure of building attributes

Fig. 2.14 represents an empirical sample of 11000 apartments located in Oslo, ac-

cessible in the empirical dataset (Section 2.3). Moving from the central node towards the peripheral (leaf) node gradually adds the exogenous categorical variables of interest shown as concentric circles: construction period (CP), building envelope material (brick, concrete or wood), source of energy for space heating (Electric (El), district heating (DH), gas, oil, heat pump (HP) or a combination of these) and the ventilation system type (balanced (B), continuous (C), natural (N) or periodical (P)). Every distinct path from the centre to the peripheral node represents a unique, theoretically possible combination of building attributes. A number of all possible combinations of building attributes is given by the cardinality of the n—fold cartesian product of the values of the categorical variables. The cardinality of the cartesian product is the product of cardinalities of each set that it forms:

$$|A \times B \times C \times ...| = |A| \cdot |B| \cdot |C| \cdot ... \tag{2.9}$$

Thus, two CPs, three different envelope materials, nine unique combinations of space heating solutions and four ventilation system types result in $2 \cdot 3 \cdot 9 \cdot 4 = 216$ combinations illustrated in Fig. 2.14. The figure also reflects upon the sample size that corresponds to the combination relative to the total (11000) number of records through the diameter of the node.

2.5 Case study design

The case study is designed to arrive at the estimates of the phenomenon in: i) bottom-up and ii) top-down manner. The phenomenon of interest is a bulk (for all buildings) total (across all energy sources) annual energy use ($kWh \cdot y^{-1}$) for an arbitrary number of buildings within certain spacial boundaries and having a given typology, similarly to papers II⇑, III⇓.

2.5.1 Bottom-up engineering-based simulation

Bottom-up modelling makes use of detailed building energy performance simulation carried out with IDA-ICE, where the choices behind designing the seed model (also used in paper II⇑) are based on the available experimental data, expert judgements and the common knowledge. Given the availability of empirical data (e.g. Section 2.3), some of these choices may be aided by the analysis of the central tendency in the univariate distribution, as elaborated in Section 2.2.3.

The occupancy pattern, as one of the most important and uncertain components in modelling of this kind, is addressed in this study by means of combinatorial analysis. Each of 100 identical units within the synthetic built stock is expected to follow one out of five empirically-defined occupancy schedules illustrated in Fig. 2.15. A number (N) of all possible distinct combinations of $n = 5$ schedules that $k = 100$ buildings may have is computed as shown in Eq. 2.10. The exact schedules that the buildings within a certain combination assume may be listed in a lexicographically ordered array (Eq. 2.11).

$$N = \frac{(n-1+k)!}{(n-1)!k!} = \frac{104!}{4! \cdot 100!} = 4598126 \tag{2.10}$$

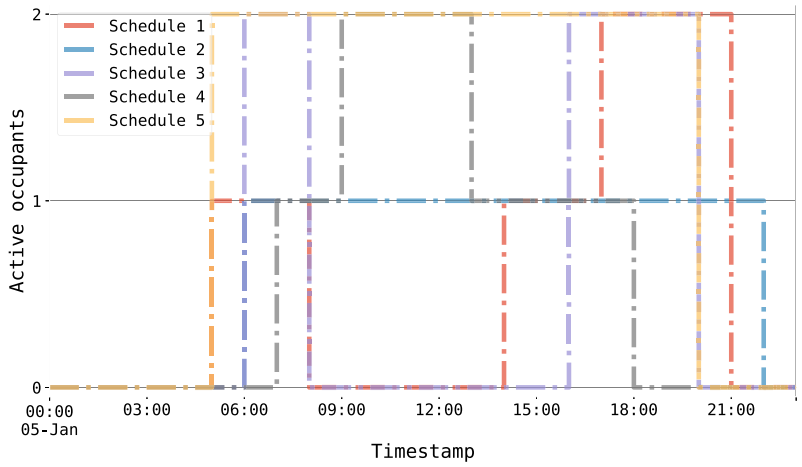


Figure 2.15: Daily occupancy patterns

$$
\begin{array}{lc|cccccc}
 & & \multicolumn{6}{c}{building:} \\
 & & [1] & [2] & \dots & [98] & [99] & [100] \\
\hline
 & [1] & 1 & 1 & \dots & 1 & 1 & 1 \\
 & [2] & 1 & 1 & \dots & 1 & 1 & 2 \\
 & [3] & 1 & 1 & \dots & 1 & 1 & 3 \\
 & [4] & 1 & 1 & \dots & 1 & 1 & 4 \\
combination: & [5] & 1 & 1 & \dots & 1 & 1 & 5 \\
 & [6] & 1 & 1 & \dots & 1 & 2 & 2 \\
 & [7] & 1 & 1 & \dots & 1 & 2 & 3 \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & [4598124] & 4 & 4 & \dots & 5 & 5 & 5 \\
 & [4598125] & 4 & 5 & \dots & 5 & 5 & 5 \\
 & [4598126] & 5 & 5 & \dots & 5 & 5 & 5
\end{array}
\tag{2.11}
$$

Combination [1] in Eq. 2.11, for example, implies that all buildings follow Schedule 1. This combination is distinct from [2], where 99 buildings follow Schedule 1 and one building has Schedule 2... For each combination in Eq. 2.11, bulk total annual energy use of 100 buildings (E_{BU}) may be computed as:

$$E_{BU} = \sum_{i=1}^{100} e_i \qquad (kWh \cdot y^{-1}) \tag{2.12}$$

where:

e_i - total annual energy use ($kWh \cdot y^{-1}$) of the i-th building, having a specific occupancy schedule.

The values of E_{BU} computed for each combination in Eq. 2.12 form an array of length 4598126 which is the uncertain output of a bottom-up model. This output has a univariate distribution with some shape and some parameters that are the subjects of analysis in Chapter 3 using the methods described in Sections 2.2.3 and 2.2.4. The estimates at the aggregated level are obtained by combining the models of individual system components, i.e. buildings, hence bottom-up reasoning.

2.5.2 Top-down probabilistic programming

Similarly to Eq. 2.12, the top-down model computes the bulk total energy use by aggregating the total energy use of individual buildings:

$$E_{TD} = \sum_{i=1}^{100} e_i \qquad (kWh \cdot y^{-1}) \tag{2.13}$$

In this case, however, the values of e_i are sampled randomly using a Pseudo-Random Number Generator (PRNG). The underlying distribution used for sampling these numbers is identified based on the empirical data (Section 2.3) that characterise the phenomenon within a given geopolitical scope and under specific architectural and technical properties of the buildings. The phenomenon at the aggregated level is estimated by knowing the distribution of the phenomenon at an even more aggregated scope, hence the top-down approach. Identifying/parameterising these distribution, as discussed in paper III⇓, necessitates density estimation (Section 2.2.4).

The values of E_{TD} in Eq. 2.13, computed within 1000 MC trials, compose an array of length 1000, which is an uncertain output of the model. A univariate distribution of this output is likewise examined in Chapter 3, based on several parameters (Section 2.2.3) and density estimation (Section 2.2.4) method.

Chapter 3

Results

3.1 Results of the individual papers

3.1.1 Paper I$_{\Leftrightarrow}$

The application of electric batteries has the potential to mitigate peak loads on the power grids and may have substantial positive economic implications. The advocated battery scheduling practice implies controlled discharging during the peak demand and controlled charging at the time of high PV output. The procedure discussed in Section 2.4.1 serves such a purpose and is capable of reshaping the grid interaction profile, similar to the illustration in Fig. 3.1.

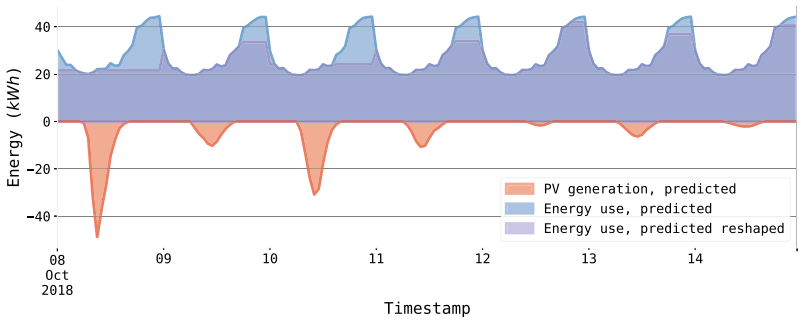


Figure 3.1: Reshaped grid interaction profile

Fig. 3.1 exemplifies how electric batteries, through the proposed scheduling procedures, may enable allocating the photovoltaic energy generated through the day to the time of peak demand on that day, thus reshaping the energy use profile. Hence, the predicted reshaped profile has daily peaks reduced by the amount of PV energy expected to be available through the same day.

The predictive accuracy of building energy demand and PV output forecasting methods govern the rationality of such scheduling approaches. The former, even with the accuracy and robustness-oriented model design and parameter selection illustrated in Fig. 2.10,

is debatable. Furthermore, paper $I_{\Leftrightarrow}$ quantifies and examines predictive (in-)accuracy of the time series model intended for forecasting building energy use. These inaccuracies are the manifestation of epistemic and aleatory uncertainties associated with the phenomenon and the model. The former stems from not considering the other potentially useful explanatory variables. The latter reflects the phenomenon's potentially important aspects of a random nature. Fig. 3.2 enables examining the model performance by illustrating: i) a subset of the original time series of electric energy use in the building; ii) the time series of this phenomenon forecasted by SARIMA $(0,1,4) \times (1,1,1)_{24}$, the parameters for which are found to yield the most accurate predictions; iii) the series of differences between model predictions and the original series at a given timestamp, a.k.a residuals. A univariate distribution of these residuals and the MLE-fitted normal distribution are illustrated in Fig. 3.2 (bottom).

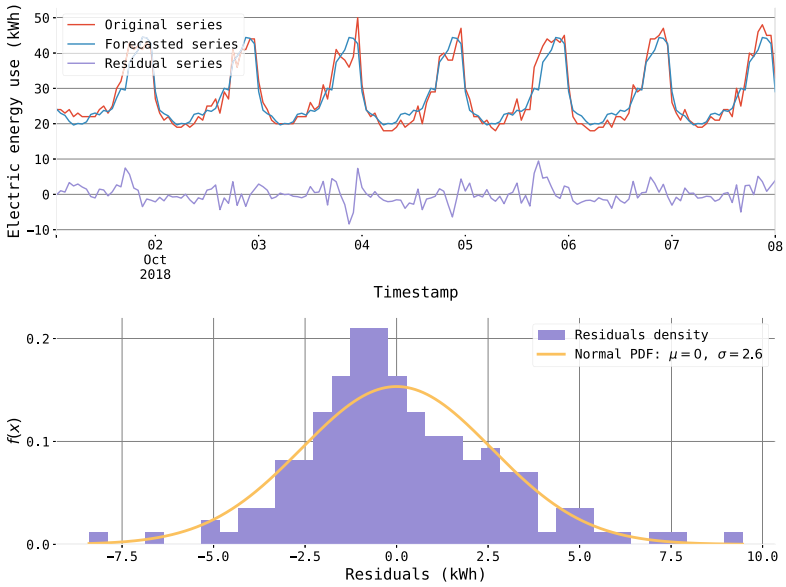


Figure 3.2: Time series and the univariate distribution of the residuals

Fig. 3.2 suggests the phenomenon, although fairly predictable by the model, exhibits non-systematic, i.e. without apparent patterns or tendencies, fluctuations. A univariate distribution illustrates several properties of the relationship between the model and the phenomenon:

1. Zero mean confirms that systematic tendencies in the residuals are absent, which is the ultimate goal since modelling by definition implies approximations;
2. Negative skewness suggests that negative residuals occur more frequently than positive, but the latter appear to have larger values. They are, as for the case study, mutually compensating (hence zero mean);
3. The measure of dispersion in the distribution conveys the magnitude and the fre-

quency of the occurring residuals.

These results demonstrate the modelling uncertainties that occur at a building level. The accuracy of predicting the phenomenon may vary depending on the wealth of factors associated with the building and the model selected/designed/parameterised to represent it. In the case of bottom-up modelling, these uncertainties (aleatory and epistemic) are propagating to the stock level.

3.1.2 Paper II⇑

The scope of modelling carried out in the article is a stock of 1000 untits that are:

- semi-detached houses divided vertically
- located in Bergen
- constructed before 1990
- featuring a technically-feasible combinations of [EV, PV, DHWT, and SWH]

The units with each combination of [EV, PV, DHWT, and SWH] are represented (modestly) by the white-box models developed in IDA-ICE. The expected deployment follows the sample paths similar to the one in Fig. 2.11 with the probabilities of deployment drawn from $norm(\mu = 0.05, \sigma = 0.0015)$. The results are:

- 1000 cumululative load duration curves (Fig. 3.3);
- 1000 values of total annual energy cost (NOK) for each of six alternative energy pricing methods (Fig. 3.4).

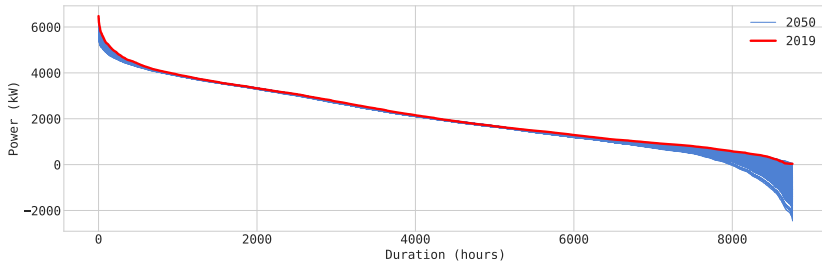


Figure 3.3: Aggregated load duration curves

Fig. 3.3 suggests that the noticeable changes, driven by the acceptance of the listed technologies up until 2050, may occur with the highest/lowest power demand and the grid feed-in, whereas the rest of the load duration remains unchanged. Hence, a minor, up to 13%, decrease in the peak demand should be expected, with a duration of fewer than ≈1000 hours per annum. An already low power demand will be reduced further, but this change will affect the no more than ≈2000 hours per annum (as shown in range $\approx [6760, 8760]$ in Fig. 3.3). It is likewise expected that the peak PV feed-in will affect the duration of ≈760 hours ($\approx [8000, 8760]$) and will not exceed ≈2000 kW.

Fig. 3.4 shows that, given the acceptance of the listed technologies and the annual patterns that the energy cost exhibits today, the time-of-use tariff is expected to be the cheapest in 2050. On the other hand, a maximum power extraction-based model with variable elements will be the most expensive. All other pricing methods are midway

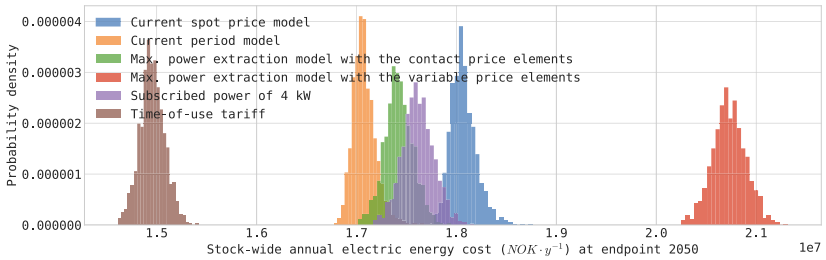


Figure 3.4: Aggregated electric energy costs per pricing model

between these two extremes, with less substantial differences and often overlapping densities.

The model design choices are intended to address its purpose by accommodating the novel technologies without knowing: i) if, ii) when, and iii) which of these technologies will be accepted in each building within the modelling scope. However, the assumptions that i) 1000 buildings have identical characteristics and follow the same patterns (homogeneity) and ii) the actual energy use or power demand of a single building strictly adheres to those modelled (Section 3.1.1 demonstrates that they indeed vary), make these conclusions rationally doubtful.

3.1.3 Paper III⇓

The procedure described in Section 2.4.3 is seen as a parsimonious top-down model suitable for numerous applications, energy mapping (Fig. 3.5) being one of them.

The map in Fig. 3.5 shows mean (μ) and STD (σ) of a normal distribution formed from the simulated bulk total annual energy use per 1×1 km geospatial grid over Trondheim. The apparent, in Fig. 3.5, areas of high energy use are the densely populated and commercially active historical centre and the industrialised southern territory. This conclusion stems from the noticeably higher colour intensity, which encodes the mean in the distribution of the simulated results. Larger estimates of energy use are typically associated with larger dispersion which is conveyed by the STD and encoded in Fig. 3.5 as the diameter of the marker.

Such results may facilitate identifying and eventually resolving urban energy hotspots in the best interest of the targeted social, environmental and economic objectives. However, the main objective of this article is to advocate the parsimony, scalability, and uncertainty-awareness that emerge from such a top-down approach.

A parsimonious, scalable and accurate model design must ignore the variables that have little or no effect on the phenomenon at the population level. Making robust, evidence-based conclusions about the causal effect is necessary to achieve these benefits. This represents a substantial analytical challenge of examining and documenting non-conformities in the empirical data. Non-conformities in the phenomenon exhibited by distinct groups often manifest the potential causal effect that is present. Fig. 3.6 demonstrates such a case.

Fig. 3.6 [A] displays a univariate distribution of energy use intensity ($kWh \cdot m^2 \cdot y^{-1}$) of apartments located in Trondheim with the construction year (CY) in range [1800, 2018]. The empirical sample size (SS) is 1842 units available in the EPC dataset (Section 2.3).

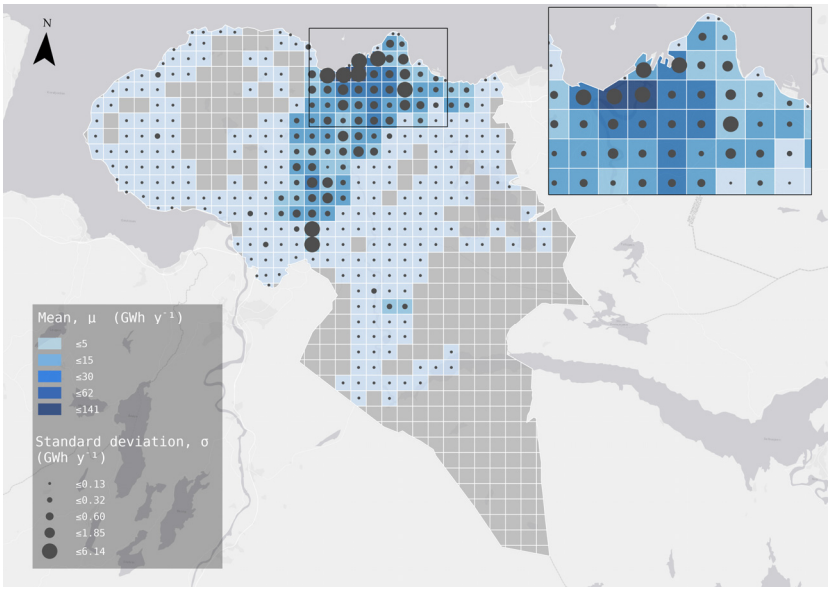


Figure 3.5: Spatial variation of bulk total annual energy use in Trondheim

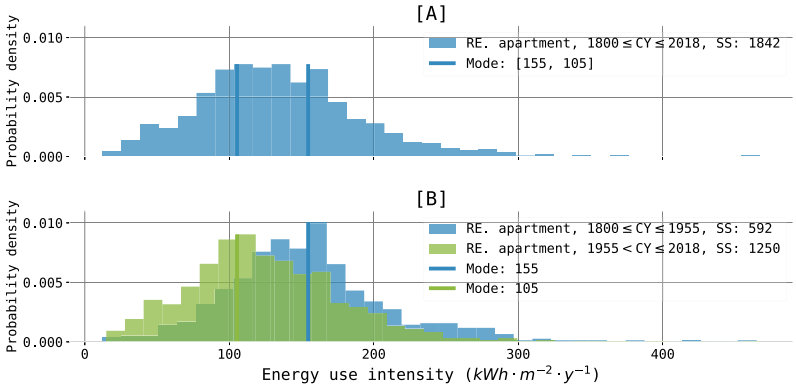


Figure 3.6: The effect of the construction period on building energy performance

This distribution has two modes (at 155 and 105 $kWh \cdot m^2 \cdot y^{-1}$). Multi- or bi- (in this case) modality is a common indicator of obscured patterns. Indeed, Fig. 3.6 [B] demonstrates that the observable bimodality is explained by the distinct construction periods of these apartments. Fig. 3.6 [B] shows the same empirical sample separated into two groups,

for the apartments constructed before and after 1955. The univariate distribution of the two groups has distinct statistical properties (mode being one of them) and suggests that newly built apartments are more energy efficient. Although it is tempting to accept this conclusion, statistical practices require answering: "could the observable effect occur by chance?" in addition to that. The p—value answers this question. The effect, in the case of KS test, is measured by the D—statistic (Section 2.2.5). A KS test carried out for the two groups shown in Fig. 3.6 [B] results in a large D—statistic (0.256, also shown in Fig. 3.7) and a negligibly low p—value ($1.212 \cdot 10^{-23}$). If there was no effect of construction period on the energy intensity, observing a test statistic that large given the (also large) sample sizes is very unlikely (as indicated by a small p—value). The asserted null hypothesis: "populations of apartments in Trondheim, constructed before and after 1955 are not different in terms of energy intensity", is rejected at a significance level of 0.05. It is therefore concluded that the energy intensity does differ within the scope of the study.

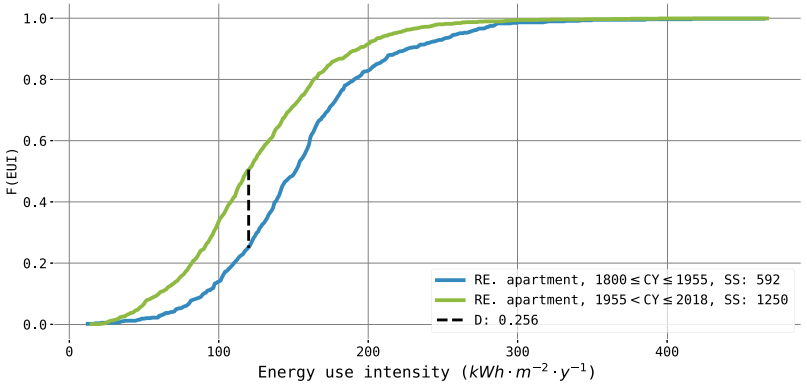


Figure 3.7: The effect measured by D-statistic

The construction period, although often important, is only one of many attributes/factors that are believed to affect the energy performance at a population level. Moreover, this effect tends to vary across cities, typologies and buildings having the other attributes in common. Is the conclusion about the effect of the construction period made above applicable to apartments equipped with heat pumps? Or, which apartments perform better: the old ones made of concrete or the new ones made of wood? What if they are, in addition to that, located in distinct districts of the city? Therefore, large-scale energy research needs a structured approach to establishing and documenting the causal effects. And hence, one of such approaches (illustrated in Fig. 2.14) suggested by paper IV$^{\Leftrightarrow}$.

3.1.4 Paper IV$^{\Leftrightarrow}$

Paper IV$^{\Leftrightarrow}$ found and documented the causal effect of four attributes of apartments in Oslo on their total energy performance: i) construction period, ii) primary envelope material, iii) space heating solution and iv) ventilation system type. Thus, the apartments in Oslo constructed after 1990 tend to perform better than those constructed earlier. The apartments having their envelope made of concrete perform better than those of brick or

wood. The apartments heated by district heating, heat pump or oil perform better than those with electric, electric and wood, or electric and gas-based solutions. There is also an apparently large and significant difference in the energy performance of apartments ventilated naturally, with periodical, continuous and balanced types.

Understanding the causal effect of individual attributes is essential for promoting or discouraging some tendencies in the built stock. It is, however, emphasised that the phenomenon is governed by the combinations of attributes rather than the attributes individually. Therefore, explanatory modelling for practical applications, such as rationalising the energy planning and the legislative mechanisms, needs to be carried out per combination. And hence, the combination-wise analysis may be facilitated by the hierarchical structure shown in Fig. 2.14.

One of the applications of explanatory modelling is demonstrated in the study and attempts to determine the combinations of attributes performing best and worse. A KS test is carried out between i) empirical sample represented by each leaf node (Fig. 2.14) provided that the sample contains more than the arbitrarily established minimum sample size of 20 records and ii) the composite sample of apartments in Oslo represented by the central node having more than 11000 records.

A favourable configuration is associated with a large positive D-statistic between the two ECDFs. For this difference to be considered significant, the p—value associated with the test should be below (also arbitrary set) significance threshold $\alpha = 0.05$. A reverse procedure, focused on large negative D—statistic only, enables the identification of the worst-performing building configurations. Fig. 3.8 and Fig. 3.9 illustrate 62 ECDFs for all configurations and highlights the best- (Fig. 3.8) and worst-performing (Fig. 3.9) relative to the composite if the difference is found to be significant.

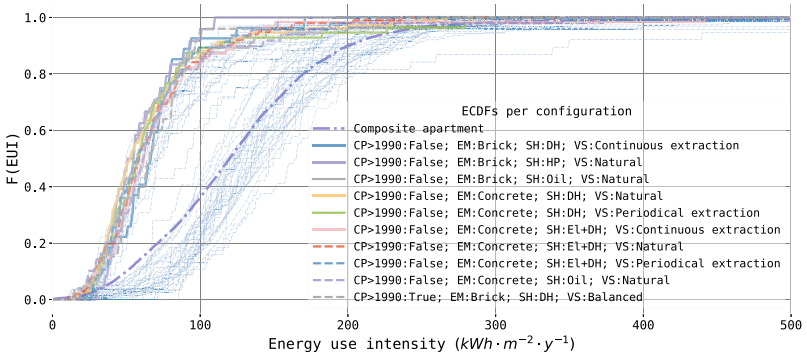


Figure 3.8: ECDFs for combinations of attributes that perform best

Fig. 3.8 suggests that the best-performing combinations of attributes of apartments in Oslo are heated by district heating alone or with electric radiators, heat pumps or oil. These combinations feature all ventilation system types, with natural ventilation appearing most frequently. Most of these combinations have concrete as the primary envelope material, brick is less frequent, and wood is absent. The majority of the best-performing combinations are constructed before 1990.

Fig. 3.9 illustrates that the poorly performing apartments in Oslo are heated by electric

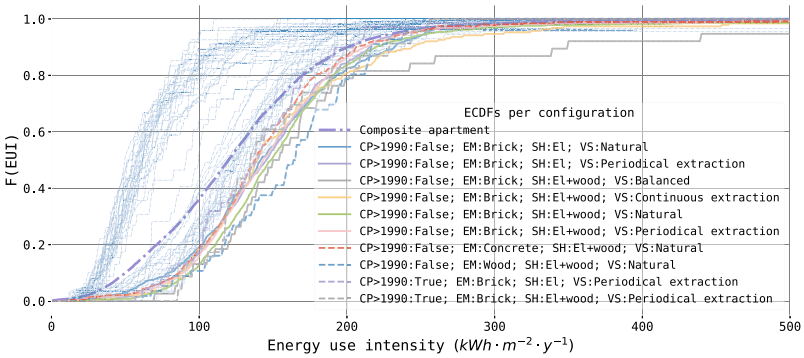


Figure 3.9: ECDFs for combinations of attributes that perform poorest

or a combination of electric with wood-based space heating solutions. These combinations involve all ventilation systems, most frequently constructed with bricks before 1990.

3.2 Case study results

Two models discussed in Section 2.5 are designed to estimate the bulk total annual energy use of an arbitrary number of buildings within certain geospatial boundaries. The case study applies these models to estimate bulk total energy use of 100 units of `RE. house, semi-detached V` constructed in Bergen, Norway. Sections 3.2.1 and 3.2.2 discuss the modelling results. Section 3.2.3 quantifies the differences between them. The key differences in handling uncertainties 3.2.4 and the appropriateness criteria 3.2.5 are examined further. Section 3.2.6 highlights the role of explanatory modelling in rationalizing the design of these predictive models.

3.2.1 Bottom-up engineering-based simulation

Bottom-up reasoning necessitates designing a seed model that reflects the targeted weather-/climate, typology and detailed architectural and engineering specifications. Because the detailed specifications are not given by the modelling objectives, these need to be determined to the best of the modeller's awareness and, whenever possible, validated by the empirical data. The associated labour, data and computational resource intensity often limit the number of seed models that are rational to design. Furthermore, these models must be representative, i.e., reflect the characteristics found amongst the buildings modelled, with negligible or infrequent deviations. An IDA-ICE model, mentioned in Section 2.5.1 and paper II⇑, is designed in such a way, has the building characteristics considered representative and therefore reused in this study. The heated floor area of 122 m^2, for example, matches the central tendency (Section 2.2.3) in the distribution of heated floor area of `RE. house, semi-detached V` in Bergen (according to the EPC registry (Section 2.3) which may be validated through Built Stock Explorer). The age category (construction year $<$ 1990) contains the prevailing 80% of EPC records for a given typology in the

municipality. Space heating with electric radiators is known to be the most common solution for detached/semi-detached houses in Norway. Therefore, such choices of building and occupancy characteristics are targeting the most representative case.

The total annual energy use by the building obtained through the IDA-ICE simulation runs is `[19189,18516,18297,18075,17645]` $kWh \cdot y^{-1}$ for occupancy schedules `[1,2,3,4,5]` (illustrated in Fig. 2.15) accordingly. The aggregation of these results to the stock level of 100 units follows the Eq. 2.12 for each possible combinations listed in Eq. 2.11. The model output is 4598126 likely values having a normal distribution characterised by mean (μ) and a standard deviation (σ) further discussed in Section 3.2.3.

3.2.2 Top-down probabilistic programming

Top-down reasoning requires an empirical sample of total energy use of the units that match the targeted geopolitical scope and the building typology. The EPC registry (Section 2.3) contains `293` records for `RE. house, semi-detached V` in Bergen. This data describes the buildings of varying size, age, architecture- and engineering-related features.

Sample density estimation with MLE (Section 2.2.4) per distribution available in `scipy.stats` following the procedure and the goodness-of-fit metrics ($D-$ statistic and $p-$ value discussed in Section 2.2.5, paper III⇓ and paper IV$^{\Leftrightarrow}$) led to the choice of logistic distribution (Eq. 3.1) parameterised by $loc = 18713$ and $scale = 3853$. The fit is characterised by $D = 0.026$ and has a corresponding p-value of 0.984. The location- and scale-adjussted logistic PDF is defined as:

$$f(x, loc, scale) = \frac{e^{\frac{-(x-loc)}{scale}}}{s(1 + e^{\frac{-(x-loc)}{scale}})^2} \tag{3.1}$$

The intermediate result of density estimation is illustrated in Fig. 3.10 and may be reproduced in Built Stock Explorer. The figure suggests that the fitted distribution approximates the cumulative sample density without noticeable deviations. This distribution, therefore, may be used for generating synthetic inputs of the model in Eq. 2.13. The output of this top-down probabilistic programming procedure is 1000 values that the bulk total annual energy use at a stock level of 100 units may be. This output likewise forms a normal distribution parameterised by mean μ and STD σ (Fig. 3.11).

3.2.3 The (mis)match of the numerical outputs

The results achieved using i) bottom-up engineering-based simulation supplemented by combinatorial analysis (Section 3.2.1) and ii) top-down probabilistic programming with prior parametric univariate density estimation (Section 3.2.2) approaches are illustrated in Fig. 3.11. Both uncertain outputs are normally distributed, which follows from the Central Limit Theorem applicable due to the summation operation in Eq. 2.12, and 2.13, but have somewhat distinct parameters.

Fig. 3.11 suggests that the outputs of bottom-up and top-down models conform. Central tendency in the distribution of the outputs characterises the best estimates of bulk total annual energy use. The central tendency, in normal distribution, is well represented by mean value which also corresponds to median and mode. The relative difference between mean values of two distributions is rather small: $\delta = \frac{\mu(E_{TD}) - \mu(E_{BU})}{\mu(E_{TD})} \cdot 100 = \frac{1871405 - 1834440}{1871405} \cdot 100 = 1.975\%$. The dispersion of the model output conveys the level of uncertainty in the results. A measure of dispersion STD of the bottom-up model is substantially smaller, which stems from distinct kinds of uncertainty considered by the models

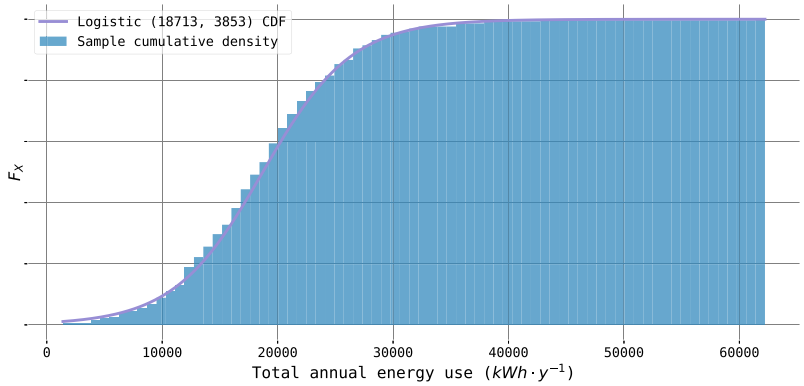


Figure 3.10: Density estimation results

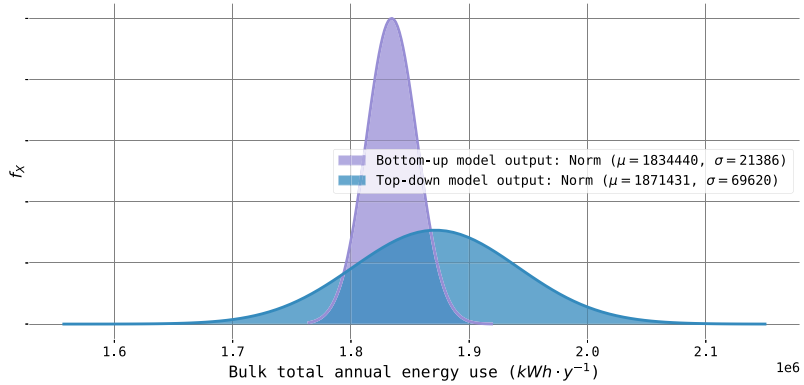


Figure 3.11: Uncertain outputs of bottom-up and top-down models

(elaborated in Section 3.2.4). Under the absence of additional information, it is rational to conclude that the true value of the phenomenon is in the area where the densities of two distributions overlap. In Fig. 3.11, this region is limited by the PDF of the bottom-up model output.

3.2.4 Handling of uncertainties

A smaller dispersion associated with the bottom-up model may (erroneously) assert more certainty in it. It must be emphasised, however, that this metric is a sole characteristic of the bottom-up model, not the quality of the phenomenon's representation by this model. Thus, the bottom-up modelling procedure exemplified in this study reflects only the epi-

stemic uncertainty behind the choice of occupancy patterns. The applicability of these schedules, given the diversity and the dynamics exhibited by real buildings, is not examined here. Addressing a wealth of other uncertainties, epistemic, aleatory and heterogeneity (Section 1.3), may substantially enlarge the dispersion of the outputs. Prioritising the uncertain elements considered by bottom-up modelling procedures may be facilitated by examining their causal relationship with the phenomenon in question, i.e. building energy performance. Papers III⇓ and IV$^{\Leftrightarrow}$ elaborate on the rationale and the instrumental aspects of establishing such causal relationships.

Alternatively, a top-down model (Section 2.5.2) is not prone to most of such uncertainties since: i) the experimental data already reflects the underlying randomness and heterogeneity and ii) model parsimony does not allow for most of the errors of the epistemic kind. The sources of uncertainty specific to probabilistic programming applications stem from the quality of data and density estimation (paper III⇓). To further reduce the dispersion of the results, empirical data must reflect geopolitical, behavioural, architectural and technical attributes of the building that have a causal effect on building energy performance (papers III⇓ and IV$^{\Leftrightarrow}$).

3.2.5 Critical assessment of model appropriateness

These modelling approaches, in light of their appropriateness criteria (Section 1.2), have the following differences:

1. *accuracy:* Both approaches require measurements to verify the accuracy of the model. Top-down modelling relies on the experimental data, which ensures that the output takes on reasonable values. The advantage of the bottom-up approach is making predictions without the data or if the data (only) partially explains the phenomenon. This property is particularly useful in modelling rare or extreme events. However, since stock-wide modelling deals with large numbers of buildings, rare outliers do not affect the output substantially. More frequent outliers, on the other hand, are likely to occur in the experimental data and, therefore - accommodated by the top-down modelling principles;
2. *sensitivity:* The model must be sufficiently sensitive to the parameter of interest to examine the implication of these parameters. In bottom-up modelling, sensitivity is the model's characteristic and can be examined through repetitive model runs with the systematic altering of these design parameters. With top-down reasoning, the causal effect of design parameters on the phenomenon governs the model's sensitivity to it. I.e, if there is no empirical evidence of the design parameter affecting the phenomenon - implicit examining and modelling this parameter is irrational (papers II⇑ and IV$^{\Leftrightarrow}$). Statistical hypothesis testing (Section 2.2.5), may facilitate the choice of sensitive parameters;
3. *versatility:* Virtually any number of design parameters can be explicitly accounted for in the bottom-up model, it is therefore only limited by the creativity/skills/resources of the modeller. However, the presence of the parameter does not necessarily improve the performance of the model. The top-down approach accounts for all design parameters implicitly and, if necessary, for the (statistically) significant ones explicitly;
4. *parsimony* (governs *transferability/reproducibility, computability, usability)*: bottom-up modelling, as shown by this study, is substantially more labor-, information- and computationally-intensive. A seed model is often case-specific since it needs many (often manual) inputs. The demand for computing to carry out physics-based simulation is further amplified by the need for a large number of runs within the uncer-

tainty and sensitivity analysis procedures. Such complexities undermine the transparency by adding black-box elements to the originally white-box model, leaving room for misinterpretation of the results. Alternatively, the implementation of the top-down approach is rather parsimonious since only a few steps are involved: i) navigating through the stock-level data to find the subset of interest; ii) carrying out density estimation; iii) running a MC simulation. Such modelling results are almost identically reproducible given that the same experimental data was used. The demand for computational resources remains modest, which favours making the models accessible via, e.g. web (similarly to Built Stock Sxplorer).

3.2.6 The effect of other design parameters

Formulating, testing and validating the causal hypotheses about the relationships between design parameters and energy performance may, in light of uncertainty (Section 3.2.4) and appropriateness criteria (Section 3.2.5), assist with: i) prioritising the uncertain factors necessary to account for in the models and elevating their accuracy; ii) governing the model's sensitivity to the design parameter and the rational level of versatility; iii) minimising model's complexity and thus, elevating their reproducibility, computability and useability.

The construction period, for example, as one of the potentially plausible model design parameters, is seen distinctly by bottom-up and top-down models. With bottom-up modelling, the construction period is a proxy for architectural and engineering measures made to comply with the energy performance standards active during the historical period with their implications on the energy performance. The other socio-economic, technological, and climate-related variables are assumed fixed. Likewise, unless explicitly modelled, the state of renovation is not accounted for. Top-down reasoning exemplified in this study sees the construction period as a mixture of all variables directly or indirectly associated with it and their relevance to the energy performance. This perspective accounts for the chance that the building is in the renovated state at the moment of data registration and for all the less apparent socio-economic tendencies in the built stock. It may, for example, happen that the older buildings tend to host more occupants or that the occupants of the older buildings tend to spend more time at home and, therefore, are likely to use more energy. However, living densely in the old house could be a matter of choice for low-income occupants, those who want and do save energy because of the associated costs. Could it also happen that the occupant age category, their professional or cultural background dominates in some building vintage or in some way determine the energy use? Hence, explanatory modelling is needed to answer, following this example: "Given all the socio-cultural, physiological, economic and other features exhibited by the occupants of today, and all the ways that they may interact with the building, do the new buildings tend to consume less energy?" The answer applies to the natural status quo of the built stock and should inform the design of both bottom-up and top-down models.

The null and the alternative hypothesis (Eq. 3.2) are defined to answer if "the populations of `RE. house, semi-detached V` constructed in Bergen [1] before and [2] after 1990 have distinct total annual energy use". The answer is given at the level of statistical significance $\alpha = 0.05$.

$$H_0 : E_1 - E_2 = 0;\ H_a : E_1 - E_2 \neq 0 \tag{3.2}$$

where:
E_1 - total annual energy use ($kWh \cdot y^{-1}$) of the population [1];
E_2 - total annual energy use ($kWh \cdot y^{-1}$) of the population [2].

The empirical data (Section 2.3) contains 293 records for `RE. house, semi-detached V` in Bergen representing two distinct construction periods (CP) of the certified units. The ECDFs for the two samples having the sample size (SS) exhibit a somewhat observable difference illustrated if Fig. 3.12.

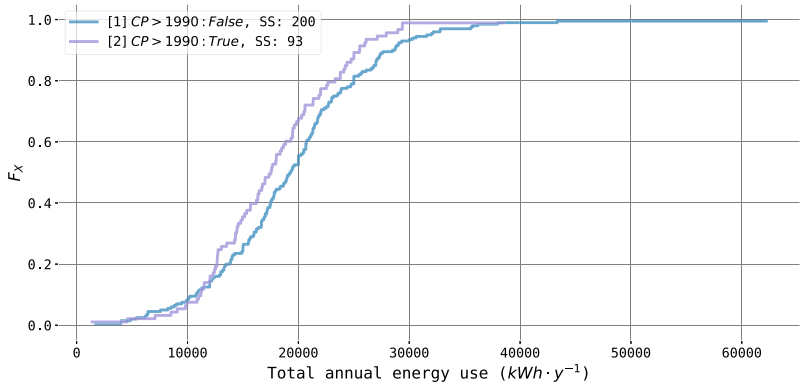


Figure 3.12: ECDFs of the empirical samples by construction period

The KS test (Section 2.2.5) with two empirical samples returns $D = 0.145$ and $p = 0.121$. The null hypothesis asserting the conformity between the two populations (Eq. 3.2) measured by D—statistic (Eq. 2.7) cannot be rejected since $p > \alpha$.

Lack of evidence against the null hypothesis, given the sample size, may have multiple reasons and their combinations, such as:

- Envelope-related:
  - In practice, energy standards/policies concerning the new buildings did not improve their thermal performance significantly
  - In practice, energy standards/policies concerning the old buildings did improve their thermal performance significantly
  - A large share of the old buildings were renovated and/or re-engineered
  - The new buildings may tend to be larger
- Occupant-related
  - There may be a demand for better comfort and/or more (powerful) appliances in the new buildings
  - The density and the time of occupancy in the old buildings may exceed those of the new buildings
  - The awareness of poor energy performance of the old buildings may trigger extra efforts in saving the energy and vice versa

This is a list made for illustrative purposes without attempting to make it comprehensive. The claims about the effect of any of these items on the phenomenon without further analysis are avoided in this study.

Because the effect that is apparent in Fig. 3.12 lacks statistical significance, making the models (Sections 2.5.1 and 2.5.2) sensitive to the construction period is not expected to

improve their performance. The opposite also holds - the design parameters of strong and significant effect on building energy use should be of primary concern in model design.

It is concluded above that there is no empirical evidence that the energy use of `RE. house, semi-detached V` in Bergen constructed [1] before and [2] after 1990 differ. Does the same conclusion apply to the other cities? Table 3.1 contains the results of a KS tests carried out to see if energy use of `RE. house, semi-detached V` constructed [1] before and [2] after 1990 differs in the other Norwegian cities. The tests result in D—statistic and p—values for the two samples representing distinct construction periods (CP) with the sample sizes (SS) shown in the table. The null and the alternative hypotheses (Eq. 3.2) are tested at a significance level $\alpha = 0.05$,

	City	D-statistic	p-value	CP>1990(SS)	CP≤1990(SS)
0	Asker	0.369	0.000	61	102
1	Askøy	0.538	0.036	51	7
2	Bergen	0.145	0.121	93	200
3	Bærum	0.099	0.646	67	238
4	Drammen	0.323	0.006	46	63
5	Horten	0.300	0.395	10	50
6	Kristiansand	0.385	0.001	41	70
7	Lillehammer	0.264	0.137	30	43
8	Lillestrøm	0.202	0.141	64	57
9	Molde	0.185	0.636	22	39
10	Nordre Follo	0.114	0.801	54	63
11	Oslo	0.246	0.000	217	341
12	Sandefjord	0.357	0.089	46	15
13	Sandnes	0.334	0.032	33	35
14	Stavanger	0.238	0.043	45	117
15	Tromsø	0.341	0.001	56	73
16	Trondheim	0.270	0.001	65	232
17	Tønsberg	0.294	0.077	38	32
18	Ålesund	0.155	0.756	70	22
19	Øygarden	0.363	0.469	59	5

Table 3.1: The effect of construction period on energy use of `RE. house, semi-detached V` measured by KS for several cities

Table 3.1 suggest that whereas for some cities, the effect of the construction period on the energy use is found i) strong and significant (Asker, Askøy, Drammen, Kristiansand, Sandnes, Tromsø), for some of them it is ii) weak but significant (Oslo, Stavanger, Trondheim), iii) strong but insignificant (Horten, Sandefjord, Tønsberg, Øygarden) and iv) weak insignificant (Bergen, Bærum, Lillehammer, Lillestrøm, Molde, Nordre Follo, Ålesund). Collecting additional data for the groups with insignificant effects, iii) and iv), may elevate the significance. Such variations of the effects between the cities are likely to reflect socio-cultural, economic and climate-related factors.

Paper IV$^{\Leftrightarrow}$ found, in a similar manner, a weak but significant effect of the construction period (before and after 1990) for apartments in Oslo. It also advocates explanatory modelling that considers multiple design parameters for practical applications.

Chapter 4

Discussion

4.1 Modelling for sustainability

This thesis and its supplementary materials emphasise that accurate, sensitive, versatile and parsimonious modelling is capable of assessing the potential developments of built stock and its energy performance. Such modelling may support an effective course of action to achieve the performance targets envisioned in sustainable future communities in the most pragmatic and timely manner. Furthermore, to achieve "multiple benefits" of energy efficiency - a comprehensive system of indicators proposed by Campbell et al. (2014) and revised by Fawcett and Killip (2019), while overcoming the critical barriers towards energy efficiency of buildings identified by Vogel et al. (2015) at contextual-, sector- and project-levels.

These objectives are, likewise with relevance to research question 1, directly or indirectly contributing to several sustainable development goals (SDGs):

- SDG7 (afforable and clean energy)
- SDG11 (cities and human settlementts)
- SDG13 (climate change)

This thesis covered several applications of modelling that may facilitate more sustainable practices and policymaking:

- Mitigating peak loads and peak feed-ins while elevating self-consumption and self-sufficiency of buildings that generate renewable energy, as shown in Section 3.1.1 and (paper $I_{\Leftrightarrow}$);
- Raising the awareness of the potential changes in the interaction between buildings and the power grids under the future deployment of distributed renewable energy technologies and the growth of the electric vehicles fleet. And facilitating the development of the rational energy pricing methods that meet the expectations of the energy consumer and the supplier (Section 3.1.2 and paper II⇑).
- Finding/quantifying the urban energy bottlenecks of high priority and potential for future energy efficiency and flexibility improvement advocated in Section 3.1.3 and paper III⇓).
- Determining which architectural and technical configurations of buildings perform well at a population level (Section 3.1.4 and paper $IV^{\Leftrightarrow}$). And hence: i) quantifying, validating and documenting the effectiveness of energy policies and practices already in place; ii) developing more rational energy policies/practices that focus

on promoting the favourable configurations while discouraging those proven to perform poorly; and iii) predicting the future transformations of the built environment with their implications on the energy performance.

The first and the last items in this list rely on modelling instruments not yet accommodated by the established model classification systems into binary bottom-up or top-down. Time series forecasting concerns the short- or long-term future development of the phenomenon and does not require distinct levels of the information available and the information modelled. The predictions, uni- (3.1.1 and paper $I_{\Leftrightarrow}$) or multivariate, are based on the past observation of the same variable. Likewise, the applications of inferential analysis (Sections 3.1.3, 3.1.4, 3.2.6, papers III$\Downarrow$ and IV$^{\Leftrightarrow}$) study the phenomenon at the population level without up- or downscaling the geospatial (or any other) scope. These are the instruments capable of invigorating the large-scale building energy research and should find their place in the model classifications (Section 1.2).

4.2 Variability in modelling

This section informs research questions 2 by generalising the findings related to variability of energy performance given the distinct modelling levels addressed in Sections 3.1.1 through 3.1.4 and the papers $I_{\Leftrightarrow}$, II$\Uparrow$, III$\Downarrow$, IV$^{\Leftrightarrow}$ accordingly. This complements the analysis of variabilities in bottom-up and top-down modelling 3.2.4 exemplified by the case study. Quantification of variability and/or judgements of model uncertainty at each level is facilitated by parameters of univariate distribution introduced in Section 2.2.3.

4.2.1 Building level modelling

A perfect forecast is achieved if the model's predictions match the actual energy use at any given time. However, any practical application of time series forecasting is associated with the residuals, i.e. the bias between model prediction and actual values. The predictive success of the model is measured through goodness-of-fit (minimum residuals) on the test set. The sole application of the model, therefore, is associated with uncertainties. These residuals (Fig. 3.2) are the manifestation of variability in the phenomenon and epistemic uncertainty in representing it in the model. The modelling objective is to obtain symmetric normal residual distribution with minor STD and the absence of critical outliers. Large dispersion suggests that the model does not accommodate a large portion of variability. A non-zero central tendency - that the model systematically under- or overestimates the phenomenon, thus the parameterisation problem. Paper $I_{\Leftrightarrow}$ also emphasises that high performance on the test set does not guarantee that the model will be suitable in the long term. Degradation of model performance is therefore natural and requires systematic recalibration.

Parametrising the model is prone to error because a particular set of parameters may be valid only over a particular subset of the time series. This is the reason for k—fold cross-validation, which implies model fitting and testing over k subsets of the historical data. To ensure the choice of model parameters amongst the other, testing each reasonable combination of them at each of the k cross-validation steps is carried out through the grid search. Considering more complex models with a larger number of controllable parameters will increase (linearly) the dimensionality of the grid and (exponentially) the number of combinations given by the Cartesian product of these parameters.

The residual analysis, therefore, builds the awareness of the natural variation of the phenomena and epistemic uncertainty behind the ignorance of exogenous factors of in-

fluence not considered by the model. The latter may be addressed by using models appropriate for multivariate time series forecasting, e.g. SARIMAX, or in combination with white-box models. However, these decisions need to be justified considering parsimony versus the expected gain of the residual variance reduction.

4.2.2 Bottom-up modelling

Bottom-up modelling, represented by Section 3.2.1 and paper II⇑, is prone to all three sources of uncertainty.

A model built with IDA-ICE is a white-box model meant to represent the energy use of the building. This phenomenon, however, is a combination of white-box and black boxes. Because the latter is poorly accounted for (idealised by asserting specific patterns, e.g. presence/behaviour of the occupants in Fig. 2.15), variations of the phenomenon are expected (aleatory uncertainty).

The building selected for the seed model is meant to represent the archetype in terms of size, age and other attributes. A considerable heterogeneity behind the assumption of representativeness is present even though the choice of the archetype is targeting the central tendency in the empirical sample distribution (e.g. Fig. 2.12). Likewise, a strong assumption is made about the adherence to one of the schedules illustrated in Fig. 2.15. This challenge may be partially resolved by defining (more) alternative schedules, although it adds model complexity and computational load.

Additionally, the uncertainty arises with not knowing if and which technology will be accepted in the future, i.e. epistemic type. Section 2.5.1 and paper II⇑ propose a solution to account for the uncertain future states through twofold probabilistic simulation: i) simulating if any of the listed technologies will be accepted in a certain year (by random sampling a technological acceptance rate from binomial distribution) and ii) simulating the exact technology that is accepted by random sampling from a uniform discrete binominal distribution of the available technologies. This procedure, illustrated in Fig. 2.11, has limitations that stem from: i) the list of available technologies, their technical variations, and installation options is never exhaustive; ii) in reality, the choice between the technologies is non-uniform since it reflects socio-economic patterns and the preferences of the decision-makers; iii) the acceptance rate may have more dynamic tendency due to societal developments.

The outputs of such discrete event simulation procedure form a distribution, e.g. Fig. 3.4 and 3.11, that is subject to the parameter and density estimation (Sections 2.2.3, 2.2.4).

4.2.3 Top-down modelling

Top-down approaches, represented by paper III⇓ and Sections 2.4.3, 2.5.2, 3.1.3 and 3.2.2, have an advantage of naturally handling variabilities and aleatory uncertainty in the status quo of the built stock. Fig. 3.6 and 3.7 illustrate such variabilities. For studies of this kind, epistemic uncertainties are behind density estimation (Fig. 2.13) results, which is affected by the availability and quality of empirical data and the goodness of fit. The available data also has implications on determining building attributes and factors identified as of strong or weak significant effect (Fig. 3.6 and 3.7). The choice of statistical significance threshold (Section 2.2.5) is also recognised as highly subjective. Long-term forecasting would entail additional sources of epistemic uncertainty similar to those associated with bottom-up modelling. The output of such top-down model is a univariate

distribution (e.g. Fig. 3.11), subject to parameter and density estimation (Sections 2.2.3, 2.2.4).

4.2.4 Built stock level modelling

Similar to the above. The variability observable in stock-level empirical data is the manifestation of inherent randomness and heterogeneity (Section 3.2.6, paper IV$^{\Leftrightarrow}$). Fig. 3.8, and Fig. 3.9 illustrate, for example, the extent to which the energy performance varies given the combinations of building attributes (individual ECDFs) and for the typology (composite ECDF). These variabilities are quantified and documented by the theoretical distribution fitted through density estimation (Section 2.2.4) procedure. Theoretical distributions then serve as the means to inform bottom-up or top-down models. The established effect of the model design parameter on the energy performance (through statistical hypothesis testing, Section 2.2.5) is likewise important to design the most rational predictive, bottom-up or top-down model. The epistemic uncertainties in analytical conclusions made at a stock level stem from data availability and quality, potential estimation bias, and the (mis)interpretation of the results.

4.3 The rationale for explanatory modelling

Papers III$\Downarrow$ and IV$^{\Leftrightarrow}$ advocate formulating and testing the causal theories about the effects of building design parameters on the energy performance, which is the answer to research question 3. Applying such practices, following the research question 4, facilitates i) a better understanding of the nature of the phenomena and ii) improving the predictive models.

Knowing if, and to which extent, building energy performance at a population level is affected by certain attribute may assist with (re)designing and rationalising the energy policies and the means for their implementation. This thesis contains several relevant examples of such knowledge discovery. Section 3.1.3 and paper III$\Downarrow$ suggest that the construction period (before and after 1955) has a strong ($D = 0.256$) effect on the energy use intensity of apartments in Trondheim, as illustrated in Fig. 3.6 and 3.7. Section 3.1.4 and paper IV$^{\Leftrightarrow}$ document the strength of the effect of several building attributes and their combinations (Fig. 2.14) on the energy use intensity of apartments in Oslo (Fig. 3.8 and 3.9). The construction period (before and after 1990) is shown in paper IV$^{\Leftrightarrow}$ to have a weak significant effect on the phenomenon for this typology. The case study results (Section 3.2.6) identified a weak insignificant effect of the construction period (before and after 1990) on the energy use of `RE. house, semi-detached V` in Bergen. This effect, however, varies amongst `RE. house, semi-detached V` for the ofther cities (Table 3.1).

Hence, the factors of strong significant effect should be prioritised in order to achieve the energy-related objectives. The factors of apparent strong effect but so far without statistical significance should be examined further. The factors of weak significant effect are likewise of interest, although secondary. Care must also be taken to systematically examine the factors earlier claimed as insignificant since the rapid transformation exhibited by the built environment may soon elevate their significance.

The validated causal effects may likewise improve the predictive modelling practices. It is rational to make the models sensitive to the attributes having a strong effect. Doing so makes the models i) less prone to overfitting and/or ii) more parsimonious, with the positive implications on data-, computing- and labour intensity behind the modelling.

4.4 Conformity of bottom-up and top-down modelling

The potential (non-)conformity of bottom-up and top-down predictive model results is raised by research question 5. With the ultimate practical goal of political support, the value of a predictive model is in certainty about its prediction. Similar results achieved with two fundamentally distinct modelling approaches are expected to assert such certainty. Therefore, mutual verification of the models and the approaches is advocated in this study to build confidence about the expected consequences of the decisions. The opposite also holds. A failure to mutually verify the results of built stock energy models may warn about potential inaccuracies and thus, prevent misleading and/or ineffective decision-making. Such practices of verifying the modelling results reinforce the linkage between the theory of domain-specific modelling and the practical application of the resulting knowledge.

The case study demonstrates that, numerically, bottom-up and top-down modelling leads to fairly similar results. This is a concise example of interoperability between distinct modelling paradigms to stimulate their mutual invigoration in the built stock energy research, policy and practice. However, making such concise examples necessitates omitting numerous details that the models and the purposes they address may be subject to. Therefore, the scope of modelling and the model coverage is one of the limitations identified for this study. The results are not universally applicable to all modelling objectives and urban system components. The case study develops bottom-up engineering-based and top-down probabilistic programming-based models for predicting the bulk total energy use at a stock level given specific exogenous information. Under the other modelling objectives and resource availability, developing two distinct approaches could be irrational or impossible. Or, this could require additional considerations, e.g. multiple continuous explanatory variables, density estimation involving the discrete random variables, and the use of bayesian as an alternative to frequentist-based statistical inference. The complexities that the discipline intends to address require more inclusive yet parsimonious modelling techniques, possibly amongst those already used in the other disciplines. And, most likely, those having the grey shades.

It follows that both bottom-up and top-down reasoning may enable accurate, sensitive and versatile modelling practices. Top-down approaches naturally handle the key uncertainties and have multiple benefits associated with addressing the problems more parsimoniously. On the other hand, bottom-up engineering-based methods address the physical modelling components at the expences of model parsimony and may neglect non-physical aspects, which are often of primary importance. Therefore, the choice between such bottom-up and top-down modelling necessitates answering "should the model include a detailed analysis of the physical aspects?" and, more importantly, "does the physical aspect indeed affect the phenomenon at a population level?". Explanatory modelling, as discussed in Section 4.3, answers these questions likewise.

4.5 The quality of experimental data

The quality of data and the ways of handling it with machine learning and classical statistics remains essential for maximising its value in the built stock energy research and well beyond. The precise meaning of data with its numerous subquestions is the fundamental question of predictive analytics. It is known that the weaknesses in study design or data quality occur more often than poorly performed analysis (Breiman 2001; Miller 2014). And that experimental design has crucial implications for correctly projecting the sample-

based analytical conclusions on the population, which is of strong emphasis in classical inferential statistics but is commonly missing in the machine learning community. The other questions are: is the data potentially biased by data collection/measurement procedures, data management tasks and the interpretation of findings? Distortion of measurements, errors in readings, and reporting and registering of the data affect the reliability of conclusions further based on it. So do the strategies for recovering the missing or incomplete data, detecting and dealing with outliers.

Many of such problems are relevant to this work, hence research question 6. Ideally, statistical inference and predictive analytics require a random sample resulting from a robust study design. Substituting a study design with the already available EPC dataset (Section 2.3) for this purpose involves:

- sampling bias;
- measurement/reporting error.

Sampling bias occurs when statistical properties of the collected sample, for some reason, do not replicate those of the population. The available sample, containing certified units only, may differ from the population having both certified and non-certified units in numerous ways. Thus, for example, certification of the new buildings is mandatory while the old buildings are certified at the moment of renovation, rent or sale only. Therefore, the proportion of new to all units differs between the EPC dataset and the population. Likewise, there may be the tendency that the occupants of the certified (or the new) buildings are wealthier/poorer, older/younger, more(less) frequently absent, more(less) demanding for comfort, more(less) equipped with appliances, more(less) environmentally conscious... Certification could also occur more frequently in some regions of the urban territory (having distinct climate-related and socioeconomic tendencies) amongst the buildings of specific sizes or any other attributes.

The measurement/reporting error applies when the data is intentionally or unintentionally misreported. For example, a good certificate elevates the property's market value, and hence the owners may be tempted to report better energy performance. Could it also happen that some of the readings/energy bills were unavailable, partially because the owner and the occupant are not necessarily the same person?

The EPC registry also contained numerous flaws that stem from poor data registration and management practices. These resulted in common outliers, duplicates, and missing values that are dealt with in this study.

Although the applicable analytical toolbox may enable mitigating such problems to some extent, their prevention is a more promising strategy. Designing, managing, and developing the systems that involve experimental data collection, such as EPC is likely to benefit from the involvement of energy analysts. This may leverage the capabilities of predictive analytics to assist with answering the important questions about energy sustainability.

Poor quality of empirical data is an essential aspect of this research, making the results, under the strongest criticism, applicable to the certified buildings rather than the built stock. However, the most important results of this work are the illustrated instrumental, methodological and philosophical means to achieve the synthesis of various modelling techniques.

This study agrees with Brøgger and Wittchen (2018) and Sousa et al. (2017) on data quality and the transparency of models that use it as crucial for their applications in political support. Research reproducibility/replicability are likewise emphasised as of high importance. Following, exemplifying and encouraging these practices is one of the objectives set for the development of Build Stock Explorer. The future development of this software

will make a more comprehensive set of analytical instruments available. Along with the growing volumes of the available data, these instruments facilitate interactive explanatory and predictive modelling to achieve the objectives of large-scale building energy research (Appendix A).

Chapter 5

Conclusions

A wealth of essential questions about the energy sustainability of the built environment are yet to be answered, which demands a better understanding of the phenomenon and better accuracy in predicting it. While the need for the latter is well-recognised and represented through either a bottom-up or top-down approach, no sufficient attention is given in the discipline to the former. Because of the radical differences in the bottom-up and top-down reasoning, their use for mutual verification of modelling results has not been sufficiently attempted. However, a rational synthesis of the available resources and modelling techniques is necessary to address the possible modelling bias and prevent false conclusions. Separating the discipline by the boundaries of familiar, convenient and/or easily interpretable approaches, methods, tools and data sources elevates the risks of the reverse. Given the importance of mediating built stock energy use through political, economic and cultural mechanisms at all levels, this study emphasises the need to focus on the incisive analysis of the subject matter by all means.

The models are shown to be helpful in stabilising building-to-grid interaction, elevating self-consumption of the renewables and self-sufficiency of the buildings in the short term; predicting the future evolution of the built stock, and the associated changes in power duration/energy costs; mapping the urban energy hotspots to prioritise for improvement; promoting and/or discouraging specific configurations of architecture-/engineering-related characteristics.

The choice between the bottom-up and top-down approach is made by considering modelling objectives, the phenomenon in question and several model appropriateness criteria. Establishing the causal effects that govern the phenomenon is crucial to each of these aspects. Explanatory modelling, as shown earlier, has the potential to infer the causality between independent and dependent variables. The implication of such a relationship empirically quantified is: i) validating the theoretical knowledge, ii) ensuring the effectiveness of energy efficiency policies and practices and iii) informing a better predictive model's design. Statistical hypothesis testing and density estimation procedures effectively aid these tasks.

A case study demonstrates that bottom-up and top-down approaches lead to fairly similar numerical results, which is a plausible conclusion given a rational model design. We, therefore, advocate mutual verification of such models, whenever possible, as a measure to i) build confidence in the model's adequacy or otherwise ii) warn about their potential inaccuracies. These aspects are essential in safeguarding the soundness of political and practical decisions based on predictive analytics. Our findings are expected to motivate numerous creative endeavours to elevate the confidence in the decisions based on

models. And to ultimately ensure the progress towards a more sustainable and inclusive built environment. It is shown that the established model classification into binary bottom-up and top-down does not accommodate some essential modelling practices and thus, needs revision.

The steps that are taken to translate a pragmatic policy- or practice-related question into a formal modelling exercise are seen as of critical importance and rely on the modeller's beliefs. Radical transparency across all modelling steps to facilitate replicability/ reproducibility/ transferability of the results is therefore emphasised. In addition to model design, this study articulates the importance of a rational design of the experiment and the downstream data transformations to ensure the consistency between the question, the model and the data. These qualities are of particular importance in further research, where the scope of modelling is expected to accommodate more technical/ architectural/ occupancy-related details, larger scales, a broader spectrum of relationships and a multitude of alternative future developments.

Bibliography

Ahmad, Tanveer, Huanxin Chen, Yabin Guo and Jiangyu Wang (2018). 'A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review'. In: *Energy and Buildings* 165, pp. 301–320. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2018.01.017.

Booth, A.T., R. Choudhary and D.J. Spiegelhalter (2012). 'Handling uncertainty in housing stock models'. In: *Building and Environment* 48, pp. 35–47. DOI: 10.1016/j.buildenv.2011.08.016.

Breiman, Leo (2001). 'Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author)'. In: *Statistical Science* 16.3, pp. 199–231. DOI: 10.1214/ss/1009213726.

Brekke, Tor, Olav Karstad Isachsen and Martin Strand (2018). *EPBD implementation in Norway. Status in December 2016*. Tech. rep. Enova, Norwegian Water Resources and Energy Directorate (NVE), Norwegian Building Authority (DIBK).

Brøgger, Morten and Kim Bjarne Wittchen (2018). 'Estimating the energy-saving potential in national building stocks – A methodology review'. In: *Renewable and Sustainable Energy Reviews* 82, pp. 1489–1496. ISSN: 1364-0321. DOI: 10.1016/j.rser.2017.05.239.

Campbell, Nina, L Ryan, V Rozite, E Lees and G Heffner (2014). 'Capturing the multiple benefits of energy efficiency'. In: *IEA: Paris, France*.

Domingos, Pedro (1999). 'The Role of Occam's Razor in Knowledge Discovery'. In: *Data Mining and Knowledge Discovery* 3.4, pp. 409–425. DOI: 10.1023/A:1009868929893.

Downey, A. (2014). *Think Stats: Exploratory Data Analysis*. O'Reilly Media. ISBN: 9781491907375.

Dubin, R. (1969). *Theory Building*. Free Press. ISBN: 9780029076309.

Economidou, M., V. Todeschi, P. Bertoldi, D. D'Agostino, P. Zangheri and L. Castellazzi (2020). 'Review of 50 years of EU energy efficiency policies for buildings'. In: *Energy and Buildings* 225, p. 110322. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2020.110322.

Fawcett, Tina and Gavin Killip (2019). 'Rethinking energy efficiency in European policy: Practitioners use of multiple benefits arguments'. In: *Journal of Cleaner Production* 210, pp. 1171–1179. ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2018.11.026.

Hamilton, Ian G., Alex J. Summerfield, Robert Lowe, Paul Ruyssevelt, Clifford A. Elwell and Tadj Oreszczyn (2013). 'Energy epidemiology: a new approach to end-use energy demand research'. In: *Building Research & Information* 41.4, pp. 482–497. DOI: 10.1080/09613218.2013.798142.

Hussain, Anwar, Muhammad Rahman and Junaid Alam Memon (2016). 'Forecasting electricity consumption in Pakistan: the way forward'. In: *Energy Policy* 90, pp. 73–80. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2015.11.028.

International Energy Agency (2020a). *Energy Technology Perspectives 2020*. Tech. rep. URL: https://www.iea.org/reports/energy-technology-perspectives-2020.

International Energy Agency (2020b). *Tracking Buildings 2020*. Tech. rep. URL: https://www.iea.org/reports/tracking-buildings-2020.

Kavgic, M., A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic and M. Djurovic-Petrovic (2010). 'A review of bottom-up building stock models for energy consumption in the residential sector'. In: *Building and Environment* 45.7, pp. 1683–1697. DOI: 10.1016/j.buildenv.2010.01.021.

Keirstead, James, Mark Jennings and Aruna Sivakumar (2012). 'A review of urban energy system models: Approaches, challenges and opportunities'. In: *Renewable and Sustainable Energy Reviews* 16.6, pp. 3847–3866. ISSN: 1364-0321. DOI: 10.1016/j.rser.2012.02.047.

Langevin, J., J.L. Reyna, S. Ebrahimigharehbaghi, N. Sandberg, P. Fennell, C. Nägeli, J. Laverge, M. Delghust, É. Mata, M. Van Hove, J. Webster, F. Federico, M. Jakob and C. Camarasa (2020). 'Developing a common approach for classifying building stock energy models'. In: *Renewable and Sustainable Energy Reviews* 133, p. 110276. ISSN: 1364-0321. DOI: 10.1016/j.rser.2020.110276.

Li, Wenliang, Yuyu Zhou, Kristen Cetin, Jiyong Eom, Yu Wang, Gang Chen and Xuesong Zhang (2017). 'Modeling urban building energy use: A review of modeling approaches and procedures'. In: *Energy* 141, pp. 2445–2457. ISSN: 0360-5442. DOI: 10.1016/j.energy.2017.11.071.

Limpens, Gauthier, Stefano Moret, Hervé Jeanmart and Francois Maréchal (2019). 'EnergyScope TD: A novel open-source model for regional energy systems'. In: *Applied Energy* 255, p. 113729. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2019.113729.

Marsaglia, George, Wai Wan Tsang and Jingbo Wang (2003). 'Evaluating Kolmogorov's Distribution'. In: *Journal of Statistical Software, Articles* 8.18, pp. 1–4. ISSN: 1548-7660. DOI: 10.18637/jss.v008.i18.

Miller, T.W. (2014). *Modeling Techniques in Predictive Analytics with Python and R: A Guide to Data Science*. FT Press Analytics. Pearson Education. ISBN: 9780133892147.

Moghadam, Sara Torabi, Chiara Delmastro, Stefano Paolo Corgnati and Patrizia Lombardi (2017). 'Urban energy planning procedure for sustainable development in the built environment: A review of available spatial approaches'. In: *Journal of Cleaner Production* 165, pp. 811–827. DOI: 10.1016/j.jclepro.2017.07.142.

Nelder, J. A. and R. Mead (Jan. 1965). 'A Simplex Method for Function Minimization'. In: *The Computer Journal* 7.4, pp. 308–313. ISSN: 0010-4620. DOI: 10.1093/comjnl/7.4.308.

Österbring, Magnus, Érika Mata, Liane Thuvander, Mikael Mangold, Filip Johnsson and Holger Wallbaum (2016). 'A differentiated description of building-stocks for a georeferenced urban bottom-up building-stock model'. In: *Energy and Buildings* 120, pp. 78–84. DOI: 10.1016/j.enbuild.2016.03.060.

Reinhart, Christoph F. and Carlos Cerezo Davila (2016). 'Urban building energy modeling – A review of a nascent field'. In: *Building and Environment* 97, pp. 196–202. DOI: 10.1016/j.buildenv.2015.12.001.

Shmueli, Galit and Otto Koppius (2009). 'The Challenge of Prediction in Information Systems Research'. In: *SSRN*. DOI: 10.2139/ssrn.1112893.

Solcast (2022). *Solar irradiance data: https://solcast.com*. (Visited on 2022).

Soto, Aner Martinez and Mark F. Jentsch (2016). 'Comparison of prediction models for determining energy demand in the residential sector of a country'. In: *Energy and Buildings* 128, pp. 38–55. DOI: 10.1016/j.enbuild.2016.06.063.

Sousa, Gustavo, Benjamin M. Jones, Parham A. Mirzaei and Darren Robinson (2017). 'A review and critique of UK housing stock energy models, modelling approaches and data sources'. In: *Energy and Buildings* 151, pp. 66–80. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2017.06.043.

Swan, Lukas G. and V. Ismet Ugursal (2009). 'Modeling of end-use energy consumption in the residential sector: A review of modeling techniques'. In: *Renewable and Sustainable Energy Reviews* 13.8, pp. 1819–1835. DOI: 10.1016/j.rser.2008.09.033.

Tian, Wei, Yeonsook Heo, Pieter [de Wilde], Zhanyong Li, Da Yan, Cheol Soo Park, Xiaohang Feng and Godfried Augenbroe (2018). 'A review of uncertainty analysis in building energy assessment'. In: *Renewable and Sustainable Energy Reviews* 93, pp. 285–301. ISSN: 1364-0321. DOI: 10.1016/j.rser.2018.05.029.

Vogel, Jonas Anund, Per Lundqvist and Jaime Arias (2015). 'Categorizing Barriers to Energy Efficiency in Buildings'. In: *Energy Procedia* 75, pp. 2839–2845. ISSN: 1876-6102. DOI: 10.1016/j.egypro.2015.07.568.

Wei, Yixuan, Xingxing Zhang, Yong Shi, Liang Xia, Song Pan, Jinshun Wu, Mengjie Han and Xiaoyun Zhao (2018). 'A review of data-driven approaches for prediction and classification of building energy consumption'. In: *Renewable and Sustainable Energy Reviews* 82, pp. 1027–1047. ISSN: 1364-0321. DOI: 10.1016/j.rser.2017.09.108.

Zhao, Hai-xiang and Frédéric Magoulès (2012). 'A review on the prediction of building energy consumption'. In: *Renewable and Sustainable Energy Reviews* 16.6, pp. 3586–3592. ISSN: 1364-0321. DOI: 10.1016/j.rser.2012.02.049.

Zhuravchak, Ruslan, Natasa Nord and Helge Brattebø (2019a). 'Control strategy for battery - supported photovoltaic systems aimed at peak load reduction'. In: *E3S Web Conf.* 111, p. 05027. DOI: 10.1051/e3sconf/201911105027.

Zhuravchak, Ruslan, Natasa Nord and Helge Brattebø (Oct. 2019b). 'Influence of emerging technologies deployment in residential built stock on electric energy

cost and grid load'. In: *IOP Conference Series: Earth and Environmental Science* 352, p. 012038. DOI: 10.1088/1755-1315/352/1/012038.

Zhuravchak, Ruslan, Natasa Nord and Helge Brattebø (2022). 'The effect of building attributes on the energy performance at a scale: an inferential analysis'. In: *Building Research & Information* 0.0, pp. 1–19. DOI: 10.1080/09613218.2022.2038537. eprint: https://doi.org/10.1080/09613218.2022.2038537.

Zhuravchak, Ruslan, Raquel Alonso Pedrero, Pedro Crespo del Granado, Natasa Nord and Helge Brattebø (2021). 'Top-down spatially-explicit probabilistic estimation of building energy performance at a scale'. In: *Energy and Buildings* 238, p. 110786. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2021.110786.

Paper I

Ruslan Zhuravchak, Natasa Nord and Helge Brattebø (2019a). 'Control strategy for battery - supported photovoltaic systems aimed at peak load reduction'. In: *E3S Web Conf.* 111, p. 05027. DOI: 10.1051/e3sconf/201911105027

Control strategy for battery-supported photovoltaic systems aimed at peak load reduction

Ruslan Zhuravchak, *Natasa* Nord and *Helge* Brattebø

Norwegian University of Science and Technology, Department of Energy and Process Engineering
Kolbjørn Hejes v 1B, NO-7491 Trondheim, Norway

Abstract. The use of photovoltaic (PV) technologies is one of the key means for achieving the balance between operational power demand and generation in net Zero Energy Buildings (nZEBs). However, direct use of PV power on-site is limited due to wide variability and uncertainty of PV output, the temporal mismatch between PV generation and load and other factors. Consequently, in addition to low self-consumption rates, the problem of peak grid load and peak PV feed into the grid persists. Batteries that are coupled to PV units may partially offer the solution to these problems, if operated under an intelligent control strategy. In this paper we proposed a forecast-based control strategy for battery-to-grid interaction aimed at enhancing self-consumption and at reducing peak load. Python programming environment was used for data processing and algorithm development. Exemplification was made based on the reported hourly energy demand in one office building of 3000 m2 heated floor area located in Trondheim, Norway. Forecasting of electricity load profiles was based on the seasonal autoregressive integral moving average (SARIMA) model. For PV power forecasting, the algorithm communicated with external service – Solcast API. The search method for optimal scheduling of operational time and the extent of charging/discharging was proposed. The results showed that as opposed to conventional battery use, this control strategy allowed to achieve significantly more consistent grid interaction. It offered highly accurate battery scheduling on a day-ahead basis while utilising minimum historical data and computational resources. The algorithm may be beneficial for end-users and grid operators, and thus, it has a high potential for future integration into building energy supply systems.

1 Introduction

Meeting strategic energy and environmental targets for nations and communities requires significant performance improvements in building sector. Net Zero Energy Buildings (nZEB) would play a crucial role in achieving these targets through low operational energy use and production of renewable energy at a quantity that meets building's delivered energy over the service lifetime [1]. Photovoltaic (PV) energy is considered as the key source that enables nZEB under various climate conditions and related political circumstances. Although grid-connected nZEBs do not require a match between instantaneous PV generation and instantaneous load, it is often preferred to increase the share of PV energy directly consumed on-site (self-consumption rate). One of the reasons for that is profitability as a result of the difference between self-produced and retail electricity prices, grid feed-in tariffs and numerous incentives [2]. Another reason is the need for more consistent grid interaction, i.e. the solutions for frequency regulations and future power grid reinforcement to handle high PV feed-in and high loads are required otherwise [3]. Maximising the building self-consumption while avoiding peak grid loads, therefore, is one of the key challenges that is focused by the research initiatives in the discipline. Two distinct but not mutually exclusive approaches can be used to tackle these challenges: demand side management (DSM) and the use of battery energy storage systems (BESS). The former is concerned with shifting the deferable loads to the time when PV system has high output. It often includes rescheduling of heating, ventilation, air conditioning and some other types of household equipment. The other approach enables energy flexibility by using PV energy stored in batteries at the time when load occurs. BESS is recognized as having a larger potential to increase self-consumption rates [4]. Currently, best practices in the field of energy flexibility involve a combination DSM and BESS [2, 5].

For the effective use of both these approaches, a reasonably accurate forecast of energy demand and PV generation in the (nearest) future is required [2]. Instantaneous load and on-site PV power, however, are the sources of variabilities and uncertainties. In addition to thermal properties of the building envelope, climate, energy supply systems, building purpose, occupant behaviour and maintenance practices determine the unique shape of load profile that changes on day-to-day basis. Solar power profile is a function of daily and seasonal variations of both direct and diffuse radiation, outdoor temperature and other factors. For a short-term load forecasting in real-time, building energy

* Corresponding author: ruslan.zhuravchak@ntnu.no

performance simulation techniques are rarely used. Detailed simulation is constrained by computational capabilities and the need for exhaustive building-related information. Data-driven approaches, on the contrary, may resolve some of these limitations [6]. Their use for electric energy demand prediction has been mainly done by regression analysis, machine learning techniques, and time series forecasting. With regression analysis, one aims at identifying a set of influential variables and the underlying (often non-linear) model structure. Machine learning techniques offer the benefits of automated pattern discovery instead. Amongst the limitations of both these approaches is the need for several explanatory variables. With time series modelling, however, the prediction of target variables can be made based only on the past observations of itself, i.e. applicable for univariate problems. Therefore, such adaptive, parsimonious time series forecasting is expected to play an important role in facilitating energy flexibility and more consistent grid interaction for nZEBs.

Likewise, real-time PV power forecasting under uncertain weather conditions is one of the fundamental challenges to enable energy flexibility. Practical solutions are offered by persistence models, physical and statistical approaches [7] that utilise sky models or data streams from either satellite imagery or reference PV systems. Recent advancements in the field succeeded at combining these into hybrid approaches which allows to achieve high forecast accuracy [8].

Given that two components, the forecasted PV output and load profiles are available, battery charge/discharge operations can be scheduled with one or more objectives, most often these are maximising self-consumption or battery service lifetime, minimising energy cost, CO_2 intensity or peak grid loads. A comprehensive overview of operational strategies for PV-coupled battery systems, their objectives and methods are available in studies [4, 9, 10].

In this study, some of the best practices related to forecast-based control were implemented and elaborated further. Analysis and modelling tasks were carried out with Python programming language using NumPy, SciPy and Sk-learn libraries for numerical computing, Statsmodels for model training/testing, Pandas for data wrangling, Matplotlib and Seaborn for data visualisation.

2 Methodology

2.1 Case study

A case study was one non-residential building located in Trondheim, Norway. It is used for health-related and recreational purposes and has the total heated floor area of more than 3000 m^2. Both electric and district heating energy use for the building were reported on hourly basis and accessible as datetime-indexed arrays through an energy monitoring platform.

Historical data over the entire building's monitoring period of more than five years is illustrated in Fig. 1. Because district heating was utilised in cold periods, electric energy use exhibits weak seasonality through the year. Also, no monthly/weekly seasonality was observed. Since such seasonality in electric energy use is negligible here, further data analysis and algorithm exemplifications are made using the most recent observation (Fig. 2) on the span of 90 days (Jul. 10^{th}, 2018 through Oct. 8^{th}, 2018, accounting for 2160 observations total).

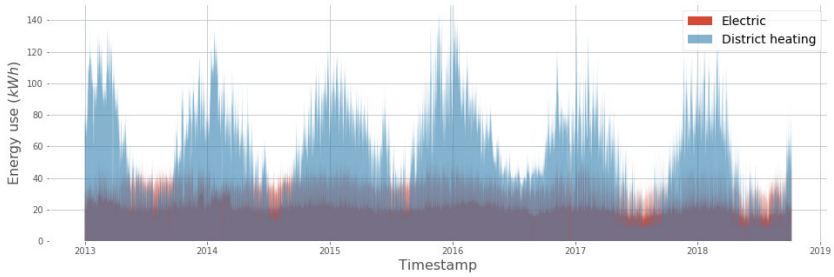


Fig. 1. Historical data (hourly measurements) on electric and district heating energy use for entire observation period.

Fig. 2. Electric energy use (hourly measurements) in most recent observations.

It can also be observed from Fig. 2 that the series has a strong daily pattern. The lowest electric energy use for the building occurs at around 6:00. It increases in linear or exponential manner through the day, achieves its peak

between 15:00 and 23:00 followed by decrease through the night.

The building is assumed to have a PV system of 150 kW total installed capacity in place with south-oriented, 30° tilted modules.

2.2 Energy use forecasting

For univariate time series forecasting, the Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model is one of the most commonly used [11, 12]. For applications with seasonal effects, ARIMA(p,d,q) has been modified to the multiplicative seasonal SARIMA$(p,d,q)\times(P,D,Q)_s$. Studies [13, 14] exemplify its application for energy forecasting purposes. Model's parameters are following:

p, P – non-seasonal and seasonal autoregressive (AR) components;

d, D – non-seasonal and seasonal order of differencing;

q, Q – non-seasonal and seasonal moving average (MA) components;

s – periodicity of the season.

Detailed formulation for both ARIMA and SARIMA is provided in literature [15]. In this study, an open-source statistical computing package Statsmodels [16] has been used for SARIMA training/testing. Package documentation contains the details on SARIMA implementation in Python.

Model's application requires the series to have constant mean, variance and autocorrelation, i.e. stationary process. This is not the case for original series, as can be seen in Fig. 2. A sequence of steps for process stationarization here involves seasonal and non-seasonal differencing. The corresponding autocorrelation and partial autocorrelation functions (ACF and PACF) are illustrated in Fig. 3.

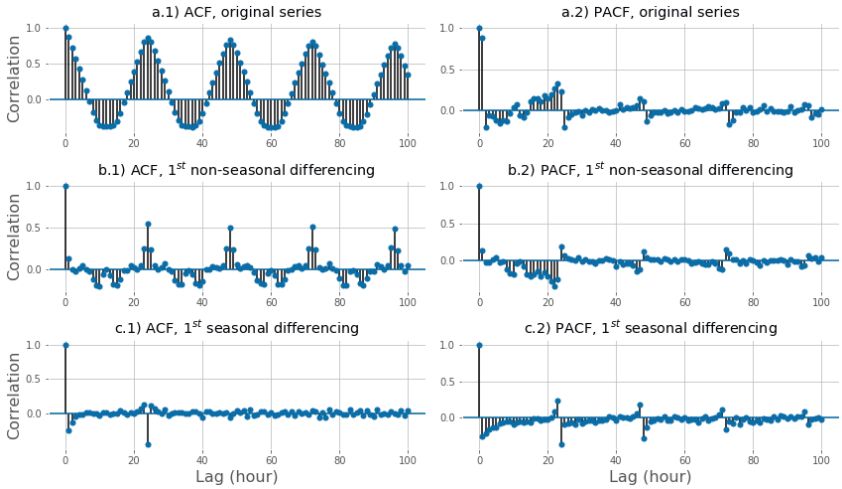


Fig. 3. ACF and PACF correlograms. a) Original data; b) Non-seasonal differencing; c) Seasonal differencing.

To induce stationarity, the first non-seasonal differencing was taken on the original series:

$$y'_t = y_t - y_{t-1} = (1-B)y_t \quad (1)$$

where:

y – original series;

B – backshift operator, such that:

$$B^k y_t = y_{t-k} \quad (2)$$

The ACF of the resulting series shown in Fig. 3b.1 still indicates the presence of statistically significant lags. Therefore, an additional seasonal differencing step was applied:

$$y''_t = (1-B^{24})y'_t \quad (3)$$

Referring to the original series, the additional seasonal differencing was obtained as:

$$y''_t = (1 - B - B^{24} + B^{25})y_t = y_t - y_{t-1} - y_{t-24} + y_{t-25} \quad (4)$$

The newly acquired time series had the mean value close to zero (0.0028) and a reasonably small number of significant autocorrelations as shown in Fig. 3c.1, at lag 1, 2, near- or at the first seasonal lag that did not repeat afterwards. This series was considered stationary, which is also confirmed by the Dickey-Fuller [17] test. The partial autocorrelation function (PACF) plot shown in Fig. 3c.2 exhibits slow decrease at the non-seasonal level and a spike at the seasonal lag. Although further differencing may result in even more consistent series, this step could induce the negative effect of overcomplication and overfitting the model. Moreover, as reported by others [18], these seasonal lags could merely reflect large correlation of non-seasonal ones.

The PACF correlogram in Fig 3c.2 indicates that the first 4 lags are significantly different from zero, thus suggesting the feasible non-seasonal component to be within the range AR ∈ [0,4]. Similarly, from the ACF plot in (Fig. 3c.1) non-seasonal MA ∈ [0,4]. Because PACF and ACF in final series indicate significant near-seasonal lag, the seasonal AR and MA chosen here are 1 and 1.

Since non-seasonal components AR (p) and MA (q) are essential for the model performance, their identification in this study is facilitated by grid search method. Grid search over the parameters' space implies training and evaluating the models with all feasible

combinations of parameters p, q. The root mean squared error (RMSE) can serve as a reference metrics for the model performance evaluation used here, defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{n=1}^{n}(y_i - \hat{y}_i)^2} \qquad (5)$$

where:

n – number of samples;

y_i – actual observation;

$\hat{y}_i$ – model prediction.

The preferred model parameters (p, q), therefore, correspond to the lowest RMSE obtained on the test set. As an additional measure used to ensure that the selection of p and q is not affected by coincidental patterns in the series, a k-fold cross-validation (CV) procedure is applied. CV for time series models implies splitting the dataset into k subsets where a training set is immediately followed by a test set. Thus, 90 most recent days of monitored data are split into $k = 5$ subsets as illustrated in Fig. 4. Grid search is conducted for each CV step. Further, the parameters (p, q) are selected based on average RMSE for all CV steps, as illustrated in Fig. 4 with color-coded heatmap.

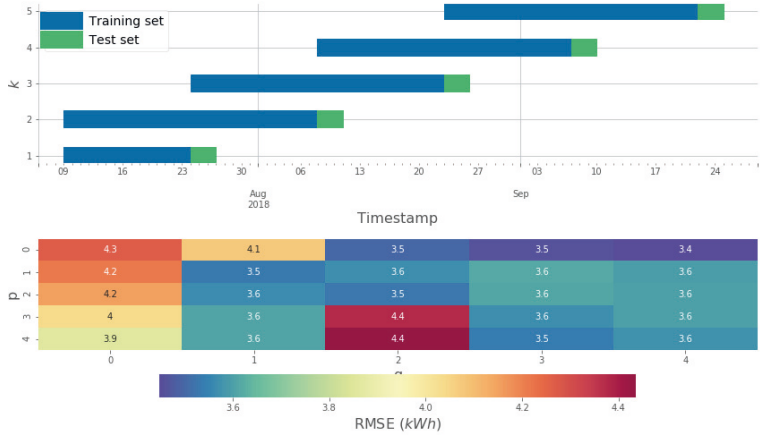


Fig. 4. Parameters tuning and validation. CV intervals split (top) and grid search results (bottom).

For a given case, the most favourable model parameters p and q are 0 and 4 accordingly, since the average over all CV steps RMSE is the least for this set of parameters. The performance of SARIMA$(0,1,4) \times (1,1,1)_{24}$ is illustrated in Fig. 5.

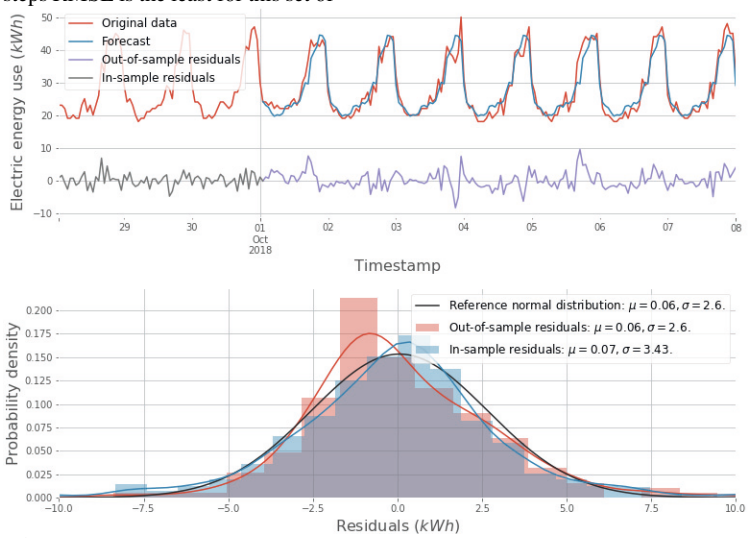


Fig. 5. Model performance. Top: original, forecasted and residuals series. Bottom: distribution of residuals.

Forecasted series overlays closely with the test series. Both, in- and out-of-sample residuals do not exceed 10 kWh. In-sample residuals follow normal distribution. Out-of-sample residuals density, however, is skewed to the left compared to the shape of equivalent normal distribution which indicates the loss of accuracy in long-term forecasting. Although such systematic error can be neglected in some applications, we restrict this control strategy to day-ahead horizon, where the accuracy is the highest.

2.3 PV power forecasting

As was mentioned before, the proposed control strategy involved communication with external service Solcast API [19] to obtain high-quality PV output forecast. Data request included geospatial (latitude and longitude) and technical information about the PV system (installed capacity, azimuth and tilt angle). The API's response is a datetime-indexed array of the forecasted PV power with 30 minutes temporal resolution over a span of 7 days. Further data wrangling involves two steps. Since the response dataset is UCT-indexed, a conversion to local time is needed. Also, original data is converted from power to energy units and from 30 minutes to hourly intervals. An example of postprocessed API's response based on PV system's parameters mentioned in Section 2.1 is illustrated in Fig. 6.

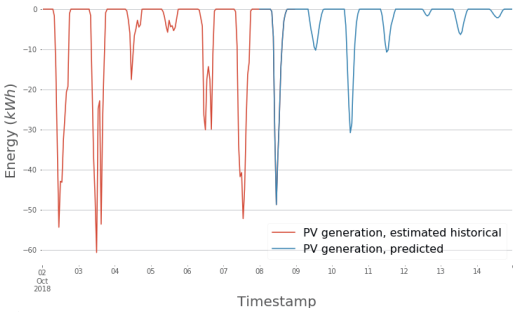


Fig. 6. PV generation (reversed), historical and forecasted.

2.4 Load matching

Given that the energy use and PV generation profiles were known, the task of load matching was to allocate the aggregated daily PV energy to the time when energy use is the highest on that day. This implies finding threshold value (p) that separates grid from battery energy use. The task could be formulated as follows:

$$\sum_1^{24} P_{PV} = \sum_{n=1}^{n}\left(\sum_{T_1}^{T_2}(P-p)\right) \pm \varepsilon \qquad (6)$$

for all $(P-p) > 0$ and if $\sum_1^{24} P_{PV} > 0$.

Here:

P_{PV} – hourly PV energy, kWh;

P– hourly energy use in the building, kWh

$p \in [0, P_{max}]$ – threshold value for the day, kWh;

$T_1, T_2 \in [0,24]$ – the beginning and the end of interval n when battery should be used;

n – number of time intervals, used to account for the possibility of battery use during more than one period through the same day;

ε – absolute error tolerance, kWh.

Solving this problem analytically would require approximating the load and PV curves and their subsequent integration. To avoid the associated loss of accuracy, a numerical search technique has been implemented instead. Namely a dichotomous search algorithm was used to find the threshold value p. Fig. 7 illustrates the search for p on Oct. 9^{th}, 2018. By gradually eliminating half of the continuous search space that was known at previous iteration, the condition (eq. 6) is achieved in a recursive manner after 8 iterations with $p = 11\ kWh$, which yields:

$$\sum_1^{24} P_{PV} = \sum_{T_1}^{T_2}(P-p) = 52.8 \pm 0.1\ kWh.$$

Finding the time intervals T_1, T_2 further, is based on the condition where $P > P - p$.

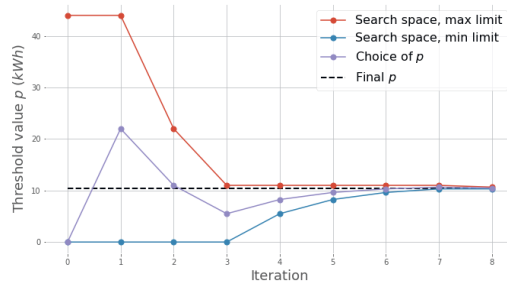


Fig. 7. Search progress.

3 Results and discussion

Given the techniques discussed above, this study proposed the algorithm for day-ahead battery operation scheduling that follows a block diagram in Fig. 8.

The first step was to send two data requests: one for historical energy use data (of a desired length, e.g. $N = 90$ days) and one for PV output forecast (installed capacity, latitude, longitude, the azimuth and the tilt of the PV modules). The response from Solcast API was then converted to local time, from power to estimated energy units and from 30 minutes to hourly time resolution. Since the day-ahead forecast was used, the series is cut at 24^{th} element. The response from the energy monitoring system's API is a datetime-indexed array of length $L = 24N$. 5 grid search and cross-validation steps were taken after which the SARIMA model parameters (p, q) were selected such that the corresponding average RMSE is the smallest. With these parameters, SARIMA$(p, 1, q) \times (1,1,1)_{24}$ was trained using the entire L-length training set. The trained model was further used for predicting energy use for the following day. Both arrays, the forecasted energy use and PV output enter the load matching unit and follow the procedure described in Section 2.4. After the load matching task was completed, the delay of 24 hours is initiated. The output of load matching unit was the reshaped grid load profile and the recommended battery charge/discharge dynamics (as exemplified in Fig. 9) to achieve such peak shaving effect.

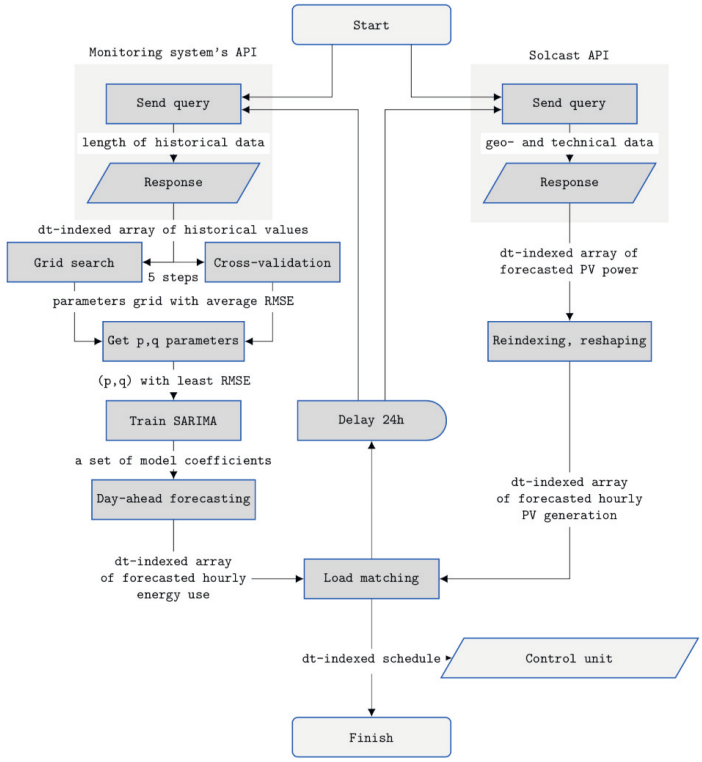


Fig. 8. Process flow diagram.

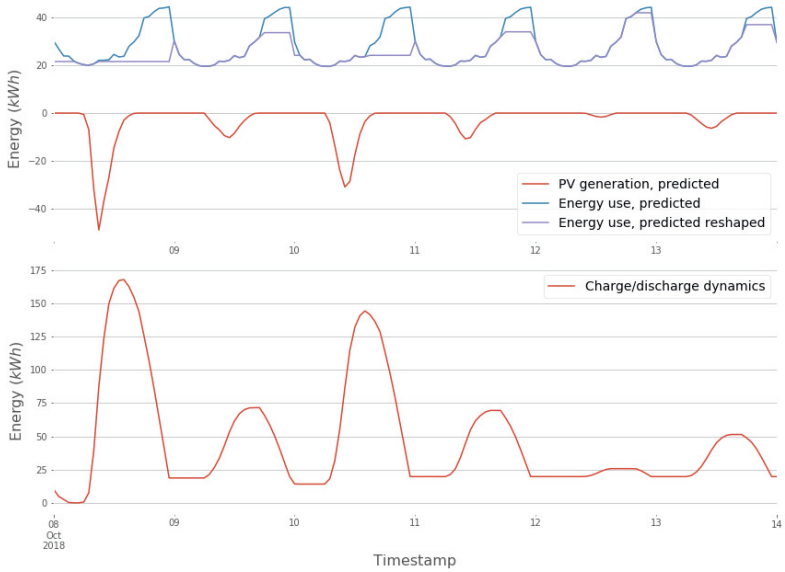


Fig. 9. Algorithm's output. Reshaped load profile (top); Charge-discharge dynamics (bottom).

Technical and climate conditions under the test case did not entail the need for controlling grid feed-in, because daily aggregated PV energy is always smaller than aggregated energy use. In case if daily PV generation was higher, a search algorithm would separate PV profile into two parts, one for self-consumption and one for grid export:

$$\sum_1^{24} P = \sum_{n=1}^{n} \left(\sum_{T_1}^{T_2} (P_{PV} - p_{PV}) \right) \pm \varepsilon \qquad (7)$$

for all $(P_{PV} - p_{PV}) > 0$ and if $\sum_1^{24} P > 0$. Here $p_{PV} \in [0, P_{PV_{max}}]$ – threshold value specific for daily PV profile, kWh. Under such circumstances, PV energy in range

$[0, p_{PV}]$ was excessive and could be exported instantaneously.

Thus, an algorithm suggests a reshaped profile, where the peak energy use and/or PV feed-in are avoided. Significantly more consistent energy use profile can be achieved for the days with high PV output, e.g. Oct. 8^{th} and Oct. 10^{th}. Over the entire period between Oct. 8^{th} and Oct. 14^{th}, 461 kWh of available PV energy offer 5% to 51% reduction of daily peak energy use from the grid.

This approach for day-ahead scheduling of battery operational time and intensity remains relatively simple and requires little historical data. It may also be used to reshape load profiles in response to dynamic energy price or CO_2 intensity.

The selection of some parameters for SARIMA was justified and automated. However, a different case study might reveal the need to conduct grid search over a larger set of parameters, e.g. seasonal components (P, Q) or to make more steps to achieve stationary series (d, D). Additional strong seasonality in building operation, e.g. weekly/holiday-based, monthly or annually, may require added complexity of SARIMA, resampling the underlying data or considering another time series model. The needed length of historical data must be satisfied accordingly. Further automation of these procedures, therefore, is needed to achieve more robust battery control that would suit the buildings of various types and more complex load profiles. For on-site algorithm deployment, as opposed to forecasting carried out in data centres, optimization of computational time and resources usage may be necessary.

Forecasting and control principles used in this study exclude the adaptation to sudden, unforeseen changes that may occur on the short term. Also, no attention is given to the potential technical limitations associated with charge/discharge intensity, critical depletion and system's losses. The authors anticipate the best application of the algorithm in combination with model predictive control or similar.

Expanding the forecasting horizon and considering a more long-term control strategy may offer even more consistent load profile. This option, however, is limited by accuracy (as can be seen in Fig. 5) and data needs in forecasting techniques currently used. Potential improvements of energy use forecasting can be achieved through including exogenous variables in time series modelling, e.g. thermal properties of building envelope, HVAC equipment specifications/setpoints and detailed data of occupant behaviour. The improvement of forecasting accuracy is also expected through supplementing time series models with other techniques, as discussed elsewhere [6].

Longer-term control strategies are likely to require more battery storage capacity. Techno-economic feasibility studies of such measures would have to address the trade-off between self-consumption rate, smoothness of grid interaction, battery capacity and utilization rate. This should be in line with the measures to prolong battery service lifetime as discussed in study [20]. As outlined in another study [4], the acceptance of such control strategies in buildings and communities highly depends on economic, technological and policy-related factors.

4 Conclusions

In order to meet energy efficiency and environmental targets without compromising grid stability, an increased self-consumption while smoothing grid interaction has to be promoted in nZEBs. This challenge requires not only capacity gains for PV systems and energy storage units, but more intelligent approaches for their utilization. The latter relies strongly on accurate energy and PV output forecasting.

With this article, an autonomous, parsimonious forecast-based control strategy was proposed. High performance of building energy use forecasting was achieved with time series model SARIMA supplemented by grid search for parameter identification and 5-fold cross validation. The proposed algorithm benefits from high quality satellite-derived PV power forecasting services. It is shown that for peak shaving purposes, the battery operation time and intensity can be scheduled on a day-ahead basis.

Such control strategies are likely to contribute to achieving the energy and environmental targets under the growing demand for efficient built environment. An increasing accessibility of PV systems and energy storage solutions are expected to positively influence their adoption.

The authors gratefully acknowledge the support from the Research Council of Norway through the research project Methods for Transparent Energy Planning of Urban Building Stocks - ExPOSe (project number 268248) under EnergiX program.

References

1. Sartori, A. Napolitano, and K. Voss, Energy and Buildings, *Net zero energy buildings: A consistent definition framework*, **48**, 220-232, (2012).
2. R. Luthander, et al., Applied Energy, *Photovoltaic self-consumption in buildings: A review*, **142**, 80-94, (2015).
3. A. Zeh, and R. Witzmann, Energy Procedia, *Operational Strategies for Battery Storage Systems in Low-voltage Distribution Grids to Limit the Feed-in Power of Roof-mounted Solar Power Systems*, **46**, 114-123, (2014).
4. G. Angenendt, et al., Applied Energy, *Comparison of different operation strategies for PV battery home storage systems including forecast-based operation strategies,* **229**, 884-899, (2018).
5. M. Castillo-Cagigal, et al., Solar Energy, *PV self-consumption optimization with storage and Active DSM for the residential sector,* **85(9)**, 2338-2348, (2011).
6. C. Deb, et al., Renewable and Sustainable Energy Reviews, *A review on time series forecasting techniques for building energy consumption*, **74**, 902-924, (2017).

7. M.Q. Raza, M. Nadarajah, and C. Ekanayake, Solar Energy*, On recent advances in PV output power forecast*, **136,** 125-144, **(**2016).
8. J.M. Bright, et al., Solar Energy, *Improved satellite-derived PV power nowcasting using real-time power data from reference PV systems*, **168**, 118-139, (2018).
9. L. Schibuola, M. Scarpa, and C. Tambani, Renewable Energy, *Influence of charge control strategies on electricity import/export in battery-supported photovoltaic systems*, **113**, 312-328, (2017).
10. J. Moshövel, et al., Applied Energy, *Analysis of the maximal possible grid relief from PV-peak-power impacts by using storage systems for increased self-consumption*, **137**, 567-575, (2015).
11. D.W. van der Meer, J. Widén, and J. Munkhammar, Renewable and Sustainable Energy Reviews, *Review on probabilistic forecasting of photovoltaic power production and electricity consumption,* **81**, 1484-1512, (2018).
12. H. Verdejo, et al., Renewable and Sustainable Energy Reviews, *Statistic linear parametric techniques for residential electric energy demand forecasting. A review and an implementation to Chile* **74**, 512-521, (2017).
13. N. Elamin, and M. Fukushige, Energy, *Modeling and forecasting hourly electricity demand by SARIMAX with interactions*, **165**, 257-268, (2018).
14. A. Tarsitano, and I.L. Amerise, Energy, *Short-term load forecasting using a two-stage sarimax model*, **133**, 108-114, (2017).
15. G.E.P. Box, et al., *Time Series Analysis: Forecasting and Control*. (Wiley, 2015).
16. S. Seabold, and J. Perktold. *Statsmodels: Econometric and Statistical Modeling with Python*, (9th Python in Science Conference, USA, 2010).
17. R.H. Shumway, and D.S. Stoffer, *Time Series Analysis and Its Applications: With R Examples*, (Springer International Publishing, 2017).
18. R.E. Abdel-Aal, and A.Z. Al-Garni, Energy, *Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis*, **22(11)**, 1059-1069, (1997).
19. Solcast, *Solar irradiance data*, http://solcast.com.au/, (2018).
20. J. Li, and M.A. Danzer, Journal of Power Sources, *Optimal charge control strategies for stationary photovoltaic battery systems*, **258**, 365-373, (2014).

Paper II

Ruslan Zhuravchak, Natasa Nord and Helge Brattebø (Oct. 2019b). 'Influence of emerging technologies deployment in residential built stock on electric energy cost and grid load'. In: *IOP Conference Series: Earth and Environmental Science* 352, p. 012038. DOI: 10.1088/1755-1315/352/1/012038

IOP Conf. Series: Earth and Environmental Science **352** (2019) 012038 doi:10.1088/1755-1315/352/1/012038

Influence of emerging technologies deployment in residential built stock on electric energy cost and grid load

Ruslan Zhuravchak, Natasa Nord and Helge Brattebø

Norwegian University of Science and Technology
Kolbjørn Hejes vei 1B, Trondheim, NO-7491, Norway

E-mail: ruslan.zhuravchak@ntnu.no

Abstract. High penetration rates of novel building energy technologies has prompted a growing concern about their microeconomic effect and grid influence. Deployment of photovoltaic (PV), solar water heating (SWH) systems and energy storage solutions, in addition to the growth of electric vehicles (EV) fleet, are reshaping the structure of built stock and lead to the changes in its electric energy demand profile. Long-term forecasting of such structural changes is necessary to guide the decision-making process that would satisfy the needs of both, energy consumers and the suppliers. Whereas electric energy price model is one of the key influencing factors of technologies acceptance for households, peak loads and grid feed-in determine the needed capacity of power grids. The objective of this study was to assess both, the aggregated cost of energy and the changes in cumulative load profiles that are excepted by 2050 for one of residential building typologies in Norway. Methodologically, it was achieved with descriptive statistics, stochastic forecasting and detailed energy performance simulation. Annual electric energy cost for consumers were evaluated under six pricing models. The results suggested that time-of-use and variable maximum power extraction models represent the lowest and the highest extremes in energy cost. At the aggregated level, peak load will decrease in range 1% to 13% compared to current level. Peak PV feed-in will reach up to 40% of peak load by 2050.

1. Introduction

Renewable energy technologies and energy conservation measures together with the emerging electric mobility solutions are penetrating the residential built stock rapidly. The motives and the consequences of such shift have to be studied in a systemic manner [1] to ensure the compliance to long-term strategic development plans for communities.

The benefits for consumers typically involve achieving the economic [2] and possibly environmental objectives [3] that may result from the reduced energy use and from power feed-in on the long term. More detailed analysis, however, suggests that these technologies may cause even further inconsistency in grid interaction [4]. Hence, their deployment at a large scale may require added power generation capacity and grid reinforcement. For the Norwegian power system this issue becomes even more challenging given that electricity is the major source of energy for space heating. Thus, Norwegian power grid development and maintenance for the

period 2016 - 2025 will be supported by NOK 140[1] billion of investments [5].

Under such investment strategies, the expected increase in electric energy nominal price for households is 30% by 2025. In addition, a number of changes in electric energy pricing model will be introduced [5] by the Norwegian Water Resources and Energy Directorate (NVE). This brings to the necessity to evaluate the future costs of electric energy under various pricing methods.

Forecasting the future state of built stock is crucial for developing further power grid investment plans and energy pricing methods. Realistic long-term, large-scale energy forecasting models need to account for: 1) Detailed information on how certain technologies affect the energy performance at a single building level; 2) Typological complexity, variability, and dynamic nature of built stock; and 3)Uncertainty in consumers' decision to accept a certain novel technology.

"White-box" energy performance simulation is considered as an effective tool used to evaluate the results of technological interventions applied at a building level. Large-scale energy planning, however, is hindered by additional challenges, associated with the heterogeneous structure of building stock. Figure 1 illustrates the variability of energy use for residential buildings in Norway. A number of factors lead to significant variance, even within the same building type [6]. It can be attributed to the properties of envelope, energy supply system, appliances and occupant behaviour. Appropriate statistical methods that quantify this variability should be used at the model development and validation steps to minimise the error [7].

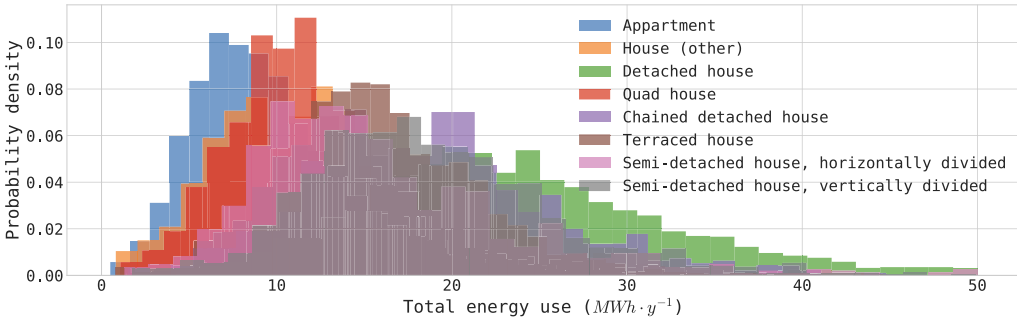


Figure 1. Energy use in the Norwegian residential built stock - univariate distribution

The deployment of novel technologies in buildings has a large number of underlying sources of variability that cannot be accommodated by deterministic modelling principles. Social and economic environment, amongst the other factors that affect the decision-makers, can be considered as random variables. Studying the dynamic evolution of such random phenomenas has to involve stochastic forecasting with the appropriate tools from probability theory [8]. This enables to account for a full spectrum of likely future outcomes, to quantify the uncertainties and thus, serves as a basis for well-informed strategic planning. Consequently, stock-wide influence of novel technologies on future electric energy prices and on power grid needs to be estimated in a quantitative, probabilistic terms. Developing the feasible methodological approach to achieve such objective is the key reasoning behind this study. It is exemplified with one residential building typology - semi-detached house divided vertically. The approach involves typological sampling, detailed building energy performance simulation, and stochastic forecasting. For the exemplification purposes, four technologies were considered: EV, PV, domestic hot water tank (DHWT) and SWH system.

[1] As of May 2019, 1 NOK = 0.1018 EUR.

2. Methodology

The suggested methodology consisted of five steps:

- Descriptive statistics - estimating central tendency, dispersion and distribution of variables within the typology to establish key properties of the representative building;
- Detailed energy performance simulation of the representative building without and with the technologies under the study and all possible combinations of those;
- Statistical simulation of stock-wide acceptance of technologies;
- Aggregation of the resulting cumulative energy costs under the available pricing methods;
- Aggregation of grid load and grid feed-in.

For detailed energy performance simulation, IDA-ICE software was used. Analytical tasks were carried out in Python programming language using the following libraries: 1) NumPy, SciPy, Statsmodels, and Networkx for numerical computing and simulation; 2) Pandas for data wrangling; 3) Matplotlib and Seaborn for data visualisation.

2.1. Descriptive statistics

The analysis of Norwegian Energy Performance Certificates (EPC) was carried out at this step. The background related to the dataset, as a component of EPBD [9] implementation in Norway, are available in source [10]. Amongst 18100 records for all the residential building categories listed in Figure 1 except apartments, semi-detached house divided vertically reached 9.7% by records count, 9.0% by heated floor area and 9.3% by energy use. The distribution of samples based on heated floor area (m^2) is illustrated in Figure 2.

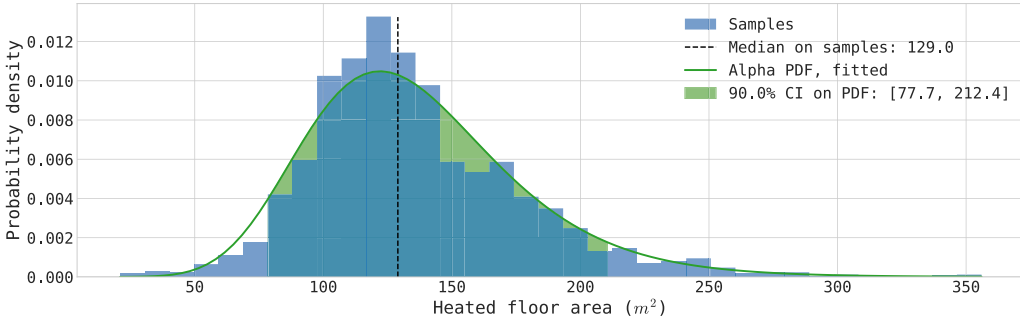


Figure 2. Heated floor area - univariate distribution (data-specific and theoretical best fit)

It was found that Alpha Probability Density Function (PDF), compared to other theoretical PDFs, describes this variable in the dataset best. For this type of PDF with its specific parameters, 10^{th} and 90^{th} percentiles, or confidence interval (CI) is [77.7, 212.4]. Subsequent steps in the methodology focused on this range only.

Further analysis of the EPC dataset revealed that 60% of buildings within this typology were constructed before 1990, as concluded earlier in source [11]. An additional step was aimed at describing the energy intensity within the building typology. Figure 3 illustrates the linear relationship between the age of buildings and the energy intensity. The figure indicates that the energy performance of buildings constructed before 1990 remains relatively poor. They contribute significantly to the stock-level energy use and their refurbishment should be of priority. The overall linear trend was further used for model validation, as elaborated in section 2.2.

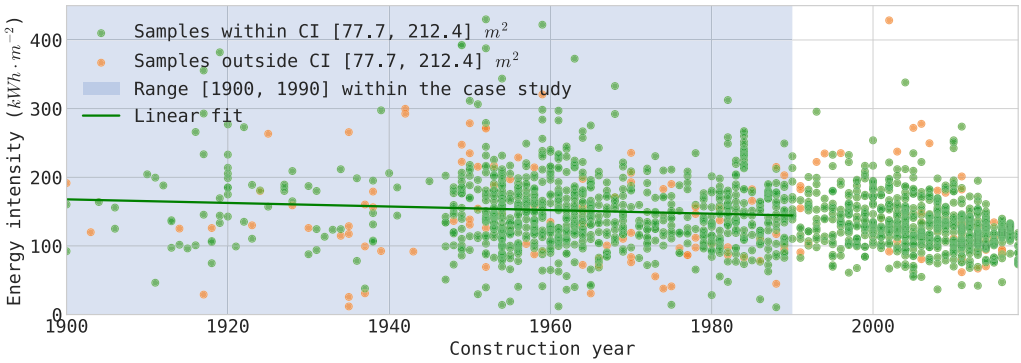


Figure 3. The relationship between building's age and energy intensity

2.2. Energy performance simulation

Based on the information from previous section, a selected representative building was two-storey, three-bedroom, single-family house built under the standards available before 1990. Heated floor area of the building modelled (122.2 m^2) close to the medium value for this typology (129.0 m^2 as specified in Figure 2). The unit was located at the end of a terraced building to represent the worst-case scenario. It was assumed that electricity is the only source of energy for space heating and domestic hot water (DHW) supply. A multi-zone model of the building was developed in IDA-ICE [12] with climate data for Bergen, Norway. Based on NS3031 [13], the heating system was designed for a desired temperature of 21/19 $°C$ in occupied/non-occupied hours during the coldest period of the year. The internal gains, electrical appliances, and DHW use were modelled in detail to take into consideration occupant behaviour as shown in [14].

The outputs of energy performance simulation for this (seed) model are load and energy use profiles on hourly basis over one year of operation. Simulated annual specific energy use was compared against standards (178 kWh/m^2 [15]), other source [16] and the EPC dataset (Figure 3) as a model calibration step.

A calibrated seed model was further extended with all unique combinations of four technologies: EV, PV, DHWS and SWH (Table 1).

Electric load profiles for EV defined based on the data for the most common commercial solutions [17]. Model inputs related to technical specifications for PV systems derived from manufacturers and distributors [18, 19]. A list of unique combinations of technologies in Table 1 accounted for the technical relationship between SWH system and the DHW tank - DHWT with the demand management mechanism is an independent component whereas SWH does requires DHWT for its operation.

The annual load and building energy use profiles were simulated for all the combinations of technologies listed in Table 1. Load profiles over the three coldest days are illustrated in Figure 4 for the simplest (seed) and the most complex models. As a convention, a reversed PV power profile displayed here. This subset of the load profiles, when the demand for space heating is highest, demonstrates the key differences between the models:

- PV power, if self-consumed, enables to decrease grid load substantially. The mismatch between PV generation and peak load, however, makes such benefit insignificant for grid stability;
- SWH system in combination with DHW tank are instrumental in reducing the total load

Table 1. Unique combinations of technologies

Index	EV	PV	DHWT	SWH	Comment
1					Seed model (none of the technologies accepted).
2	■				Battery capacity 22 kWh. 50% charging daily.
3		■			PV area 16 m^2. 4000 kWp installed capacity.
4			■		Volume 200 l. Heater capacity 2 kW.
5			■	■	Collector area 16 m^2. 200 l storage tank.
6	■	■			
7	■		■		
8		■	■		
9	■	■	■		
10	■		■	■	
11		■	■	■	
12	■	■	■	■	All technologies accepted.

during the peak periods. Their performance varies strongly in response weather conditions;

- EV may contribute substantially to the afternoon peak if no advanced charging control/scheduling used.

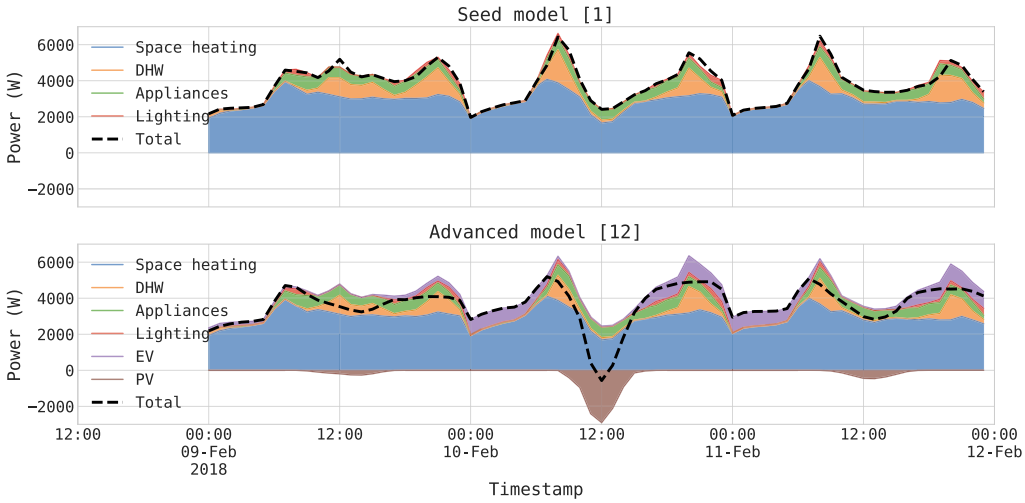


Figure 4. Load/feed-in profiles for model [1] (top) and [12] (bottom)

2.3. Stochastic model

This section is concerned with developing a stochastic model used to predict a stock-wide penetration level of novel technologies. The key concept is stochastic process of building's evolution. It defines the sequence and the time steps with which the technologies can be accepted in one building. The acceptance of any technology in a building is denoted by a discrete (binary) random variable X:

$$X = \begin{cases} 0, & \text{if technology is } \textbf{not accepted} \\ 1, & \text{if technology is } \textbf{accepted} \end{cases} \tag{1}$$

Variable X is characterised by the synthetic parameter - technological acceptance rate (TAR). TAR reflects the probability P with which any of the available technologies will be accepted $(X = 1)$ at the time step t, such that:

$$P\{X = 0\} + P\{X = 1\} = 1 \tag{2}$$

Under the absence of perfect knowledge, TAR is meant to encapsulate all kinds of judgments and influencing factors that drive the decision to accept any technology. Through social advertisement, economic incentives and other energy-related programs at the municipal or national scale, this parameter can be influenced.

Given that at some time step t, variable X yields an acceptance, the exact technology is drawn from a random, uniformly distributed variable denoted as U which is one realization of an exhaustive list of the available technologies $u = \{EV, PV, DHWT, SWH\}$. Thus, when the first acceptance occurs:

$$P\{U = EV\} = P\{U = PV\} = P\{U = DHWT\} = P\{U = SWH\} = \frac{1}{4} \tag{3}$$

A general form of Equation 3 for the acceptance n is:

$$\forall n \in [1,\ 4], \quad \forall U \in u \qquad P\{U = u\} = \frac{1}{5 - n} \tag{4}$$

The overall process resembles a random walk over the discrete filed with inconsistent time step. Figure 5 illustrates one sample path with TAR = 0.07 as a network graph. Current state of the building always corresponds to the seed model. The final state, however, depends on if/how many acceptances occurred and which technologies were drawn, but always has a reference model in Table 1.

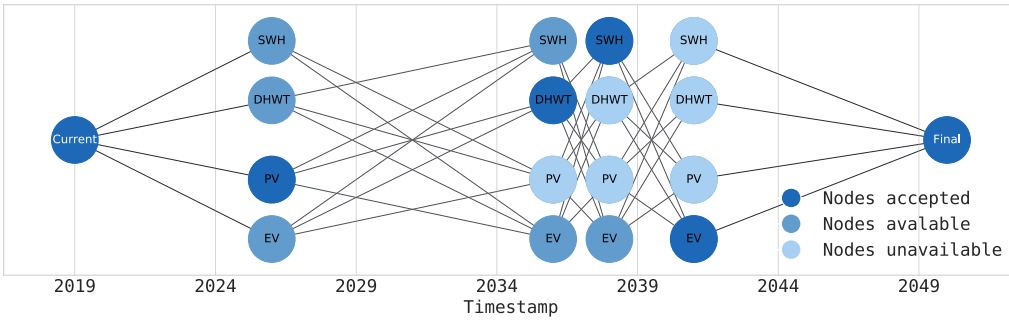


Figure 5. Sample path - acceptance of technologies in the building

Such simulations were carried out for each out of 1000 individual buildings within the synthetic built stock. Its final state can be presented as a bar chart in Figure 6, illustrating a number of times when each unique combination of technologies can be observed in 2050 under TAR = 0.07. The results in Figure 6 represent one possible outcome (trial) of the built stock evolution. To address the objectives of the study, 1000 trials were simulated, where TAR takes

a random value from continuous normal (Gaussian) PDF with mean value $\mu = 0.05$ and the standard deviation $\sigma = 0.015$.

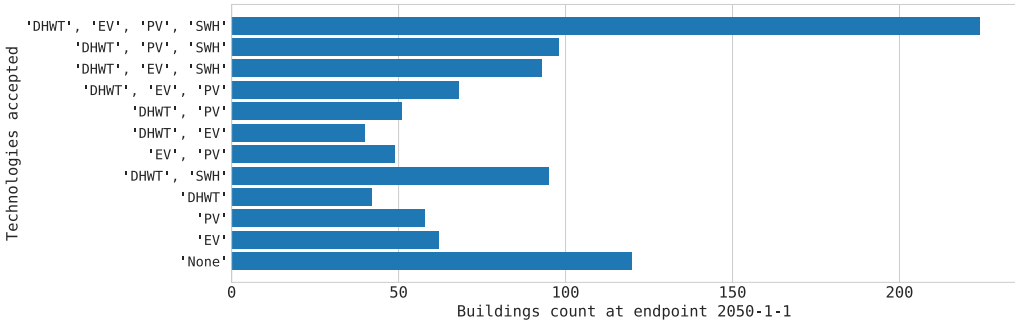


Figure 6. Number of buildings per combination of technologies accepted

Given that simulated final states of built stock are available, the corresponding data on grid interaction from the reference models and the energy price can be aggregated to the stock level.

2.4. Energy price calculation methods

Currently in Norway, the electricity bill for residential building owners consists of two elements: energy and electricity grid fee. Further, the electricity grid fee consists of energy part and a constant part. This constant part should reflect the power demand of a user, but in the current grid fee model it is not considered. Finally, this means that the electricity grid fee is not reflecting the real electricity grid cost caused by the power demand of residential buildings.

In this study, the energy cost was based on the Nord Pool hourly data for 2017. The energy cost was calculated by using hourly electricity demand and the hourly electricity cost from the Nord Pool market. The grid fee part may be defined in different ways depending on the model. In this study, the six models were analysed: 1) the current spot price model, 2) the current period model, 3) the maximum measured power model with the constant coefficients over the year, 4) the maximum measured power model with the variable coefficients over the year, 5) the subscribed power model with 4 kW subscribed power, and 6) the time of use model. The specific values for each of these models were provided from Haugaland Kraft. In the case of the electricity export, the feed-in tariff, when the building installed PV, a possible income for the building was calculated. This income was calculated in the same way as the grid fee cost for each model, except that each element was weighted 80% and 25% was taken for the taxation. An incentive to motivate building owners to install PVs would be to decrease the taxation, but this was not analysed in this study. For each of the scenarios in Table 1, monthly and annual electricity cost for each of the six pricing models were calculated. Due to effectiveness of the paper, the pricing models are not introduced in detail mathematically. Currently, in Norway there are some suggestions for the feed-in tariff, but it is not yet widely used. The values used in this paper were discussed with the company Haugaland Kraft.

3. Results

Given the simulation procedures and assessment methods discussed above, the results represent the likely levels of penetration of technologies under the study. The associated grid interaction and the prices for energy are elaborated here.

3.1. Aggregated grid interaction

In Figure 7 illustrates 1000 load duration curves aggregated to the synthetic stock level. Each curve reflects an evolved unique stock, composed of buildings with various combination of technologies (Table 1) in 2050. For comparative purposes, aggregated duration curve from seed model is displayed also.

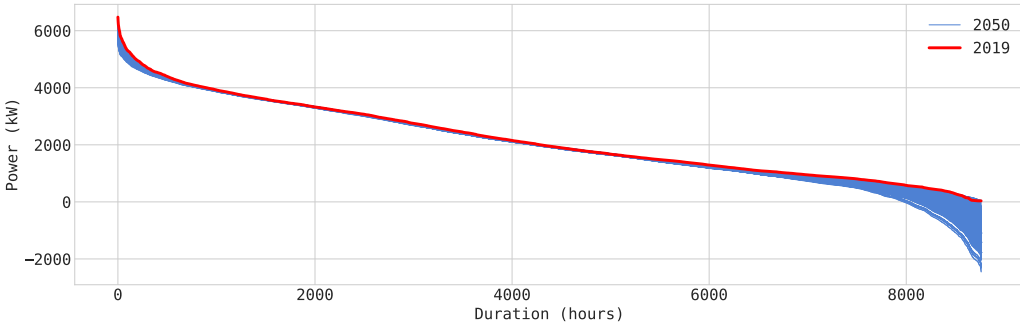


Figure 7. Aggregated load duration curves

It can be observed that the middle part of the curves remains unchanged. Substantial changes, however, are expected in peak loads and at grid feed in. Peak load observed on the annual basis is expected to decrease between 1% and 13%. The running time with low loads will change, which can be attributed to higher levels of self-sufficiency. The figure indicates that under this case study, extra peak power generation will not be necessary since additional load caused by the the EVs will be compensated by PV and SWH systems. Penetration of PV systems would yield significant levels of peak grid feed-in which can reach up to 40% of peak load.

3.2. Cumulative electricity cost

Cumulative annual cost for electricity, according to six pricing methods discussed in Section 2.4 and under various TAR, are illustrated in Figure 8.

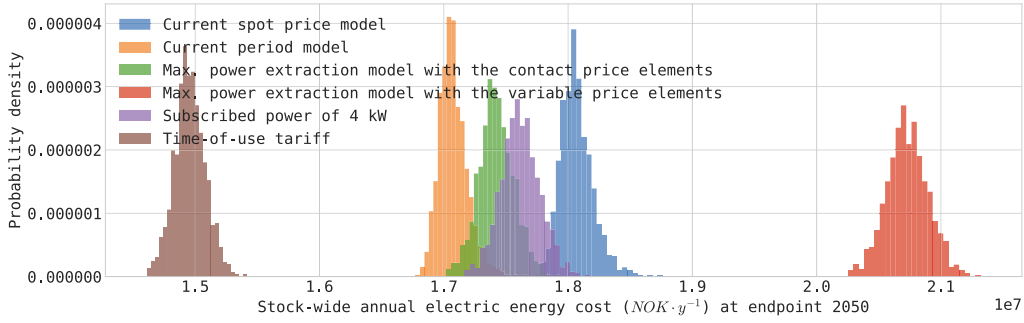


Figure 8. Aggregated electric energy costs per pricing model

It is evident from the Figure 8, that time-of-use tariff and variable maximum power extraction models refer to the lowest and the highest electric energy price under any acceptance of novel

technologies. Overlapping variance makes it more challenging to establish preferences between the other four models. Thus, under high TAR (left part of each histogram), cost for energy with the spot price model will be lower compared to subscribed power (or contract maximum power) model with low acceptance of novel technologies.

4. Discussion

A methodological procedure, proposed in this study, accounts detailed information at the building level and random causality behind the acceptance of novel technologies. Modelling principles (and the results) are not idealised by deterministic scenario modelling, but benefit from probabilistic approaches instead. The likely developments of grid interaction and energy price are evaluated, providing the necessary background for informed energy planning.

As elaborated in Section 2.1, a dynamic stochastic process relies strongly on synthetic TAR. More detailed analysis of consumers' willingness to deploy [20] each particular technology, their economic, environmental and social motives, is needed to produce more accurate results.

The dataset and the pricing methods used to assess the stock-wide cost for energy are limited by those currently available. The results are sensitive to any future changes in e.g. taxation mechanisms, incentives for energy efficiency measures and/or low energy use. With methodological and instrumental toolset used in this study, the feasibility of such strategic initiatives can be assessed.

5. Conclusions

Meeting future strategic energy- and environment-related plans for cities and communities requires a consideration of structural changes in the built environment. These changes are, to the large extent, shaped by the penetration of novel technologies, particularly in the residential buildings. This study evaluated two key aspects of such changes in the future grid stability and cost for energy. The scope covered EV, PV, DHWT, and SWH systems deployment. These objectives were achieved through a comprehensive methodological approach that involves descriptive statistical analysis, building energy performance simulation with model calibration step and stochastic forecasting of built stock evolution. It is exemplified with one residential building typology in Norway. The results suggested that for the given case study, additional power demand is not likely to occur, regardless of increased use of EVs. PV feed-in, however, would reach substantial levels (up to 40% of peak load), depending on the penetration rates for the PV systems. Considering different pricing methods, those based on the pricing of power extraction would cover better electricity grid cost, while they would result in highest electricity cost for the building users. The limitations of the study are associated with narrow scope and data scarcity. However, it was shown that the methodology is applicable for the analysis of future developments of built stock and answering the given research questions. Large scale energy planning and policy making, therefore, may benefit from the approaches and methods provided in study.

Acknowledgments

The authors gratefully acknowledge the support from the Research Council of Norway through the research project Methods for Transparent Energy Planning of Urban Building Stocks - ExPOSe (project number 268248) under EnergiX program.

References

[1] Nykamp H 2017 *Environmental Innovation and Societal Transitions* **24** 83 – 93
[2] Becchio C, Bottero M C, Corgnati S P and Dell'Anna F 2018 *Land Use Policy* **78** 803 – 817
[3] Lausselet C, Borgnes V and Brattebø H 2019 *Building and Environment* **149** 379 – 389

[4] Sørnes K, Fredriksen E, Tunheim K and Sartori I 2017 *Energy Procedia* **132** 610 – 615 11th Nordic Symposium on Building Physics, NSB2017, 11-14 June 2017, Trondheim, Norway
[5] Hansen H, Jonassen T, Løchen K and Mook V 2017 Forslag til endring i forskrift om kontroll av nettvirksomhet - Utforming av uttakstariffer i distribusjonsnettet Tech. Rep. 5 Norges vassdrags- og energidirektorat
[6] Yoshino H, Hong T and Nord N 2017 *Energy and Buildings* **152** 124 – 136
[7] Fonseca J and Panão M J O 2017 *Energy and Buildings* **152** 503 – 515
[8] Moral P and Penev S 2017 *Stochastic Processes: From Applications to Theory* Chapman & Hall/CRC Texts in Statistical Science (CRC Press)
[9] Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency Official Journal of the European Union
[10] Brekke T, Isachsen O K and Strand M 2018 EPBD implementation in Norway. Status in December 2016 Tech. rep. Enova, Norwegian Water Resources and Energy Directorate (NVE) and Norwegian Building Authority (DIBK)
[11] Statistisk sentralbyrå (SSB) / Statistics Norway 2013 Statistisk årbok 2013
[12] EQUA Simulation AB 2019 IDA indoor climate and energy - a new generation building performance simulation software
[13] Standard Norge 2014 NS-3031: Beregning av bygningers energiytelse - Metode og data
[14] Nord N, Tereshchenko T, Qvistgaard L H and Tryggestad I S 2018 *Energy and Buildings* **159** 75 – 88
[15] Statistisk sentralbyrå (SSB) / Statistics Norway 2011 Energy consumption in households, 2009
[16] Enova SF 2017 Enovas byggstatistikk 2017
[17] Sørensen Å L, Jiang S, Torsæter B N and Völler S 2018 Smart ev charging systems for zero emission neighbourhoods - a state-of-the-art study for norway. zen report no. 5 Tech. rep. SINTEF and NTNU
[18] Norsk solenergiforening (NSF) 2019 Produser din egen strøm med solcellepanel
[19] FME SUSOLTECH-IFE, BIPVNO-SINTEF, BIPVNO-NTNU, Solenergiklyngen and Solenergiforeningen 2018 Muligheter og utfordringer knyttet til bygningsintegrerte solceller (BIPV) i norge 2018
[20] Zheng D, Yu L, Wang L and Tao J 2019 *Sustainable Cities and Society* **44** 291 – 309

Paper III

Ruslan Zhuravchak, Raquel Alonso Pedrero, Pedro Crespo del Granado, Natasa Nord and Helge Brattebø (2021). 'Top-down spatially-explicit probabilistic estimation of building energy performance at a scale'. In: *Energy and Buildings* 238, p. 110786. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2021.110786

Energy & Buildings 238 (2021) 110786

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enb

Top-down spatially-explicit probabilistic estimation of building energy performance at a scale

Ruslan Zhuravchak [a,*], Raquel Alonso Pedrero [b], Pedro Crespo del Granado [b], Natasa Nord [a], Helge Brattebø [a]

[a] Department of Energy and Process Engineering, NTNU. Kolbjørn Hejes v. 1B, NO-7491 Trondheim, Norway
[b] Department of Industrial Economics and Technology Management, NTNU. Alfred Getz v. 3, NO-7491 Trondheim, Norway

ARTICLE INFO

Article history:
Received 7 August 2020
Revised 14 December 2020
Accepted 23 January 2021
Available online 4 February 2021

Keywords:
Built stock
Energy performance
Geographical information systems (GIS)
Urban energy planning
Top-down modelling
Probabilistic sampling
Parametric density estimation
Frequentist inference

ABSTRACT

Achieving the energy-related and environmental targets for nations and municipalities is largely dependent on the existing built stock. It plays a pivotal role in the accomplishment of these targets through the implementation of energy efficiency and flexibility programs, involving the deployment of distributed energy resource management technologies, refurbishment of building envelopes and upgrading of indoor environmental control equipment. Spatial awareness about urban energy use enables to prioritise the areas where these solutions will be most effective and balanced with the plans for new constructions. Large-scale building energy mapping, however, must cope with heterogeneity of buildings within the built stock, absence of detailed information and multiple sources of uncertainty that stem from the complex and dynamic properties of the phenomenon at a building level. One of the key challenges in the discipline is to account for these uncertainties while maintaining the rational model complexities and data needs. This study, therefore, suggests a parsimonious top-down probabilistic modelling recipe to enable geospatial energy mapping and analysis. Under such modelling principles, an inverse propagation of uncertainties is carried out from the status quo of the built stock. The proposed framework is based on probabilistic sampling with prior parametric univariate density estimation and statistical hypothesis testing. Consolidation with the exogenous influencing factors is facilitated through the measure of statistically significant difference. This approach is exemplified with the data from two sources: the cadastral system and the energy performance certificates registry. A case study developed for Trondheim (Norway) quantified the central tendency and dispersion in the distributions of the simulated bulk total annual energy use by buildings per 1×1 *km* grid cell over the urban territory. The results suggest that best estimates of these values vary between 11 $MWh \cdot y^{-1}$ and 141 $GWh \cdot y^{-1}$ depending on the grid cell. A measure of dispersion in the simulated results is highly correlated with these estimates. Robust handling of uncertainties and the possibility to accommodate a variety of modelling objectives make this approach practical for energy mapping with a flexible spatial resolution that may facilitate numerous applications in energy planning. A collection of methods for univariate density estimation discussed in this study together with the empirical data are accessible through Built Stock Explorer:https://builtstockexplorer.indecol.ntnu.no. This open web application for knowledge discovery in building energy data enables to reproduce some of the results presented in the article.

1. Introduction

Built stock is perceived as holding a large potential for mitigating the environmental impacts directly or indirectly associated with its final energy use [1], which reached 128 *EJ* globally in 2019 [2]. Improving the energy performance of buildings, therefore, is being supported through regulatory mechanisms at various levels of governance. These mechanisms, usually initiated at a national or municipality levels, are targeting the solutions at districts or neighbourhoods [3] and focused primarily on well-reasoned infrastructural transformations [4], retrofitting and upgrading programs [5] and more sustainable energy management technologies [6].

An effective strategic energy planning of these and the related solutions relies on geospatial information in several ways. Spatial awareness enables to priortise the areas of high energy use, where the technical, economic, and environmental feasibility of relevant measures may be justified. Such solutions could lead to multiple benefits, e.g. decrease the total energy use and reduce the costs

* Corresponding author.
E-mail address: ruslan.zhuravchak@ntnu.no (R. Zhuravchak).

https://doi.org/10.1016/j.enbuild.2021.110786
0378-7788/

Nomenclature

CDF	Cumulative distribution function, page 12	r.v.	Random variable, page 7
i.i.d.	Independent identically distributed (sample), page 24	SD	Standard deviation, page 9
KS	Kolmogorov–Smirnov (test), page 14	SS	Sample size, page 14
MC	Monte-Carlo (simulation), page 7		
MLE	Maximum likelihood estimation, page 13		
PDF	Probability density function, page 5		

of energy by avoiding transmission losses, support the integration of non-dispatchable renewables, enhance the reliability and resilience of power grid through peak load smoothing and frequency control. Given the anticipated growth of the electric vehicles fleet, a more favourable deployment of charging stations can also be accommodated by spatially-informed energy planning [7]. Another important reason for mapping the energy bottlenecks is studying the contribution of buildings to the atmospheric heating – a phenomenon known as urban heat island [8,9].

The need for spatial analysis of urban energy use prompted numerous attempts to complement built stock energy information with geospatial data [10,11]. The availability of multiple studies indicates a widespread interest in developing the means to enable such analysis. Building level [12,13], block area [14], square grid cell [8,15,9] and local authority [16] are the most common choices of spatial or administrative resolution that such analysis is focused on.

Energy performance of buildings, however, is characterised as highly complex, complicated and dynamic phenomena, which is attributed to numerous factors of physical and occupancy-related origins [17]. At a larger scale, the associated uncertainties are amplified by the diversity of buildings, variations in their exposure to the outdoor conditions, ageing processes, use/maintenance practices and other. Determining the energy performance of buildings at the large scale given these explanatory factors is the subject of built stock energy modelling [18]. It is, amongst other important applications, an essential component of geospatial energy mapping. Following both, the original classification system proposed by Swan and Ugursal [19] and its recent revision [20], the prevailing practices for spatially-explicit built stock energy modelling correspond to the bottom-up approach. Bottom-up reasoning enables to infer energy use at the aggregated level based on the information available at a lower spatial level. Several studies suggested engineering-based (“white-box”) models as the means for building energy mapping [15,9,12,21]. Other authors make use of “black-box” methods [22] instead [14,23,13]. As opposed to bottom-up methods, top-down approach implies spatial downscaling procedures from a broader aggregated scope to the city- district- or building-level. Studies [8,9,24–26] suggested top-down methods as suitable for examining spatial variations of energy use for numerous purposes.

Regardless of which, bottom-up or top-down, approach is used, the attempts are made to model the phenomenon that lacks either detailed and complete knowledge or order or pattern or coherence or a combination of these. Booth et al. [27], systematized the uncertainties behind bottom-up engineering-based built stock models by their origins: a) variability of energy use due to chance within the identical buildings (aleatory uncertainties); b) heterogeneity of buildings within groups or typologies; and c) epistemic uncertainties which accommodate lack of knowledge about the phenomenon, the choice of inadequate model parameters and/or the risk of obtaining a biased model. These are also applicable to “black-box” methods that seek to approximate the uncertainties and assume a likewise deterministic relationship between the variables.

Probabilistic modelling enables to account for uncertainties and to address the limitations of approximating them in a deterministic alternative [27]. At a building level, the available studies quantify the uncertainties probabilistically in either forward or inverse manner [28]. Forward uncertainty propagation is dominated by sampling methods, where the inputs of the model are intentionally varied to obtain the likely variations of model outputs. Amongst the built stock energy studies, forward propagation principles were used to account for the epistemic uncertainties [27,29,30]. Inverse uncertainty propagation methods aim to relate the observed empirical data to both known and unknown model parameters and/or built stock properties. Given the underlying statistical inference approach, Tian et al. [28] categorised the inverse uncertainty analysis practices in the discipline as either frequentist or Bayesian. Whereas the former consists of methods for operating solely on empirical data, the latter also accommodates the prior knowledge and beliefs to aid the inferential statistics. In the context of built stock energy modelling, inverse Bayesian-based inference is advocated in studies [16,31,32]. Built stock energy studies with frequentist inference were not found in the domain-specific literature.

If combined with the inverse uncertainty propagation principles, top-down modelling reveals numerous useful properties. Both, aleatory uncertainties and heterogeneity of buildings are reflected in the built stock energy data at the aggregated spatial level. A step-wise disaggregation of this data with exogenous factors naturally conveys the associated uncertainties in an inverse manner. An empirically inferred probability density function (PDF) is a proxy for central tendency and variability due to yet unexplained uncertainty at each of these steps. Consequently, every subsequent disaggregation may lead to better estimates of uncertainties and thus, higher modelling accuracy. If the rational disaggregation steps are reflected in the structure of the model, numerous advantages directly or indirectly stem from the following:

- Uncertainties can be quantified at each level of the step-wise disaggregation, which enables a modeller to make judgements about the quality of the model and to control the trade-off between the expected accuracy gains versus additional data feed;
- The levels of sensitivity to adding the exogenous factors can be quantified through statistical hypothesis testing. This can prevent from using the redundant or insignificant model inputs and thus, to address the overfitting;
- Data requirements to achieve the necessary level of modelling details can be calculated beforehand. This enables to set up and efficiently manage the data collection process.

Despite these substantial advantages, top-down probabilistic modelling remains poorly explored within the discipline. This article, therefore, is motivated by the need to elaborate on the workflow, methods, and procedures that such modelling may involve. It is shown that this modelling approach may facilitate spatially-explicit energy mapping with a flexible spatial resolution and the

varying levels of details. Exemplification is made through the case study developed for Trondheim, Norway. Section 2 Method explains the non-parametric probabilistic model, data resources that the model relies on and a collection of methods which are illustrated using the available data. All these components are further synthesised into a coherent computational procedure used to obtain the estimates of bulk total annual energy use over 1×1 km geospatial grid. Section 3 Results provides both, intermediate and final outputs obtained through this procedure. Section 4 Discussion evaluates the strengths, weaknesses and potential further developments for such modelling. Extra care is taken to elaborate on the capabilities of top-down models to account for an increasingly detailed architectural and technical information needed for built stock energy research. And to discuss the role of statistical inference in further shaping the available domain knowledge. The Conclusions (Section 5) summarise the findings made in this study, reflect upon potential opportunities and barriers for further developments.

2. Method

The proposed procedural framework facilitates estimation of bulk (for all buildings) total (for all energy sources) annual energy use in geospatial zones. This can be achieved with a non-parametric model described in Section 2.1. Two data sources, namely the National Cadastral System (Section 2.2.1) and the Energy Performance Certificates (EPC) dataset (Section 2.2.2) provide the inputs into the model. Whereas accurate information on geospatial positioning, size and type of buildings and dwellings is available, the data on real energy performance are relatively scarce and available only at a city level. Therefore, the computational procedure involves Monte-Carlo (MC) simulation and considers the energy use intensity of buildings as a continuous random variable (r.v.). Inferring the properties of this r.v. from the available sample is the subject to density estimation procedure described in Section 2.3. Section 2.4 provides the description of a comprehensive computational procedure used to achieve the desired results – the estimates of central tendency and dispersion of bulk total annual energy use by buildings per geospatial zone.

The smallest element of built stock accounted for in the model (Section 2.1) is a building unit,[1] which enables to harmonise the data and to preserve the consistency across all steps of the study. Also, this allows to explicitly account for the energy performance of buildings that have a mixed use purpose, e.g. offices and apartments across multiple floors of the same building.

2.1. Probabilistic model

The simplest non-parametric model for estimating the bulk total annual energy use $E_{zone\ tot}$ of a geospatial zone consisting of $j \in [0, m]$ units can be defined as:

$$E_{zone\ tot} = \sum_{j=0}^{m} (a_j \cdot r_j) \qquad (kWh \cdot y^{-1}) \tag{1}$$

where:

a_j – heated floor area (m^2) of j^{th} unit;
r_j – energy use intensity ($kWh \cdot m^{-2} \cdot y^{-1}$) of j^{th} unit.

[1] The smallest element registered (in the cadastral system) or certified (in the EPC system): dwelling if residential use purpose; the section or the whole building otherwise.

The need to account for exogenous variables (disaggregation), in order to reflect the properties specific to the group of units, entails modifying the Eq. (1). A generalised form of the top-down model in Eq. (1) accepts t categorical variables, each of which has $k_t \in [0, l_t]$ categories:

$$E_{zone\ tot} = \sum_{k_1=0}^{l_1} \sum_{k_2=0}^{l_2} \cdots \sum_{k_t=0}^{l_t} \sum_{j=0}^{m} (a_{k_1,k_2,\ldots,k_t,j} \cdot r_{k_1,k_2,\ldots,k_t,j}) \qquad (kWh \cdot y^{-1}) \tag{2}$$

For this study, the generalised model in Eq. (2) is adapted to accommodate the typology-specific information. Thus, a model for estimating bulk total annual energy use $E_{zone\ tot}$ of a geospatial zone consisting of n building types with m units is defined as:

$$E_{zone\ tot} = \sum_{i=0}^{n} \sum_{j=0}^{m} (a_{i,j} \cdot r_{i,j}) = \sum_{i=0}^{n} A_i \cdot R_i^T \qquad (kWh \cdot y^{-1}) \tag{3}$$

where:

$a_{i,j}$ – heated floor area (m^2) of j^{th} unit of the i^{th} type;
$r_{i,j}$ – energy use intensity ($kWh \cdot m^{-2} \cdot y^{-1}$) of j^{th} unit of the i^{th} type;
$A_i = [\, a_{i,0} \ldots \ a_{i,m} \,]$ – a row matrix containing the known values of heated floor area (m^2) of all m units of the i^{th} type;
$R_i^T = [\, r_{i,0} \ldots \ r_{i,m} \,]^T = \begin{bmatrix} r_{i,0} \\ \ldots \\ r_{i,m} \end{bmatrix}$ – a column matrix containing the unknown values of energy use intensity ($kWh \cdot m^{-2} \cdot y^{-1}$) of all m units of i^{th} type.

Obtaining $E_{zone\ tot}$ without knowing the exact values of R_i^T represents an inverse probability estimation problem: "given the known univariate distribution of the uncertain model input, estimate the distribution of uncertain model output by repetitive random sampling of these inputs". This procedure is referred to as Monte Carlo simulation. With top-down reasoning, the distribution of R_i^T is inferred from higher spatial level, i.e. from the data available for the city.

Since the model in Eq. (3) applies in-sample summation, the simulated output tends towards normal distribution (Fig. 1) as the number of simulations gets larger, according to the Central Limit Theorem. Therefore, the output across simulation trials is well described by two parameters: a measure of central tendency given by the mean value μ and the dispersion properties quantified by standard deviation (SD) σ.

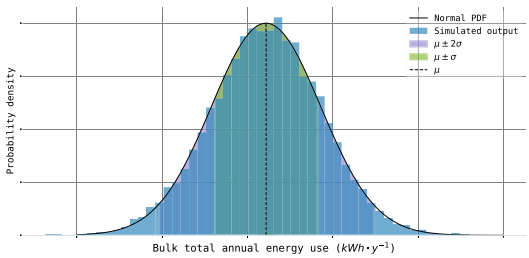


Fig. 1. Mean and standard deviation in the univariate distribution of simulated results.

2.2. Data sources

2.2.1. The size and structure of built stock

The known inputs of the Eq. (3) are the values of heated floor area and building type for each unit within the geospatial zone. This information is being collected, updated and made available through the Norwegian Cadastral system which is managed by the Norwegian Mapping Authority. The Norwegian Cadastral system was established in 2010 following the guidelines and the requirements of the Land Act 2005, offering a reliable, transparent and updated registry of all land users [33]. Currently, the registry contains almost 5 millions of registered properties nationwide, classified according to the Standard for building types [34].

Concerning the Trondheim municipality, the cadastral system's registry contains more than 92000 units covering 12 km^2 of total constructed floor area, 83% of which is residential. A spatial join of the attributes of these units enables the analysis and simulation per geospatial zones with a flexible spatial resolution, ranging from the individual building to the city level. A square grid of 1×1 km is an arbitrary choice of spatial resolution made to exemplify the computational procedures proposed in this study. Fig. 2 illustrates this geospatial grid over the urban territory and some attributes of the built stock per grid cell analysed: total constructed floor area (colour intensity) and the share of residential in the constructed floor area (marker size). The figure suggests that high construction density is present in the historical centre of the city and in the more industrialised southern part. These areas are often associated with a higher rate of non-residential units. The majority of the urban territory, however, is represented by sparse construction density and is dominated by residential buildings.

2.2.2. Energy performance of built stock

Inferring the statistical properties of the r.v. energy use intensity ($kWh \cdot m^{-2} \cdot y^{-1}$) per building type in Trondheim is based on the Norwegian EPC dataset. EPC dataset is the component of the Norwegian Energy Labelling System for Houses and Dwellings – a mechanism established to support the progress towards low energy use in communities and nationwide. The Norwegian EPC Scheme follows the implementation of the Energy Performance of Buildings Directive (EPBD), similarly to the other EU's Member States [35,36].

Hence, the Norwegian EPC scheme has been in place since 2010 intended to ensure Norway's compliance with the EPBD 2002/91/EC, to improve building energy awareness and to promote high energy performance. By 2016, more than 670 000 certificates were issued. The background for certification, legislative and practical framework in the Norwegian context was discussed in source [37].

The total annual energy use ($kWh \cdot y^{-1}$) per certified unit is voluntarily specified and registered in approx. 10% of all certificates. These values, normalised per unit of heated floor area (m^2), constitute an empirical sample of 4660 records representing dwelling/building units registered in Trondheim. Fig. 3 illustrates the univariate distribution of energy use intensity in this sample for both, non-residential (NR) and residential (RE) units. The figure demonstrates that the statistical properties of energy use intensity, accommodated by the shape of the density histogram, vary significantly per building type. This applies to such parameters as dispersion (range of values and variance), central tendency (mean, median, and mode), skewness and kurtosis. Accounting for these distinct properties is expected to positively contribute to the accuracy of best estimates and the margins of error provided by the built stock model.

2.3. Density estimation

To simulate R_i^T for the arbitrary number of units, one must know the relative likelihood of its values to occur. This information is communicated with two properties of a parameterised theoretical r.v: PDF and cumulative distribution function (CDF). Deciding which distribution type and parameters characterise the theoreti-

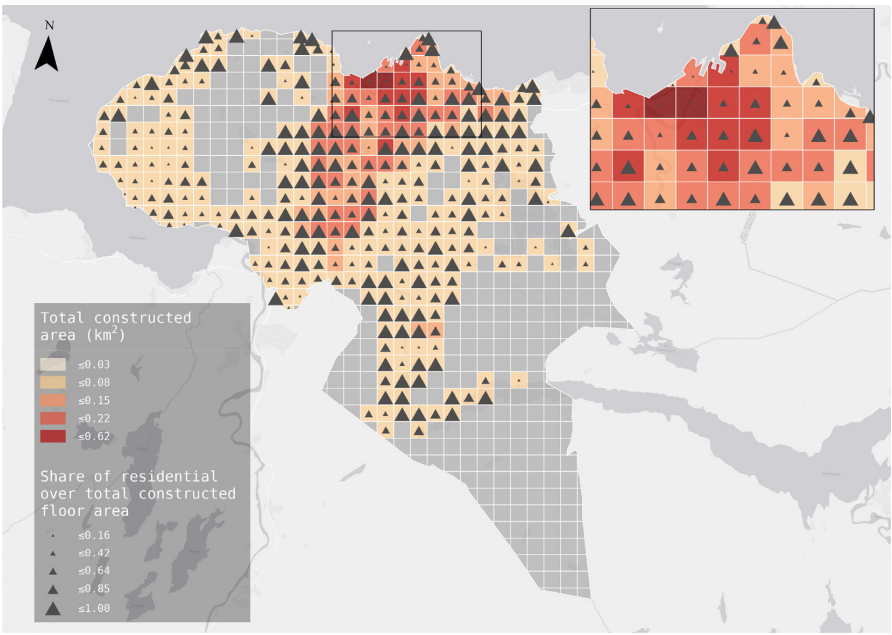


Fig. 2. Spatial variabilities of total constructed floor area and the share of residential buildings per geospatial zone in Trondheim.

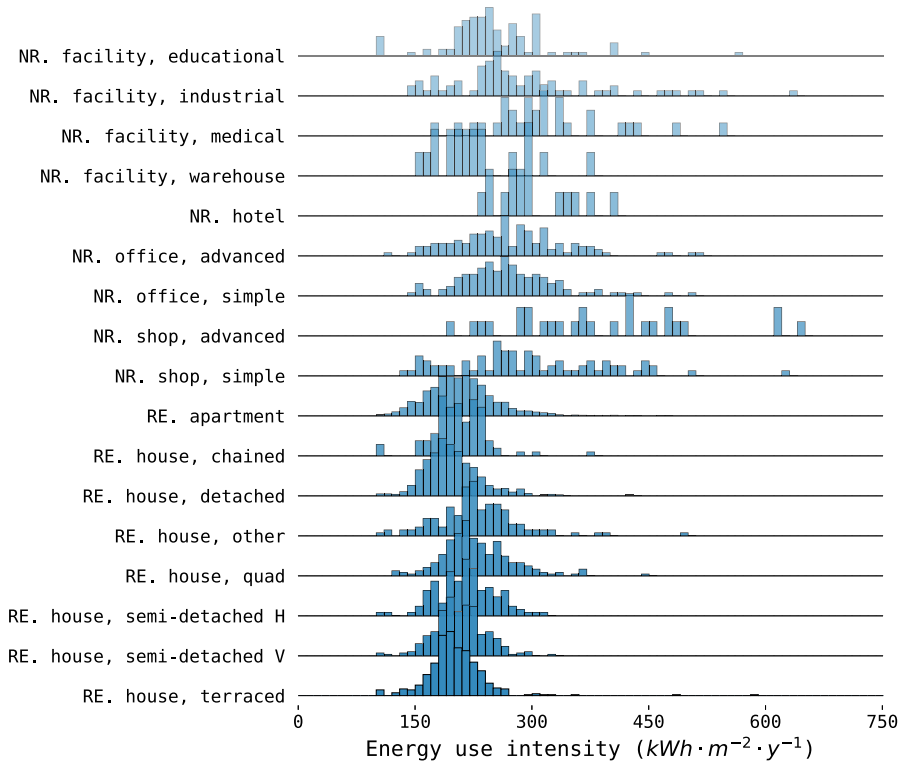


Fig. 3. Univariate distribution of energy use intensity per building type in Trondheim.

cal continuous r.v. can be carried out within the density estimation procedure consisting of:

1. Fitting (parameterising) each available distribution type individually based on the empirical sample;
2. Quantifying the goodness-of-fit;
3. Finding the theoretical parameterised distribution that characterises the sample best.

Maximum likelihood estimation (MLE) [38,39] is one way to fit the parameters of the theoretical distribution to the sample. In MLE, the objective is set as: given the observed sample $x : [x_1, x_2, x_3, \ldots, x_n]$ and the theoretical continuous PDF $p_X(x|\theta)$, find the vector θ of parameters that are most likely to generate such sample. This is achieved through maximising the log-likelihood function:

$$f(\theta, x) = max_\theta \left\{ ln \left[\prod_i^n p_X(x_i|\theta) \right] \right\} = max_\theta \left\{ \sum_i^n ln[p_X(x_i|\theta)] \right\} \quad (4)$$

Finding the objective function in Eq. (4), represents a multivariate unconstrained optimisation problem with potentially noisy (non-smooth) functions. An effective search method for the problems of this kind is downhill simplex (Nelder–Mead) method [40–42] which is also known as a generalisation of dichotomic search to higher dimensions. Depending on the distribution type, vector θ may have between two and five parameters, meaning that the simplex takes a form of a triangle, tetrahedron, pentachoron or 5-simplex accordingly. Convergence to the optima is carried out through stepwise improvement of the initial guess without computing the gradients. The exit condition is either achieving the desired error tolerance or lack of progress in objective function compared to previous iterations.

Fig. 4 illustrates the results of the MLE-based fitting of some theoretical distributions to the empirical sample (also shown in Fig. 3) of detached houses in Trondheim. It is shown that the PDFs (Fig. 4 [A]) and CDFs (Fig. 4 [B]) follow the shape of sample distribution with various precision. This entails deciding which distribution describes the sample best and requires quantitative metrics to facilitate the decision.

The goodness-of-fit between the continuous theoretical distribution and the empirical sample may be studied with a non-parametric Kolmogorov–Smirnov (KS) test [43,44]. The KS test quantifies the difference between the empirical CDF represented by the step-function and the CDF of the theoretical distribution (as shown in Fig. 4 [B]). The test returns two values of interest: the D_n statistic (Eq. 5) and the measure of statistical significance (p-value).

D_n for the sample with sample size (SS) n is the supremum (the maximum or the bound) of the absolute difference between the CDF of a theoretical distribution $P_0(x)$ and the empirical CDF $\hat{P}_n(x)$ [45]:

$$D_n = \sup_x |P_0(x) - \hat{P}_n(x)| \quad (5)$$

The p-value corresponds to the survival function $(1 - CDF)$ in the asymptotic distribution of D_n at $\sqrt{n} \cdot D_n$. In statistical hypothesis testing, p-value serves as the basis for accepting or rejecting the hypotheses about the conformity between distributions. Low p-value suggests statistically significant evidence against the asserted

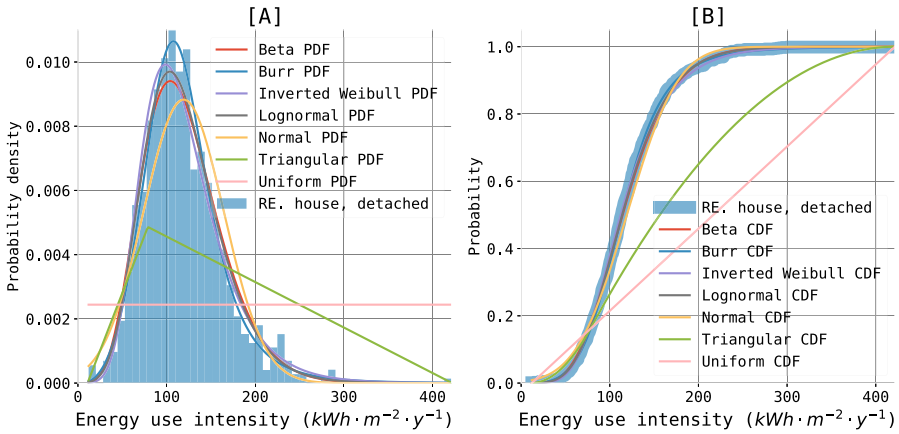


Fig. 4. Sample density [A] and the empirical CDF [B] with fitted PDFs and CDFs accordingly for some theoretical distributions.

null hypothesis: "sample x is generated by the r.v. X with the PDF $p_X(x|\theta)$". The threshold value α for statistical significance has to be chosen prior to the experiment. The null hypothesis, therefore, is rejected under the condition $p < \alpha$ and is failed to be rejected otherwise.

2.4. Computational framework

The components of the probabilistic model (Section 2.1) together with methods and procedures discussed in Section 2.3 are organised in a computational framework (Fig. 5).

Density estimation component is designed to find the parameterised theoretical distributions that describe the energy use intensity of distinct building types. The process starts with obtaining a subset of the sample with energy use intensity corresponding to particular building type in the city. For each of the available theoretical distributions, their parameters are fitted with MLE using the downhill simplex method. MLE is terminated either if the objective function is found with the absolute error tolerance $\epsilon \leqslant 1 \cdot 10^{-10}$ or if the maximum number of iterations $N \times 200$ (N – the number of simplex's dimensions) is achieved. The KS test is then carried out with the CDFs of an empirical sample and of a fitted distribution. The null hypothesis is rejected under the condition $p < 0.05$. At the end of the loop, the most suitable distribution amongst those passing the test is selected. This choice is based on comparing the associated D statistic and selecting the smallest. The procedure is then repeated for all building types in Trondheim.

Within the simulation component, the primary loop carries out iterations over the grid cells. In each cell, a secondary loop iterates over the building types that are present which is followed by retrieving a row matrix A_i. A series of 10000 Monte-Carlo trials are then carried out using the Mersenne Twister [46,47] pseudo-random number generator. At each trial, a column matrix R_i^T is simulated as a r.v. using the previously found parameterised distribution that characterises this building type. The dot product $A_i \cdot R_i^T$ is computed per trial and stored as one likely value of total energy use by the typology in the cell. When the iterations over the building types are complete, the total energy use across simulation trials per typology are aggregated to the grid cell level. This output forms a normal distribution, the mean value μ and the standard deviation σ for which are computed.

3. Results

The output of Density Estimation component, as discussed in Sections 2.3 and 2.4, are the parameterised distributions that are found to represent the data generation processes for individual building types in Trondheim. This information is summarised, together with the metrics for goodness-of-fit and sample statistic (minimum/maximum values and the sample size) in Table 1.

Vector θ of distribution's parameters in Table 1 is structured as $\theta : [\theta_1, \ldots, \theta_{k-1}, \theta_k]$. Two last elements in the list θ_{k-1} and θ_k are location (l) and scale (s) parameters accordingly. Any additional shape parameters, if applicable, are at the beginning of this list.

An example of interpreting the information provided in Table 1 is the following: energy use intensity of "RE. house, terraced" in Trondheim conforms to Johnson SU distribution parameterised by vector $[-0.392,\ 1.309,\ 108.848,\ 40.094]$. The difference between the empirical CDF of a sample with the size 407 and the CDF of this theoretical distribution is found to be 0.02. This difference is insignificant ($p > \alpha:\ 0.96 > 0.05$), thus implying a failure to reject the asserted null-hypothesis "the empirical sample is generated by Johnson SU r.v. with these parameters". The empirically evident range of values taken by the r.v. is $[25,\ 623]\ kWh \cdot m^{-2} \cdot y^{-1}$. Energy use intensity of "RE. house, terraced" in Trondheim within this range, therefore, can be simulated as the Johnson SU r.v. that has the PDF of a form:

$$f(x,a,b,l,s) = \frac{b}{s\sqrt{\left(\frac{x-l}{s}\right)^2+1}}\phi\left(a + b\log\left(\frac{x-l}{s} + \sqrt{\left(\frac{x-l}{s}\right)^2+1}\right)\right) \tag{6}$$

where:

- x – energy use intensity ($kWh \cdot m^{-2} \cdot y^{-1}$);
- ϕ – normal PDF;
- a, b, l, s – a list of parameters identified with MLE:
  - $[a,\ b] = [-0.392,\ 1.309]$ – distribution-specific shape parameters;
  - $[l,\ s] = [108.848,\ 40.094]$ – location and scale parameters accordingly.

The outputs of a probabilistic model (Eq. (3) in Section 2.1), produced within the Simulation component (Section 2.4) using the parameterised distributions listed in Table 1 are the estimates of the mean (μ) and the SD (σ) of the bulk total annual energy use

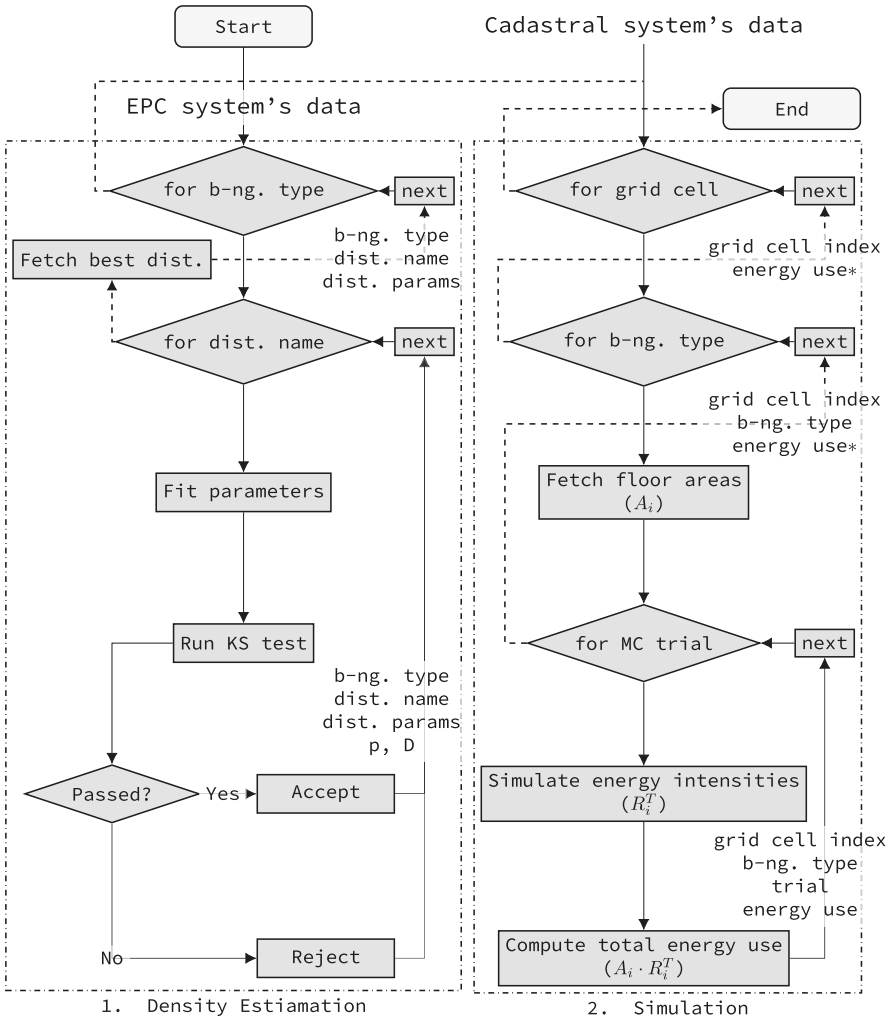


Fig. 5. The flowchart of computational procedures (Column or row matrices are denoted by asterisk $*$. Otherwise, a single categorical or numerical value is returned.).

Table 1
Sample information and parameterised distributions identified per building type in Trondheim.

Building type	Distribution	Parameters	D	*p*-value	Min	Max	SS
NR. facility, educational	Mielke Beta-Kappa	[6.444, 5.037, −0.446, 172.923]	0.06	0.85	55	602	117
NR. facility, industrial	Folded Cauchy	[2.612, 47.979, 57.731]	0.06	0.87	48	698	80
NR. facility, medical	Log-laplace	[3.662, −0.615, 264.554]	0.06	1.00	90	573	31
NR. facility, warehouse	Alpha	[3.785, −68.437, 771.015]	0.11	0.98	65	353	17
NR. hotel	Inverse Gaussian	[0.365, 136.117, 324.874]	0.12	0.93	167	390	17
NR. office, advanced	Logistic	[210.129, 51.094]	0.03	1.00	15	535	128
NR. office, simple	Tukey-Lambda	[−0.087, 205.013, 38.051]	0.04	0.99	51	524	129
NR. shop, advanced	Maxwell	[20.864, 219.505]	0.08	0.99	115	703	28
NR. shop, simple	Exponentially modified normal	[1.036, 154.399, 88.784]	0.06	0.93	33	680	80
RE. apartment	Mielke Beta-Kappa	[2.641, 5.967, −0.348, 163.23]	0.01	0.91	12	467	1844
RE. house, chained	Vonmises (non-circular)	[3.808, 131.935, 69.196]	0.07	0.67	66	349	103
RE. house, detached	Exponentially modified normal	[1.374, 82.776, 26.539]	0.01	1.00	12	422	881
RE. house, other	Tukey-Lambda	[−0.156, 157.477, 27.537]	0.04	0.98	17	516	136
RE. house, quad	Johnson SU	[−0.726, 1.561, 121.128, 67.087]	0.04	0.83	22	438	243
RE. house, semi-detached H	Rice	[1.477, 44.969, 53.336]	0.04	0.99	51	276	124
RE. house, semi-detached V	Exponentially modified normal	[0.645, 111.087, 34.003]	0.03	0.98	16	292	295
RE. house, terraced	Johnson SU	[−0.392, 1.309, 108.848, 40.094]	0.02	0.96	25	623	407

per grid cell. These attributes are displayed in Fig. 6. Colour intensities in the figure illustrate spatial variations of μ, whereas the diameter of markers is proportional to σ.

Fig. 6 facilitates the analysis of the city's energy hotspots (areas with high μ) and where the additional information may be needed (high σ). Examining these results against the data on the built stock in Fig. 2 demonstrates that the mean of the simulated bulk total annual energy use in Fig. 6 is correlated with the built area density and the share of energy-intensive non-residential buildings. The energy hotspots are located in the centre of the city and the industrial southern suburbs. Remote, mostly residential areas, which are known to have low unit density, are associated with relatively low energy use. Standard deviation correlates with mean of simulated bulk total annual energy use in the geospatial zones. Further analysis suggests a roughly linear relationship (Fig. 7) between μ and σ.

The scatter plot in Fig. 7 presents the results separated into two groups based on the arbitrary condition $c \geqslant 0.1$ and $c < 0.1$ where c is the coefficient of variation ($c = \sigma/\mu$). More detailed analysis suggested that c exceeds 0.1 for those grid cells where constructed units density is sparse or, alternatively, dense but with a high share of non-residential build area (above 60%). For most of the areas where energy use is high, c remains below 0.1. According to the empirical interpretation of normal distribution, $c < 0.1$ suggests at least 67% of confidence that the true value of bulk total annual energy use in the spatial zone is within the range $\mu \pm 10\%$. Similarly, $\mu \pm 20\%$ is the 95% confidence range.

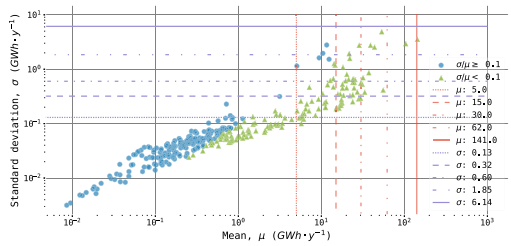


Fig. 7. The relationship between mean and standard deviation of simulated bulk total annual energy use per geospatial zone.

4. Discussion

Through the case study developed for Trondheim, this article demonstrates a top-down modelling approach with the inverse uncertainty propagation for urban energy mapping purposes. Methodically, it implies spatial downscaling of the energy use intensity values from the city-level to the finer resolution. As a result, the probabilistic estimates of bulk total annual energy use per geospatial zone may be obtained. Random sampling used to compute these estimates enables to address aleatory uncertainties and heterogeneity discussed by Booth et al. [27]. The model in Eq. (3) does not contain any parameters and does not assume any rigorous knowledge about the factors that drive the phenomenon, thus eliminating the associated epistemic uncertainties.

Typology-specific density estimation of energy use intensity carried out at the city-level enabled to downscale the analysis to 1×1 *km* square grid. This choice of spatial resolution was arbitrary and can be substituted in the model with any other spatial or administrative boundaries. The parsimonious model design options enable the upper and lower boundaries to be anywhere between national and the building levels accordingly.

Within the modelling framework, disaggregation by exogenous influencing factors is supported and may lead to more accurate estimates. The simplest top-down model (Eq. (1)), for example, would contain a single parameterised r.v. that simulates energy intensity of all units with no regards to any other exogenous fac-

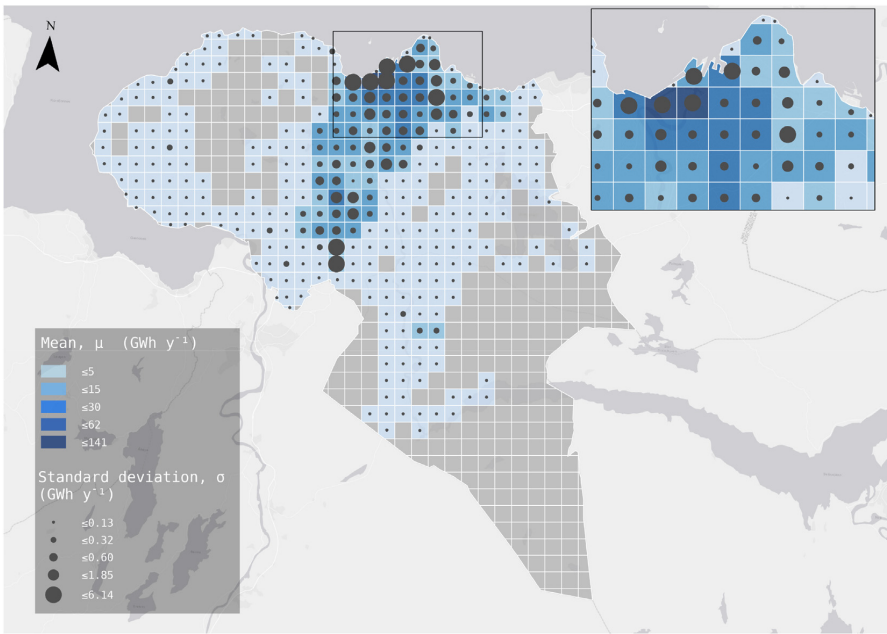


Fig. 6. Spatial variation of the mean and the standard deviation in simulated bulk total annual energy use per grid cell in Trondheim.

tors in the computational framework (Section 2.4). In this study, disaggregation by building types is beneficial because typology-specific r.v. evidently conveys a more detailed empirical information about the unique data generation process, i.e. relative likelihood of energy use intensity values to occur. Such disaggregation in a top-down manner may be done further to reach the necessary levels of technical details. Fig. 8 illustrates one of the plausible next steps in the disaggregation procedure to reflect distinct energy performance in "NR. apartment" given the Construction Year (CY).

Fig. 8 illustrates that the originally bimodal distribution at a higher (Fig. 8 [A]) level may be successfully disaggregated into at least two unimodal groups (Fig. 8 [B]). The basis for disaggregation in this case is arbitrary set to the year 1955, which is the first threshold between construction year classes defined for the Norwegian built stock within the TABULA [48] system. The empirical evidence supports such archetypes definition, since the two samples are characterised by substantially distinct statistical properties. Stricter energy efficiency standards for buildings, together with the other contributing factors, directly or indirectly led to the observable shift of central tendency between the two distributions. Dispersion is also affected by regulatory, technological, and socio-economic transformations.

A measure of dispersion quantifies yet unexplained uncertainties. By examining this parameter, a modeller may decide whether the remaining level of uncertainty is acceptable to address the purpose of the modelling or if further disaggregation is needed. If the latter, additional exogenous factors may be tested. Two distributions in Fig. 8 [B] are associated with the dispersion smaller than the composite (Fig. 8 [A]), meaning that the construction period explains a portion of the original uncertainties.

Statistical significance of the difference between the disaggregated distributions suggests the level of sensitivity to disaggregation by the exogenous factors. Since the resulting distributions in Fig. 8 [B] are distinct, it is plausible to disaggregate by construction period. The reverse statement also holds true – mutually conforming distributions exhibited by disaggregated groups suggest a little or no benefit from disaggregation. The previously mentioned KS test may be used to quantify the difference between the two empirical samples. A null hypothesis for testing is formulated as: "two samples are drawn from the same continuous distribution", and high p-value ($p > \alpha$) implies a failure to reject this null hypothesis. An example of such pairwise testing of samples with the significance threshold $\alpha = 0.05$ is presented in Table 2.

Test 1 in Table 2 resulted in a large D-statistic. If the null hypothesis is true, obtaining such a large value of D by chance is unlikely given the samples sizes. This likelihood is reflected by the negligibly low p-value which suggests to reject the asserted null-hypothesis. This indicates high sensitivity to the construction period. The null hypothesis cannot be rejected in Test 2 and therefore, these two samples are found to conform even though they represent distinct archetypes in TABULA. Higher likelihood of apartments from Period 1 to be in their renovated state may explain the absence of statistically significant difference between energy use intensities of samples in Test 2.

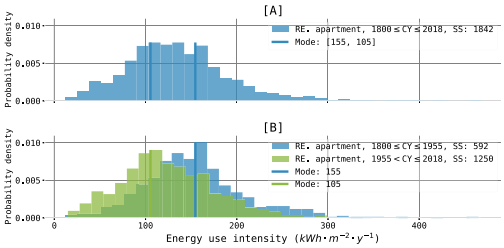


Fig. 8. Univariate distribution of energy use intensity of apartments in Trondheim: [A] composite; [B] disaggregated by construction period.

The disaggregation procedure discussed above reveals the source of both, advantages and limitations of the proposed approach: if exogenous factors of influence on energy use intensity are not represented in the model, they are assumed consistent between the upper and the lower spatial levels. Practically, it suggests that the model needs to reflect only those factors that are known to lead to spatial variations of energy use. Otherwise, the unexplained uncertainty of energy use intensity entails larger dispersion in the simulated results. Although the most essential operations for density estimation and probabilistic simulation can be automated, the choices behind disaggregation procedures remain manual. This entails that the choice of the acceptable unexplained uncertainty level and the number of categories for disaggregation must involve domain knowledge even if these judgements are supported by quantitative metrics.

The available sample size is an important aspect that represents a source of epistemic uncertainty associated with the proposed modelling approach. Density estimation with insufficient sample size may suggest the type of distribution or the parameters that poorly describes the data generation process and should be avoided whenever possible. A frequentist-based density estimation discussed in this article may inform empirically the choice of prior distribution in Bayesian inference, which may lead to the need for fewer samples and higher reliability of the latter. Therefore, complementing the two approaches may be beneficial for future studies. A rational threshold α of statistical significance for the p-value needs to be established in the discipline to support the coherence between the studies alike. Moreover, a rigorous recommendation on the domain-specific smallest sample size for density estimation is not yet available. Therefore, this study agrees with Booth et al. [27] on the need for adapting the existing practices from other domains, e.g. physics, medicine, and economy to assist overcoming these challenges in built stock energy modelling. The implications of data quality can be regarded as an additional source of epistemic uncertainty in modelling. High accuracy and soundness of conclusions made through statistical inference, similarly to other techniques that rely on data, require independent and identically distributed (i.i.d.) random samples. A practical way to obtain such data is through stratified (e.g. block) design of the experiment. Substituting such sample with potentially biased data may entail inaccuracies. It is, for example, debatable if any of the EPC system's designs is capable of providing randomised i.i.d. samples, because the sole phenomenon of certification is under the strong influence of numerous socio-techno-economic tendencies that may cause the bias.

Given the availability and high quality of empirical data, however, virtually any level of technical details and end-uses may be reached through such modelling. No obstacles are anticipated in evaluating the implications of altering building envelope, energy supply systems and/or indoor environmental quality at a large scale. Similar approaches can also accommodate energy use for source-specific space heating, hot water supply, plug loads and other through multivariate distributions, e.g. copulas. Currently, these capabilities of top-down modelling are underestimated and poorly explored in the domain. It is, however, shown through this study that handling aleatory uncertainties and heterogeneity of buildings yields numerous benefits, mitigating the "performance gap" being one of them. It is also evident that data-enabled knowledge discovery and modelling, facilitated by statistical inference and probabilistic programming, may complement already established architectural and engineering-based foundations of built stock energy research. Synthesising the methods from these

Table 2
Pairwise KS test results on conformity between samples from various construction periods for apartments in Trondheim

Test	Period 1 (SS)	Period 2 (SS)	D	p-value	Null-hypothesis
1	$1800 \leqslant CY \leqslant 1955(592)$	$1955 < CY \leqslant 2018(1250)$	0.256	$1.212 \cdot 10^{-23}$	$p < 0.05$: Reject
2	$1955 < CY \leqslant 1980(642)$	$1980 < CY \leqslant 2000$ (147)	0.086	0.323	$p > 0.05$: Failed to reject

domains is also the key objective set for the future developments of Built Stock Explorerhttps://builtstockexplorer.indecol.ntnu.no.

5. Conclusions

This study draws attention to the topic of urban energy mapping, where the uncertainties must be eliminated to the best possible extent while keeping data collection efforts rational. A proposed probabilistic top-down modelling approach is shown to have a high potential for addressing this trade-off. Probabilistic origins naturally account for aleatory uncertainties behind the phenomena and for the heterogeneity of buildings through parameterising the random variables. Disaggregation by exogenous factors conveys these uncertainties without the loss of information. Non-parametric model structure enables to address the epistemic uncertainties, associated with approximations and simplifications that are necessary otherwise. A suggested modelling approach offers adaptiveness to the purpose of the modelling and the associated level of details. The key benefits of the approach emerge from the ability to quantify and control the uncertainties while adding the explanatory (exogenous) factors to the model.

The results suggest that the typology-specific energy use intensity can be represented by parameterised random variables. With these random variables and the information on geospatial coordinates, size and type of buildings, bulk total annual energy use can be estimated at a spatially downscaled area, e.g. 1×1 km grid cell.

The coefficient of variation for most of Trondheim's energy hotspots remains below 0.1, which makes the results already suitable for many practical applications. Urban areas of high energy use, for example, can be prioritised for developing refurbishment strategies and/or deploying more efficient energy supply solutions. By resolving the energy-related bottlenecks first, higher energy and environmental performance of built stock may be achieved within a shorter time horizon. These results may also aid the planning of new construction and the energy-intensive units with a minimum intervention into the existing infrastructure for energy generation and distribution purposes.

Considering the additional factors of influence on building energy performance may further improve the accuracy of the modelling. The feasibility of using these factors can be guided by statistical hypothesis testing. Currently, a substantial barrier for such modelling is the absence of both, established practices for defining the levels of statistical significance and the recommendations on sample size for such tasks. These and the related challenges entail a yet unresolved epistemic uncertainty associated with the proposed modelling approach.

The instruments and the techniques discussed in this article may produce reliable insights into the spatial variabilities of the building energy use. They lay the foundations for the work ahead which will synthesise the probabilistic status quo with the probabilistic forecasting of future developments in the built stock. And hence, will assist with establishing the pathways towards higher efficiency of the built environment.

CRediT authorship contribution statement

Ruslan Zhuravchak: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Visualization. **Raquel Alonso Pedrero:** Conceptualization, Software, Formal analysis, Writing - original draft, Visualization. **Pedro Crespo del Granado:** Validation, Writing - review & editing. **Natasa Nord:** Validation, Writing - review & editing, Supervision. **Helge Brattebø:** Validation, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the support from the Research Council of Norway through the research projects:- Methods for transparent energy planning of urban building stocks - ExPOSe (project number 268248) under EnergiX program;- The value of end-use flexibility in the future Norwegian energy system - Flexbuild (project number 294920).

References

[1] F. Johari, G. Peronato, P. Sadeghian, X. Zhao, J. Widén, Urban building energy modeling: State of the art and future prospects, Renewable and Sustainable Energy Reviews 128 (2020), https://doi.org/10.1016/j.rser.2020.109902 109902.
[2] T. Abergel, C. Delmastro, K. Lane, Tracking buildings 2020, Tech. rep., International Energy Agency (IEA), 2020.
[3] J.-P. Petersen, E. Heurkens, Implementing energy policies in urban development projects: The role of public planning authorities in Denmark, Germany and the Netherlands, Land Use Policy 76 (2018) 275–289, https://doi.org/10.1016/j.landusepol.2018.05.004.
[4] M. Kljajic, A.S. Andelkovic, I. Mujan, Assessment of relevance of different effects in energy infrastructure revitalization in non-residential buildings, Energy and Buildings 116 (2016) 684–693, https://doi.org/10.1016/j.enbuild.2015.02.033.
[5] N.H. Sandberg, I. Sartori, O. Heidrich, R. Dawson, E. Dascalaki, S. Dimitriou, T. Vimm-r, F. Filippidou, G. Stegnar, M.Š. Zavrl, H. Brattebø, Dynamic building stock modelling: Application to 11 European countries to support the energy efficiency and retrofit ambitions of the EU, Energy and Buildings 132 (2016) 26–38, 10.1016/j.enbuild.2016.05.100.
[6] L. Tozer, Catalyzing political momentum for the effective implementation of decarbonization for urban buildings, Energy Policy 136 (2020), https://doi.org/10.1016/j.enpol.2019.111042 111042.
[7] M. Shepero, J. Munkhammar, Spatial Markov chain model for electric vehicle charging in cities using geographical information system (GIS) data, Applied Energy 231 (2018) 1089–1099, https://doi.org/10.1016/j.apenergy.2018.09.175.
[8] D.J. Sailor, L. Lu, A top–down methodology for developing diurnal and seasonal anthropogenic heating profiles for urban areas, Atmospheric Environment 38 (17) (2004) 2737–2748, https://doi.org/10.1016/j.atmosenv.2004.01.034.
[9] S. Heiple, D.J. Sailor, Using building energy simulation and geospatial modeling techniques to determine high resolution building sector energy consumption profiles, Energy and Buildings 40 (8) (2008) 1426–1436, https://doi.org/10.1016/j.enbuild.2008.01.005.
[10] S.T. Moghadam, C. Delmastro, S.P. Corgnati, P. Lombardi, Urban energy planning procedure for sustainable development in the built environment: A review of available spatial approaches, Journal of Cleaner Production 165 (2017) 811–827, https://doi.org/10.1016/j.jclepro.2017.07.142.
[11] U. Ali, M.H. Shamsi, M. Bohacek, K. Purcell, C. Hoare, E. Mangina, J. O'Donnell, A data-driven approach for multi-scale gis-based building energy modeling for analysis, planning and support decision making, Applied Energy 279 (2020), https://doi.org/10.1016/j.apenergy.2020.115834 115834.
[12] M. Österbring, É. Mata, L. Thuvander, M. Mangold, F. Johnsson, H. Wallbaum, A differentiated description of building-stocks for a georeferenced urban bottom-up building-stock model, Energy and Buildings 120 (2016) 78–84, https://doi.org/10.1016/j.enbuild.2016.03.060.
[13] S.T. Moghadam, J. Toniolo, G. Mutani, P. Lombardi, A GIS-statistical approach for assessing built environment energy use at urban scale, Sustainable Cities and Society 37 (2018) 70–84, https://doi.org/10.1016/j.scs.2017.10.002.

[14] B. Howard, L. Parshall, J. Thompson, S. Hammer, J. Dickinson, V. Modi, Spatial distribution of urban building energy consumption by end use, Energy and Buildings 45 (2012) 141–151, https://doi.org/10.1016/j.enbuild.2011.10.061.
[15] Y. Yamaguchi, Y. Shimoda, M. Mizuno, Proposal of a modeling approach considering urban form for evaluation of city level energy management, Energy and Buildings 39 (5) (2007) 580–592, https://doi.org/10.1016/j.enbuild.2006.09.011.
[16] R. Choudhary, Energy analysis of the non-domestic building stock of Greater London, Building and Environment 51 (2012) 243–254, https://doi.org/10.1016/j.buildenv.2011.10.006.
[17] H. Yoshino, T. Hong, N. Nord, Iea ebc annex 53: Total energy use in buildings—analysis and evaluation methods, Energy and Buildings 152 (2017) 124–136, https://doi.org/10.1016/j.enbuild.2017.07.038.
[18] A.A.A. Gassar, S.H. Cha, Energy prediction techniques for large-scale buildings towards a sustainable built environment: A review, Energy and Buildings 224 (2020), https://doi.org/10.1016/j.enbuild.2020.110238 110238.
[19] L.G. Swan, V.I. Ugursal, Modeling of end-use energy consumption in the residential sector: A review of modeling techniques, Renewable and Sustainable Energy Reviews 13 (8) (2009) 1819–1835, https://doi.org/10.1016/j.rser.2008.09.033.
[20] J. Langevin, J. Reyna, S. Ebrahimigharehbaghi, N. Sandberg, P. Fennell, C. Nägeli, J. Laverge, M. Delghust, É. Mata, M. Van Hove, J. Webster, F. Federico, M. Jakob, C. Camarasa, Developing a common approach for classifying building stock energy models, Renewable and Sustainable Energy Reviews 133 (2020), https://doi.org/10.1016/j.rser.2020.110276 110276.
[21] C.F. Reinhart, C.C. Davila, Urban building energy modeling – a review of a nascent field, Building and Environment 97 (2016) 196–202, https://doi.org/10.1016/j.buildenv.2015.12.001.
[22] G. Tardioli, R. Kerrigan, M. Oates, J. O'Donnell, D. Finn, Data driven approaches for prediction of building energy consumption at urban level, Energy Procedia 78 (2015) 3378–3383, https://doi.org/10.1016/j.egypro.2015.11.754.
[23] A. Mastrucci, O. Baume, F. Stazi, U. Leopold, Estimating energy savings for the residential building stock of an entire city: A GIS-based statistical downscaling approach applied to Rotterdam, Energy and Buildings 75 (2014) 358–367, https://doi.org/10.1016/j.enbuild.2014.02.032.
[24] H.C. Gils, J. Cofala, F. Wagner, W. Schöpp, GIS-based assessment of the district heating potential in the USA, Energy 58 (2013) 318–329, https://doi.org/10.1016/j.energy.2013.06.028.
[25] M. Berger, J. Worlitschek, A novel approach for estimating residential space heating demand, Energy 159 (2018) 294–301, https://doi.org/10.1016/j.energy.2018.06.138.
[26] D. Meha, T. Novosel, N. Duić, Bottom-up and top-down heat demand mapping methods for small municipalities, case Gllogoc, Energy 199 (2020), https://doi.org/10.1016/j.energy.2020.117429 117429.
[27] A. Booth, R. Choudhary, D. Spiegelhalter, Handling uncertainty in housing stock models, Building and Environment 48 (2012) 35–47, https://doi.org/10.1016/j.buildenv.2011.08.016.
[28] W. Tian, Y. Heo, P. [de Wilde], Z. Li, D. Yan, C.S. Park, X. Feng, G. Augenbroe, A review of uncertainty analysis in building energy assessment, Renewable and Sustainable Energy Reviews 93 (2018) 285 – 301. doi:10.1016/j.rser.2018.05.029..
[29] M. Kavgic, A. Summerfield, D. Mumovic, Z. Stevanovic, Application of a Monte Carlo model to predict space heating energy use of Belgrade's housing stock, Journal of Building Performance Simulation 8 (6) (2015) 375–390, https://doi.org/10.1080/19401493.2014.961031.
[30] M. Hughes, J. Palmer, V. Cheng, D. Shipworth, Global sensitivity analysis of England's housing energy model, Journal of Building Performance Simulation 8 (5) (2015) 283–294, https://doi.org/10.1080/19401493.2014.925505.
[31] F. Zhao, S.H. Lee, G. Augenbroe, Reconstructing building stock to replicate energy consumption data, Energy and Buildings 117 (2016) 301–312, https://doi.org/10.1016/j.enbuild.2015.10.001.
[32] D. Yu, A two-step approach to forecasting city-wide building energy demand, Energy and Buildings 160 (2018) 1–9, https://doi.org/10.1016/j.enbuild.2017.11.063.
[33] Kommunal- og moderniseringsdepartementet, Lov om eigedomsregistrering (matrikkellova) LOV-2005-06-17-101 (2005). http://lovdata.no/lov/2005-06-17-101..
[34] STATISTISK SENTRALBYRÅ, Seksjon for eiendom-, areal- og primærnæringsstatistikk, Standard for bygningstype/matrikkelen (2000). https://www.ssb.no/klass/klassifikasjoner/31..
[35] M. Prieler, M. Leeb, T. Reiter, Characteristics of a database for energy performance certificates, Energy Procedia 132 (2017) 1000–1005, 11th Nordic Symposium on Building Physics, NSB2017, 11–14 June 2017, Trondheim, Norway. doi:10.1016/j.egypro.2017.09.704..
[36] Y. Li, S. Kubicki, A. Guerriero, Y. Rezgui, Review of building energy performance certification schemes towards future improvement, Renewable and Sustainable Energy Reviews 113 (2019), https://doi.org/10.1016/j.rser.2019.109244 109244.
[37] T. Brekke, O.K. Isachsen, M. Strand, EPBD implementation in Norway. Status in December 2016, Tech. rep., Enova, Norwegian Water Resources and Energy Directorate (NVE), Norwegian Building Authority (DIBK) (2018).
[38] C. Robert, G. Casella, Monte Carlo Statistical Methods, Springer Texts in Statistics, Springer New York, 2013. doi:10.1007/978-1-4757-3071-5..
[39] N.T. Thomopoulos, Statistical Distributions. Applications and Parameter Estimates, 1st Edition,., Springer International Publishing, 2017, 10.1007/978-3-319-65112-5.
[40] J.A. Nelder, R. Mead, A simplex method for function minimization, The Computer Journal 7 (4) (1965) 308–313, https://doi.org/10.1093/comjnl/7.4.308.
[41] F. Gao, L. Han, Implementing the Nelder-Mead simplex algorithm with adaptive parameters, Computational Optimization and Applications 51 (1) (2012) 259–277, https://doi.org/10.1007/s10589-010-9329-3.
[42] W. Härdle, O. Okhrin, Y. Okhrin, Basic elements of computational statistics, Springer International Publishing (2017), https://doi.org/10.1007/978-3-319-55336-8.
[43] R. Feldman, C. Valdez-Flores, Applied Probability and Stochastic Processes, 2nd ed., Springer-Verlag, Berlin Heidelberg, 2010, 10.1007/978-3-642-05158-6.
[44] R. Bhattacharya, L. Lin, V. Patrangenaru, A Course in Mathematical Statistics and Large Sample Theory, Springer Texts in Statistics, Springer New York, 2016. doi:10.1007/978-1-4939-4032-5..
[45] G. Marsaglia, W.W. Tsang, J. Wang, Evaluating Kolmogorov's distribution, Journal of Statistical Software, Articles 8 (18) (2003) 1–4, https://doi.org/10.18637/jss.v008.i18.
[46] M. Matsumoto, T. Nishimura, Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator, ACM Transactions on Modeling and Computer Simulation 8 (1) (1998) 3–30, https://doi.org/10.1145/272991.272995.
[47] R. Kneusel, Random Numbers and Computers, Springer International Publishing, 2018. doi:10.1007/978- 3- 319- 77697- 2..
[48] T. Loga, B. Stein, N. Diefenbach, Tabula building typologies in 20 european countries—making energy-related features of residential building stocks comparable, Energy and Buildings 132 (2016) 4–12, towards an energy efficient European housing stock: monitoring, mapping and modelling retrofitting processes. doi:10.1016/j.enbuild.2016.06.094..

Paper IV

Ruslan Zhuravchak, Natasa Nord and Helge Brattebø (2022). 'The effect of building attributes on the energy performance at a scale: an inferential analysis'. In: *Building Research & Information* 0.0, pp. 1–19. DOI: 10.1080/09613218.2022.2038537. eprint: https://doi.org/10.1080/09613218.2022.2038537

Building Research & Information

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rbri20

The effect of building attributes on the energy performance at a scale: an inferential analysis

Ruslan Zhuravchak, Natasa Nord & Helge Brattebø

To cite this article: Ruslan Zhuravchak, Natasa Nord & Helge Brattebø (2022): The effect of building attributes on the energy performance at a scale: an inferential analysis, Building Research & Information, DOI: 10.1080/09613218.2022.2038537

To link to this article: https://doi.org/10.1080/09613218.2022.2038537

Published online: 21 Feb 2022.

BRI BUILDING RESEARCH & INFORMATION
https://doi.org/10.1080/09613218.2022.2038537

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

The effect of building attributes on the energy performance at a scale: an inferential analysis

Ruslan Zhuravchak, Natasa Nord and Helge Brattebø

Department of Energy and Process Engineering, NTNU. Kolbjørn Hejes v. 1B, Trondheim, Norway

ABSTRACT

The commitments to mitigate the negative impacts associated with final energy use stipulate the increase of energy efficiency of the built environment. This is the focus of urban energy policies and of built stock energy models that aid them. The complexities behind the phenomenon, however, hinder the development of the means for controlling and unbiased modelling. Such tasks necessitate the empirical evidence of causal relationships between architectural and technical attributes and building energy performance at the population level. This study, therefore, elaborates on the methods of inferential statistics for establishing such causal effects. The focus is on the methods of frequentist inference, active use of which may advance the understanding of the phenomenon and foster more accurate modelling practices. The case study examines the energy performance exhibited by distinct configurations of construction periods, envelope materials, sources of energy for space heating and the ventilation system types. The empirical sample consists of more than 11,000 records registered in the Norwegian energy performance certification system. The results document the effects and their significance. These methods are applicable in any urban context and may provide the empirical basis for promoting/discouraging certain technological and architectural tendencies, and simulating the phenomena through probabilistic programming.

ARTICLE HISTORY
Received 8 October 2021
Accepted 28 January 2022

KEYWORDS
Built stock; energy performance; probabilistic programming; combinatorial analysis; frequentist inference; univariate density estimation; statistical hypothesis testing

Introduction

Energy use in buildings is seen as one of the key bottlenecks in the transition towards more sustainable cities, communities and nations. Improving the energy efficiency and energy flexibility of the built stock, therefore, is amongst the central components of urban development strategies which are being initiated, supported and/or supervised through a variety of political mechanisms (Kennedy et al., 2014; Tozer, 2020). Developing, implementing and reviewing such mechanisms rely on long-term urban energy planning that seeks to accurately predict and rationally match future energy use to the generation capacities. Building energy use at the urban level, however, is a complex phenomenon governed by multiple factors of socioeconomic, architectural, technical, environmental and other kinds. Such complexities, amplified by the heterogeneity and the continuous evolution of the built environment, entail uncertainties that undermine the plausibility of energy planning. Developing the means for analysing and predicting the phenomenon while overcoming these challenges is the subject of built stock energy modelling and has important implications for achieving sustainability targets.

The significance of stock-wide energy modelling for practical and policy-related applications is one of the reasons for elevated attention to such models in building energy research (Johari et al., 2020; Moghadam et al., 2017; Reinhart & Davila, 2016). Given a variety of objectives that build stock energy modelling may pursue, their design, methodological foundations and resource needs may differ substantially. These characteristics are the basis for the hierarchical model classification proposed by Swan and Ugursal (2009). A more recent study by Langevin et al. (2020) suggested an extended classification of models by design and several additional criteria to determine their taxonomic affiliation, namely the degree of transparency, system boundaries, spatial resolution, temporal dynamics and the approach to handle the uncertainties.

The prevailing modelling practices, although indisputably instrumental at addressing their objectives, do not rest on empirically validated causal relationships

CONTACT Ruslan Zhuravchak ruslan.zhuravchak@ntnu.no

between building attributes and the energy performance at a population level. Even if such relationships may have been established for individual buildings, projecting them to the urban scale implies the risks of biased results. The bias is caused by a virtually infinite number of aspects that, directly and indirectly, statically and dynamically, individually and jointly, affect the real energy use at a population level. Addressing such complexities represents a significant challenge. To maintain the data, labour and computational resources rational, modelling procedures often involve the assumptions and approximations in relationships. These are, ultimately, some of the key drivers of the performance gap (van den Brom et al., 2018; Menezes et al., 2012) and elevate the risks of developing irrelevant theories, making misleading conclusions and pursuing implausible or ineffective energy strategies.

Establishing the causal relationships with explicit account for variabilities in final energy use at the population level is advocated by energy epidemiology - a framework for incisive analysis and modelling proposed by Hamilton et al. (2013). Methodologically, it implies synthesizing the analytical instruments used in health sciences with already established architecture- and engineering-based foundations of building energy research. Epidemiological approaches articulate the need for robust conclusions about the direct and indirect, individual and joint causal effect of a certain factor on the phenomenon of interest. It is, for example, expected that within the built stock renovation program, insulation of building envelope leads to energy savings (direct positive effect) (Jones et al., 2013). The occupant may, however, prefer elevated indoor temperature setpoints once the renovation is done which leads to higher energy use (indirect negative effect), a.k.a. rebound effect (Guerra Santin, 2013; Hamilton, 2016). Furthermore, the joint influence of multiple factors is likely to result in 'effect modification', *e.g.* the energy savings that follow envelope insulation together with the deployment of renewable energy technologies is not equal to the sum of savings from these measures if implemented separately.

Documented evidence-based causal relationships, considering the inherent variabilities handled through the epidemiological approach, may enable to (i) acquire a better theoretical understanding of the phenomenon; (ii) improve the accuracy of modelling practices; (iii) rationalize the energy planning and the associated legislative mechanisms. Although these needs are understood, building energy research lacks the means to address them. The causalities are commonly documented by comparing the parameters of central tendency (mean, median, mode), dispersion (variance, standard deviation, interquartile range, support) and shape (skewness, kurtosis) of the empirical univariate sample distribution, as shown in van den Brom et al. (2019); Gangolells et al. (2016); Hjortling et al. (2017). Individually, neither of these parameters characterize the phenomenon through the entire range of possible values. More comprehensive metrics must be used to facilitate the conclusions about the causal relationships. Probability density function (PDF) and probability mass function (PMF) are therefore used as parsimonious representations of continuous and discrete phenomena accordingly. PDF or PMF accommodate the central tendency, dispersion and shape of the distribution, represent a statistical model capable of generating synthetic data and hence, used for carrying out simulations facilitated by the methods of probabilistic programming (Zhuravchak et al., 2021). The procedure of identifying the underlying PDF or PMF given the empirical sample is referred to as density estimation.

Documenting the properties of the population based on the available subset (empirical sample) is the objective of statistical inference, density estimation being one of its components. The focus on population makes statistical inference distinct from the descriptive analysis which is focused on the empirical sample only. The practices of statistical inference follow either of the two established paradigms: frequentist and Bayesian inference. The debates on theoretical correctness, practical benefits and the possible synthesis of frequentist and Bayesian approaches last for a century (Bayarri & Berger, 2004; Cox, 2006; Raue et al., 2013). In numerous applications dealing with knowledge discovery and modelling, the choice between these approaches is driven primarily by the objectives, considerations on the accessibility and quality of data, availability of prior information and computational resources. In the built stock energy research, density estimation is often approached using Bayesian inference, as a part of either forward or inverse uncertainty analysis procedures (Tian et al., 2018). The distributions of parameters related to architectural and operational characteristics of the buildings, for example, are documented in several studies (Booth et al., 2012; Heo et al., 2015; Tian & Choudhary, 2012; Zhao et al., 2016). The variability of typology-specific actual building energy performance is quantified by Choudhary (2012), Choudhary and Tian (2014), Braulio-Gonzalo et al. (2016). Frequentist methods are scarcely represented in the domain literature, with one example of density estimation by Fonseca and Panão (2017).

Unless the data for the entire population is collected, any conclusions about the population based on a randomly collected sample are prone to errors. That is,

there is the risk that the causal relationships apparent in the sample do not apply to the population. In such a case, projecting the sample-based analysis on the population leads to biased results. Most of the time, however, collecting the data for the whole population is irrational or impossible to carry out. To address the problems of this kind, both frequentist and Bayesian inference provide the methods of hypothesis testing (Silva, 2018) that yield a measure of confidence in any claims related to causal relationships. Statistical hypothesis testing, despite its appreciation in epidemiological studies (Rigby, 1998) and scientific practices in general (Mizrahi, 2020), is not being used systematically in building energy research. One of the likely causes is that hypothetical reasoning is amongst the most challenging statistical concepts to explain and to comprehend (Park, 2019).

This study, therefore, is motivated by the need to elaborate on some methods for determining the causal relationships in the context of energy epidemiology, with the focus on frequentist approaches (Section Methodology). Statistical hypothesis testing, in this study, is represented by Kolmogorov–Smirnov (KS) test. Density estimation - by maximum likelihood estimation (MLE) and several metrics for judging the goodness-of-fit. The case study is based on the empirical dataset, as described in Methodology section, and exemplifies (in Results section): (i) hypothesis testing procedures to find if the age, envelope material, source of energy for space heating and the type of ventilation system have direct or indirect implications on the actual energy performance of the population of apartments in Oslo, Norway; (ii) density estimation to document the variability of energy performance between apartments having distinct combinations of these attributes. This enabled identifying the combinations of attributes that exhibit relatively high and low energy performance within the scope of the case study. Discussion section points towards several practical applications of the methods discussed, evaluates the possibilities for upscaling and diversifying the scope and outlines several alternative methods of interest. A summary of findings is provided in Conclusion section. These findings are partially based on and can be reproduced/replicated through Built Stock Explorer (https://buildingstockexplorer.indecol.no) – an open access research software for knowledge discovery and modelling of the Norwegian built stock.

Methodology

Empirical data

Following the Energy Performance of Buildings Directive (EPBD) 2002/ 91/ EC, the Norwegian strategy for advancing towards low energy use in buildings is assisted by the Energy Labelling System for Houses and Dwellings (Brekke et al., 2018). One of the outputs of this system is the energy performance certificate (EPC) registry. EPCs contain the reported total annual energy use ($kWh \cdot y^{-1}$) per certified unit and its source-specific annual energy use if more than one source is used. The values are averaged over 3 years of the building's operation to account for the varying weather conditions, occupancy- and maintenance-related factors. The reported total annual energy use normalized per unit of heated floor area is a continuous random variable that reflects the actual energy performance - a reported total energy use intensity (EUI) in $kWh \cdot m^{-2} \cdot y^{-1}$.

The residential built stock in Oslo consists of apartment blocks and several typologies of houses, namely detached, semi-detached, chained, terraced and quad house. In the EPC registry, the apartment is the most frequent typology, reaching 74% by records count, 54% by heated floor area and 52% by total annual energy use amongst all the residential units. Certified apartments have four attributes relevant to this study, namely:

(1) Construction period (CP) > 1990: binary [True, False];
(2) Primary envelope material (EM) used: either of [Concrete, Brick, Wood];
(3) Source of energy for space heating (SH): either or a combination of [Electricity (El), District heating (DH), Wood, Gas, Oil, Heat pump (HP)];
(4) Ventilation system (VS) type: either of [Natural (N), Periodical (P) extraction, Continuous (C) extraction, Balanced (B)].

The construction period is seen as a proxy to architectural and envelope-related measures made to comply with the energy performance standards active during a certain period. Historically, in Norway, substantial improvements in the energy performance requirements occurred in 1990 (Sandberg et al., 2016; Sartori et al., 2009). The year 1990, therefore, is used in this study as the basis for separating the empirical sample into two groups. The other attributes reflect structural, technological and indoor environmental comfort considerations which are likewise expected to affect the actual energy performance. In addition to the direct effects triggered by the building attributes, they could be the cause and/or the result of more indirect tendencies of socio-cultural, physiological, economic and other kinds exhibited by the occupants. The inferential analysis presented in this study, therefore, examines both

direct and indirect effects associated with building attributes.

The subset of the EPC registry, limited to apartments in Oslo, with the reported total EUI and the four attributes specified, constitutes a sample of 11,163 records. These units have distinct combinations of attributes, each of which is defined explicitly in this study. The number of all possible combinations of two CPs, three EMs, nine unique combinations of energy sources for SH and four VS types is given by the cardinality (216) of the Cartesian product of these attributes. All possible combinations of these attributes are illustrated as a circular tree structure in Figure 1. A central (root) node in Figure 1 represents the entire sample (11,163 records) of apartments in Oslo. This node has two child nodes representing the subsets of apartments in Oslo constructed (i) before and (ii) after 1990. Similarly, each of these nodes has the child nodes that represent the subsets of distinct EMs, sources of energy for SH and VS types. This tree structure has 216 leaf nodes that form the outermost circle. Every distinct path from the root to the leaf node defines a unique combination of four attributes that characterize the apartments in Oslo. The diameter of the node is proportional to the size of the sample represented by this node.

Statistical hypothesis testing

Identifying the direct or indirect effects of building attributes on the energy performance involves comparing the samples of the energy performance of buildings with and without this attribute. However, since the available empirical sample is only a subset of the population, the effect observed in the data may occur by chance. Statistical hypothesis testing, therefore, answers the question: 'If the attribute does not affect the energy performance in the population, how likely is it to observe this effect in the empirical sample that represents this population?'

A formal hypothesis testing requires (i) the choice of test statistic -- the metrics that quantify the effect of interest, *e.g.* differences in population means, population proportions, etc.; (ii) the formulation of the null hypothesis, which is an initial assumption about the absence of direct or indirect effect measured by the test statistic; (iii) computing the p-value (Figure 2), which is the likelihood of observing a certain effect provided that the null hypothesis is true; (iv) judging the statistical significance of the results: small p-value suggests that the observed effect is not likely to occur by chance (implies rejecting the null hypothesis), a large p-value implies a failure to reject the null hypothesis. This decision is typically based on comparing p-value to the threshold α of statistical significance established prior to the experiment.

In Figure 2, the shaded area illustrates p-value as the likelihood of observing the values of the test statistic as large as x or more extreme. The illustration applies to two-sided tests since p-value accounts for extreme values on both sides of the distribution. Alternatively, one-sided tests can be used (outside the scope of this study).

This study tests the null hypothesis formulated as 'distinct combinations of building attributes do not affect the energy performance'. Rejecting this hypothesis is made at a significance level $\alpha = 0.05$. The test statistic and the calculation of p-value are based on the KS test (Bhattacharya et al., 2016; Feldman & Valdez-Flores, 2010). In the (two-sided) KS test, a measure D of conformity between two empirical samples is the supremum of the difference between their cumulative distribution functions (CDF) (Marsaglia et al., 2003):

$$D = \sup_x |F_1(x) - F_2(x)| \tag{1}$$

where: x, random variable; $F_1(x)$, CDF of sample being tested; $F_2(x)$, CDF of the sample against which the test is carried out.

Figure 3 illustrates the empirical CDFs of two arbitrary samples and the associated D-statistic. Because the CDF fully characterizes the central tendency and the dispersion of the empirical sample, the KS test is recognized as a comprehensive and convenient method for hypothesis testing. This is a non-parametric test, applicable to any empirical and theoretical distributions. The p-value associated with the test can be found from the asymptotic distribution of the KS test statistic.

Density estimation

Density estimation seeks to fit (and to evaluate the goodness-of-fit) a set of parameters θ that characterize the PDF $f_X(x \mid \theta)$ of the theoretical random variable X to the empirical sample $x{:}[x_1, x_2, x_3, \ldots, x_n]$. MLE (Robert & Casella, 2013; Thomopoulos, 2017) is a method for finding the parameters θ by solving a multivariate unconstrained optimization problem of maximizing the log-likelihood function that has a form:

$$\begin{aligned} L(\theta, x) &= \max_\theta \left\{ \ln \left[\prod_i^n f_X(x_i \mid \theta) \right] \right\} \\ &= \max_\theta \left\{ \sum_i^n \ln \left[f_X(x_i \mid \theta) \right] \right\}. \end{aligned} \tag{2}$$

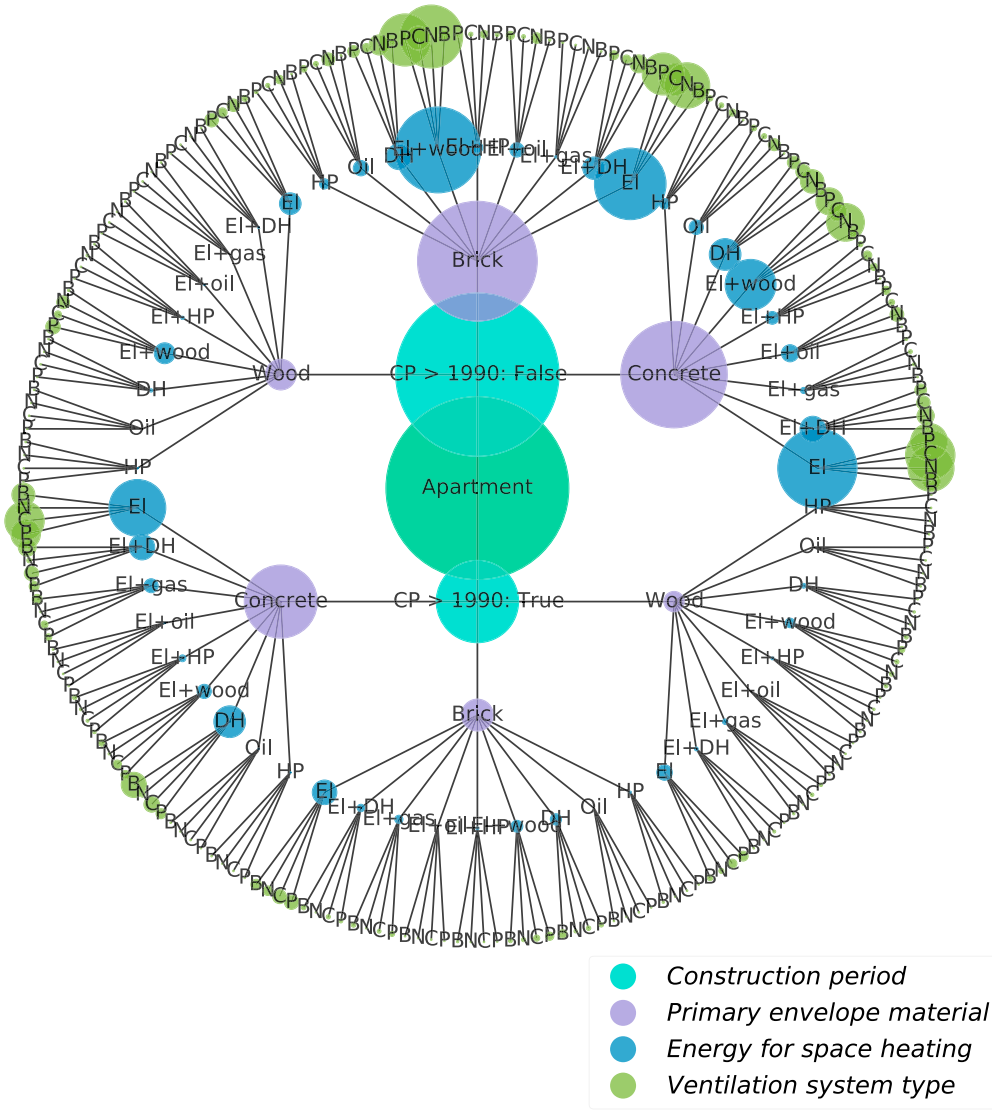


Figure 1. Building attributes structured as a circular tree.

The objective of MLE may be achieved with numerous optimization techniques, such as the downhill simplex (Nelder–Mead) method (Gao & Han, 2012; Härdle et al., 2017; Nelder & Mead, 1965), where the objective function converges by adjusting the arbitrary-selected initial parameters step-wisely. The objective is met either if the error tolerance satisfies a certain criterion or if the objective function exhibits no changes for several past iterations.

Goodness-of-fit between the theoretical parameterized distribution and the empirical data must be quantified further. A KS test, discussed above, may be adapted for this purpose, with *D*-statistic and the associated *p*-value used as reference metrics. High *p*-value suggests that the null hypothesis 'the observed sample is the outcome of the fitted PDF' cannot be rejected. The conformity between the sample density and the theoretical parameterized PDF can be also quantified

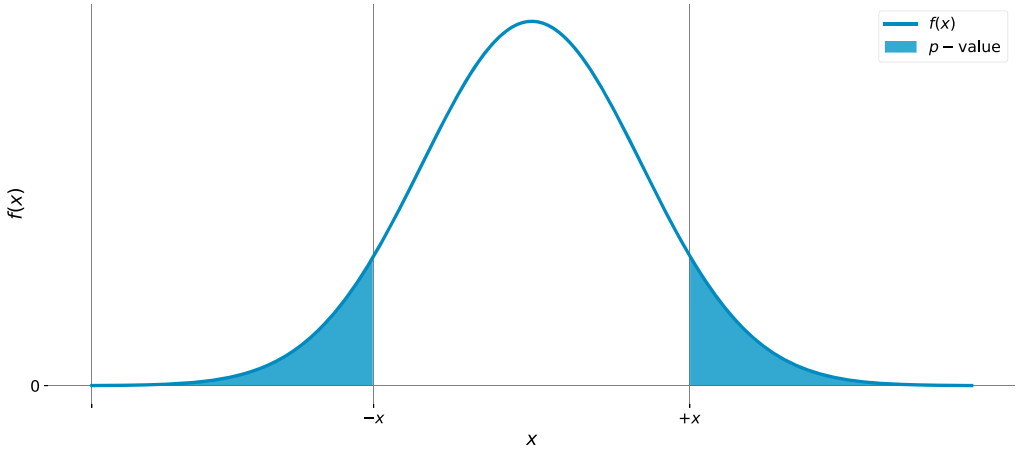


Figure 2. *p*-value under the distribution $f(x)$ of test statistic x.

with the sum of squared errors (SSE, Equation (3)), where a small SSE indicates a better fit. Another instrumental metric is the coefficient of determination R^2 (squared coefficient of correlation R) of the linear least-squares fit between the quantiles of the theoretical distribution and the ordered values of the sample. R^2 quantifies the total variation in the sample described by the variation in the theoretical quantiles. High R^2 suggests a good fit and vice versa.

$$SSE = \sum_{i=0}^{n} (y_i(x) - f_i(x))^2 \tag{3}$$

where $y_i(x)$, sample density at the ith interval; $f_i(x)$, density of a fitted PDF at the ith interval; n, number of intervals considered.

Figure 4 exemplifies the metrics for goodness-of-fit associated with fitting the exponentially modified Normal (Exponnorm) distribution to the empirical sample. The figure suggests overall conformity between the empirical sample and the parameterized distribution. The theoretical PDF (Figure 4a) approximates the density histogram, with occasional underestimated spikes compensated by overestimating the neighbouring density. This fit is associated with a small SSE. A theoretical continuous CFD (Figure 4b) follows the step function of the empirical CDF, with rare minor deviations. The corresponding D-statistic

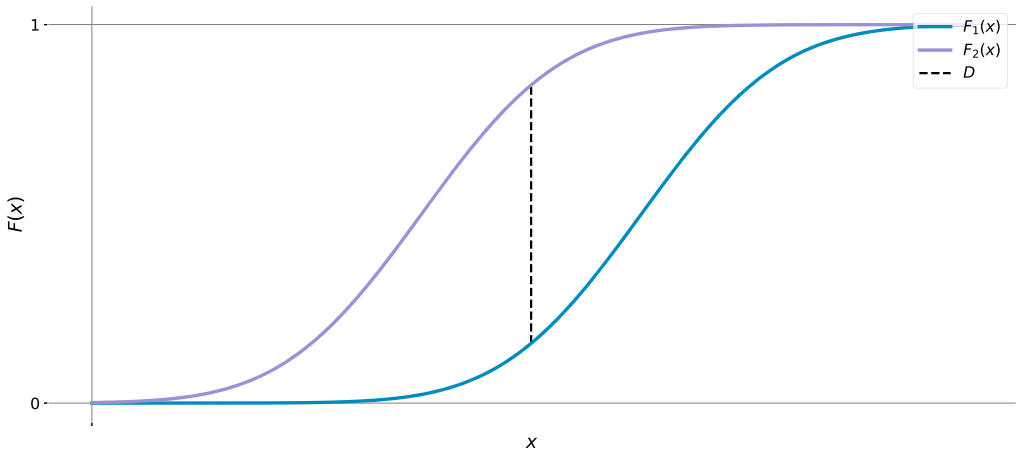


Figure 3. Empirical CDFs of two arbitrary samples and the D-statistic.

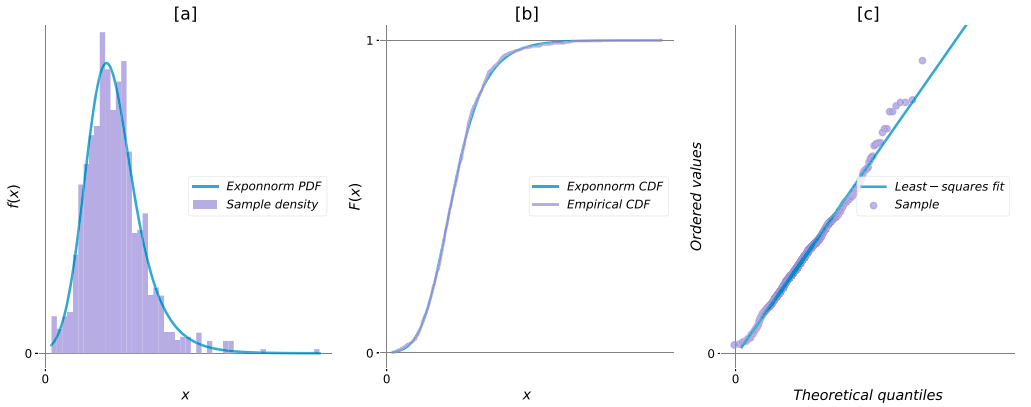


Figure 4. Goodness-of-fit between the arbitrary sample and the MLE-parameterized exponentially-modified Normal distribution: (a) sample density histogram over 100 bins with the PDF of a fitted distribution; (b) an empirical CDF with the CDF of a fitted distribution; (c) probability plot, *i.e.* quantiles in the PDF of a fitted distribution against the ordered values of the empirical sample with the linear fit.

is small, and the p-value is high. A strong positive correlation between theoretical quantiles and the empirical values can be observed (Figure 4 c), with negligible deviations from the linear fit and, therefore, high R^2. The example demonstrates the key goal of density estimation - obtaining an approximate parametric description of the data generating process. It is not possible and not attempted to assure that the observed empirical sample is generated by one distribution and not another. The p-value solely suggests that the observed D-statistic is not too rare to reject the choice of the distribution. SSE and the R^2 provide the quantitative metrics to support the choice of the distributions amongst the alternatives and to better understand the performance of the probabilistic model based on this fit.

Results

The method of statistical hypothesis testing introduced above assists with concluding if buildings characterized by distinct attributes have distinct energy performance. The difference and the statistical significance of such difference are quantified and documented for individual attributes. A significance level for hypothesis testing through this study is set to $\alpha = 0.05$. The KS test statistic enabled identifying which building configurations amongst those analysed are most- and least favourable for better energy efficiency of the built stock within the scope of the study. This section also elaborates and documents the results of density estimation per individual building configuration. The reference is made to Figure 1 whenever applicable to explain which sample(s) are considered.

Attribute-wise analysis of conformity

The first question of interest is formulated as: 'do the populations of apartments in Oslo, constructed [1] before and [2] after 1990 have a significant difference in their energy performance?'. The null- and the alternative hypothesis are set as follows:

$$H_0 : E_1 - E_2 = 0; \quad H_a : E_1 - E_2 \neq 0; \tag{4}$$

where E_1, energy use intensity (kWh $\cdot$ m^{-2} $\cdot$ y^{-1}) of the population [1]; E_2, energy use intensity (kWh $\cdot$ m^{-2} $\cdot$ y^{-1}) of the population [2].

The empirical samples used for testing correspond to all the EPC records for apartments in Oslo where [1] CP $>$ 1990: *False* and [2] CP $>$ 1990: *True*. These samples are accommodated by all the nodes in [1] upper and [2] lower semicircles accordingly in Figure 1. The results of the test are illustrated in Figure 5.

Empirical cumulative distribution functions (ECDF) in Figure 5(a) suggest that the phenomenon exhibits distinct properties in the two groups, particularly in the range $[100 \ldots 200]$ kWh $\cdot$ m^{-2} $\cdot$ y^{-1}. The energy performance of recently built apartments is evidently better. The non-diagonal elements in the colour-encoded matrix (Figure 5b) illustrate the largest absolute difference (D-statistic) found within pairwise KS testing. Obtaining this D-statistic is associated with rather small p-values illustrated as the non-diagonal elements (Figure 5c). Because the p-values obtained

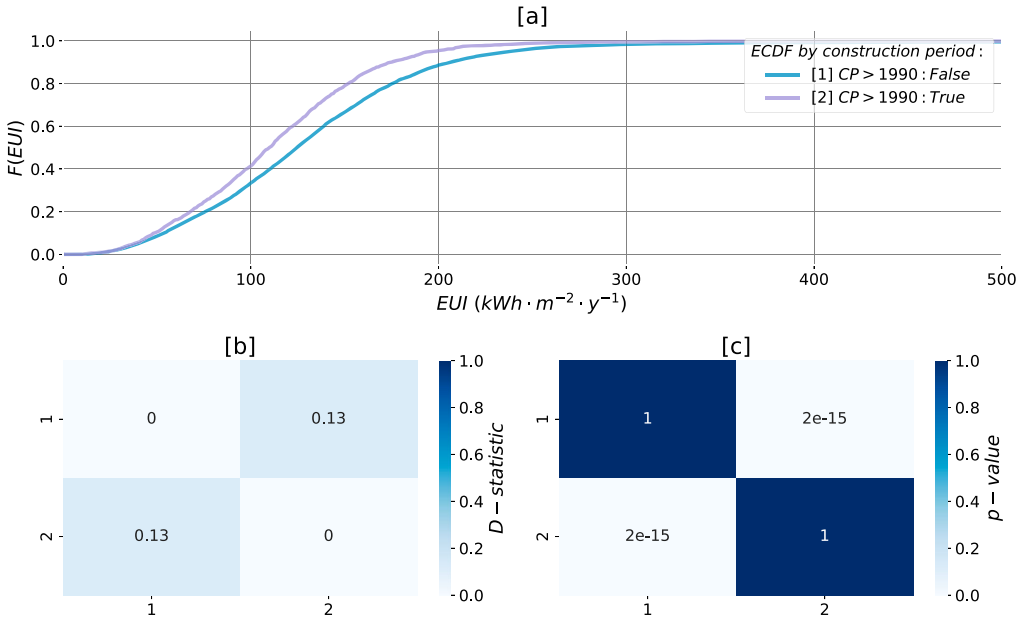


Figure 5. Energy performance of apartments in Oslo, by construction periods: (a) ECDFs of the samples; (b) *D*-statistic found with KS test; (c) *p*-values associated with the test.

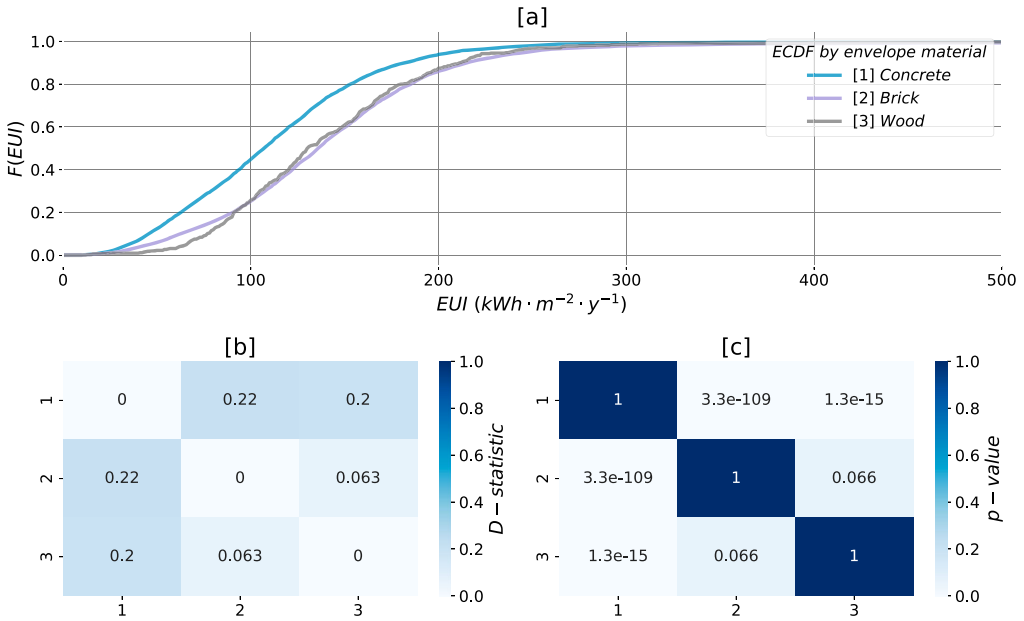


Figure 6. Energy performance of apartments in Oslo, by envelope material: (a) ECDFs of the samples; (b) *D*-statistic found with KS test; (c) *p*-values associated with the test.

through the test are substantially smaller than the established significance level $\alpha = 0.05$, the null hypothesis is rejected. With the available empirical samples, energy use intensity in the populations of apartments in Oslo constructed [1] before and [2] after 1990 is found to differ significantly.

Pairwise KS testing enables examining the implications of the other attributes on building energy performance. Figure 6 illustrates the results of KS testing to answer the next question: 'is there a significant difference in the energy use intensities between populations of apartments in Oslo that have [1] concrete, [2] bricks or [3] wood as a primary construction material in their envelope?'. All nodes matching a specific envelope material in either of construction periods (upper or lower semicircle in Figure 1) constitute the empirical samples.

Figure 6(a) suggests that apartments constructed with [1] concrete exhibit better energy performance compared to the alternatives. Energy use intensity in this population is significantly different from apartments built with bricks [2] or wood [3], which is conveyed by small p-values associated with the tests [1] – [2] and [1] – [3] accordingly in Figure 6(c). The null hypothesis adapted from Equation (4) which asserts the conformity between populations, in this case, can be rejected. The p-value returned by the test [2] and [3], however, exceeds the significance level α, meaning that the significant evidence against the null hypothesis is absent which implies a failure to reject it. It may be concluded that, given the empirical samples, the populations of apartments in Oslo constructed from [2] bricks and [3] wood do not have a significantly different energy use intensity.

Similarly to the previous attributes, Figure 7 implies rejecting or failing to reject the asserted null hypothesis for populations of apartments given the distinct sources of energy for space heating. All nodes matching a specific energy supply solution in any of construction periods and any of envelope materials (Figure 1) form the samples.

Figure 7(a) suggests mutual conformity in the distribution of total energy use intensity amongst the apartments heated by [3,4] district heating, [5,6] oil and [9] heat pump. The differences (Figure 7(b)) between their ECDFs are insignificant (Figure 7c). Energy performance of apartments with space heating solutions based on [1] electricity only and the combinations of electricity with [2] wood, [7] heat pump and [8] natural gas is significantly different from any other alternatives considered (Figure 7a–c). Electric combined with wood [1], on the one hand, and a group of solutions [3,4,5,6,9] on the other are the two extremes in the energy performance (Figure 7a). Whereas for the latter, 80% of records use less than 100 $\text{kWh} \cdot \text{m}^{-2} \cdot \text{y}^{-1}$, the share of such efficient units representing the former is only 20%.

Figure 8 enables concluding that energy use intensity amongst the populations of apartments in Oslo featuring various types of ventilation systems differs significantly. The empirical samples include all the nodes matching [1] natural, [2] periodical extraction, [3] continuous extraction or [4] balanced ventilation systems in Figure 1. Figure 8(a) suggests that the units equipped with the balanced [4] system perform better compared to the alternatives. With the empirical interpretation of Figure 8(a), 50% of such units use less than 100 $\text{kWh} \cdot \text{m}^{-2} \cdot \text{y}^{-1}$ which is followed by continuous extraction [3] (40%). The units having natural [1] and periodical [2] types appear as the least efficient, and often mutually conforming (Figure 8a). The difference between these two types, however, is found to be significant given the choice of the test statistic and the significance threshold.

The most and the least favourable configurations

The section above provides the empirical evidence that there are differences in the energy performance of populations of apartments in Oslo given the energy sources for space heating, envelope materials, vintage and ventilation systems used. These differences, provided that they are significant, tend to vary across the distinct attributes, *e.g.* the energy sources for space heating are associated with a larger D-statistic compared to the construction period. Such variability hinders the understanding of which combination of attributes that constitute building configurations perform relatively better or worse. Additionally, actual building energy performance is likely to be governed by the attributes jointly rather than individually. This section, therefore, is concerned with finding which building configurations exhibit significantly better and significantly worse energy performance relative to the entire stock of apartments in Oslo. The task requires analysing all the members in the Cartesian product of four attributes provided that the sample of minimum size is available. In this study, the minimum required sample size is set arbitrary to 20 records. 62 out of 216 possible building configurations (Figure 1) met this requirement.

Identification of the best performing configurations amongst those available involves pairwise KS testing of the corresponding sample (leaf nodes in Figure 1) versus the composite sample (root node in Figure 1) of all the apartments. A favourable configuration is associated with a large positive D-statistic between the

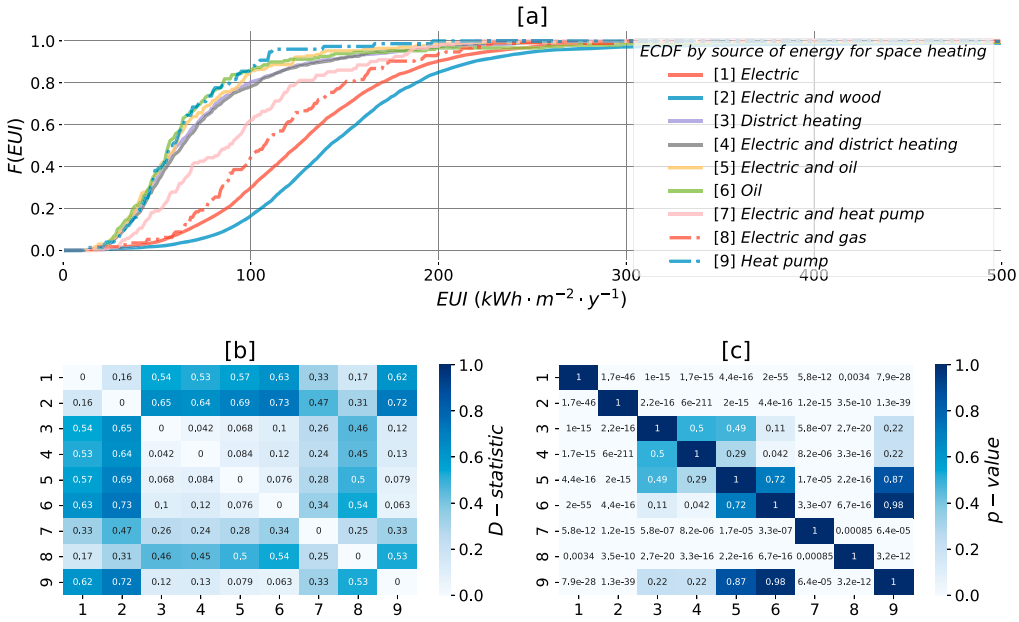


Figure 7. Energy performance of apartments in Oslo, by the source of energy for space heating: (a) ECDFs of the samples; (b) *D*-statistic found with KS test; (c) *p*-values associated with the test.

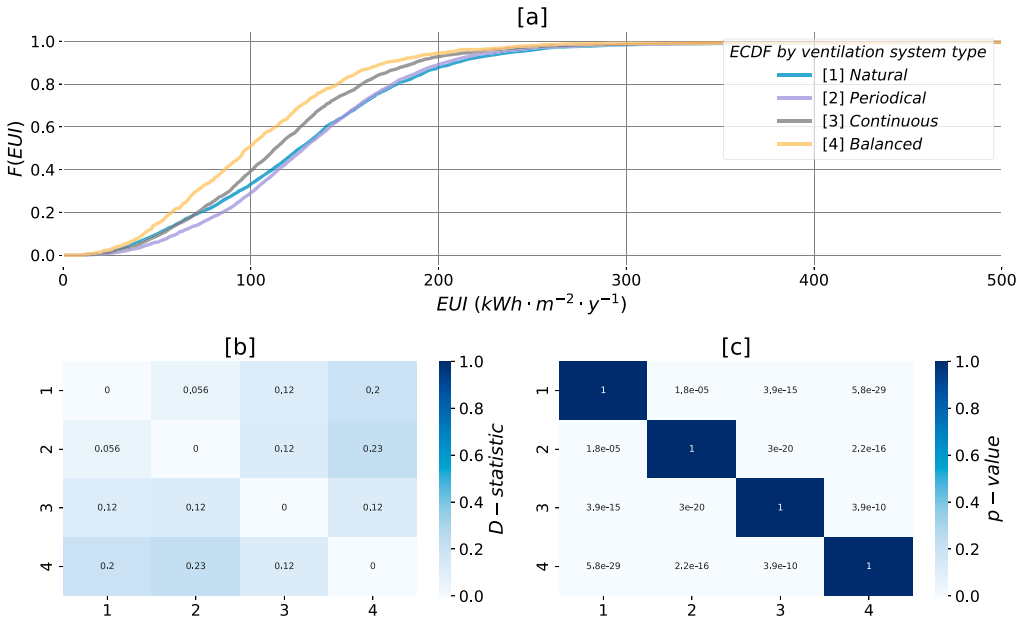


Figure 8. Energy performance of apartments in Oslo, by type of ventilation system: (a) ECDFs of the samples; (b) *D*-statistic found with KS test; (c) *p*-values associated with the test.

two ECDFs. For this difference to be considered significant, the p-value associated with the test should be below $\alpha = 0.05$. A reverse objective focused on large negative D-statistic enables the identification of the worst-performing building configurations. Figure 9 illustrates 62 ECDFs for all building configurations and highlights the best- (Figure 9a) and worst-performing (Figure 9b) compared to the composite if the difference is significant.

Figure 9 enables drawing several conclusions about the implications of building configurations on the actual energy performance. Building configurations already present in the built stock are associated with a wide spectrum of high and poor energy performance. Poorly performing apartments are typically more common (correspond to larger sample size), which governs a generally poor performance of the entire stock of apartments, as shown by the composite ECDF. To the largest extent, the distinction between the best- and poorest-performing configurations follows the distinct energy source for space heating, which, as shown earlier, reflect the largest significant differences in total energy use intensity. The least favourable configurations are featuring purely electric or electric with wood-based space heating solutions (Figure 9b). Amongst the most favourable ones, the demand for space heating is fulfilled through district heating alone or in combination with electric heaters, through heat pumps or oil (Figure 9a). The concrete- or brick-based building envelope is the most common amongst the apartments with low energy intensity (Figure 9a). The walls made of bricks, however, also appear frequently amongst the least-favourable configurations (Figure 9b). Despite having a high energy performance standard, new apartments are not common amongst those with the best energy performance (Figure 9a). Some configurations involving new apartments appear amongst the worst-performing ones (Figure 9b). Various types of the ventilation system are equally frequent amongst the least and most favourable configurations (Figure 9a,b).

Density estimation for building configurations

Varying shapes and locations of the ECDFs in Figure 9 indicate that the populations of apartments with distinct building configurations exhibit distinct statistical properties, which can be effectively characterized by the PDFs. Accounting for these properties is crucial for probabilistic simulation of the populations. This objective necessitates selecting, parameterizing and evaluating the goodness-of-fit of the PDF for each building configuration.

Parameterized distributions that are found to characterize the empirical data per individual building configuration are documented in Appendix. The distribution and the parameters represent the MLE-based best fit (the smallest D statistic) amongst the 97 distributions available in `scipy.stats` if p-value exceeds 0.05. The ranges (Min–Max) of values that the random variables are known to take are specified. The metrics used to evaluate the goodness-of-fit (D-statistic, p-value, SSE and R^2) is present in the table together with the sample size (SS) that the fit is based on.

Discussion

The development of large-scale energy efficiency strategies necessitates building energy research to establish better theoretical foundations and more accurate modelling practices. It becomes apparent that these needs cannot be addressed through architectural and engineering knowledge alone, because of the underlying complexity, magnitude, dynamics and genuinely stochastic aspects that govern the phenomenon. The instruments of inferential statistics, intended for making robust conclusions under the acute variability, uncertainty and data scarcity, are often used in population health sciences and may tackle similar challenges in building energy research. This study, therefore, agrees with (Hamilton et al., 2013) on energy epidemiology as capable to address the performance gap at the population level.

From the epidemiological point of view, three mutually related questions about the population-level energy use are essential: (i) given all possible direct or indirect effects that the technical or architectural measures may have, is there evidence of causal relationships between these measures and the phenomenon? (ii) under such complex relationships, which combinations of building attributes exhibit the most- and the least-favourable energy performance? (iii) how can this knowledge inform more accurate modelling practices? This study exemplifies the achievable answers and elaborates on some of the applicable instruments.

The results suggest that individually, the construction period, some primary envelope materials, some of the energy sources for space heating and all ventilation system types considered in this study entail a causal effect on the energy performance of apartments in Oslo. These are inferred at the significance level $\alpha = 0.05$ using the KS test. The energy policies that are targeting the building attributes that have a causal effect on the phenomenon can be used to mediate this phenomenon. The picture is less clear once more building attributes are under consideration since multiple

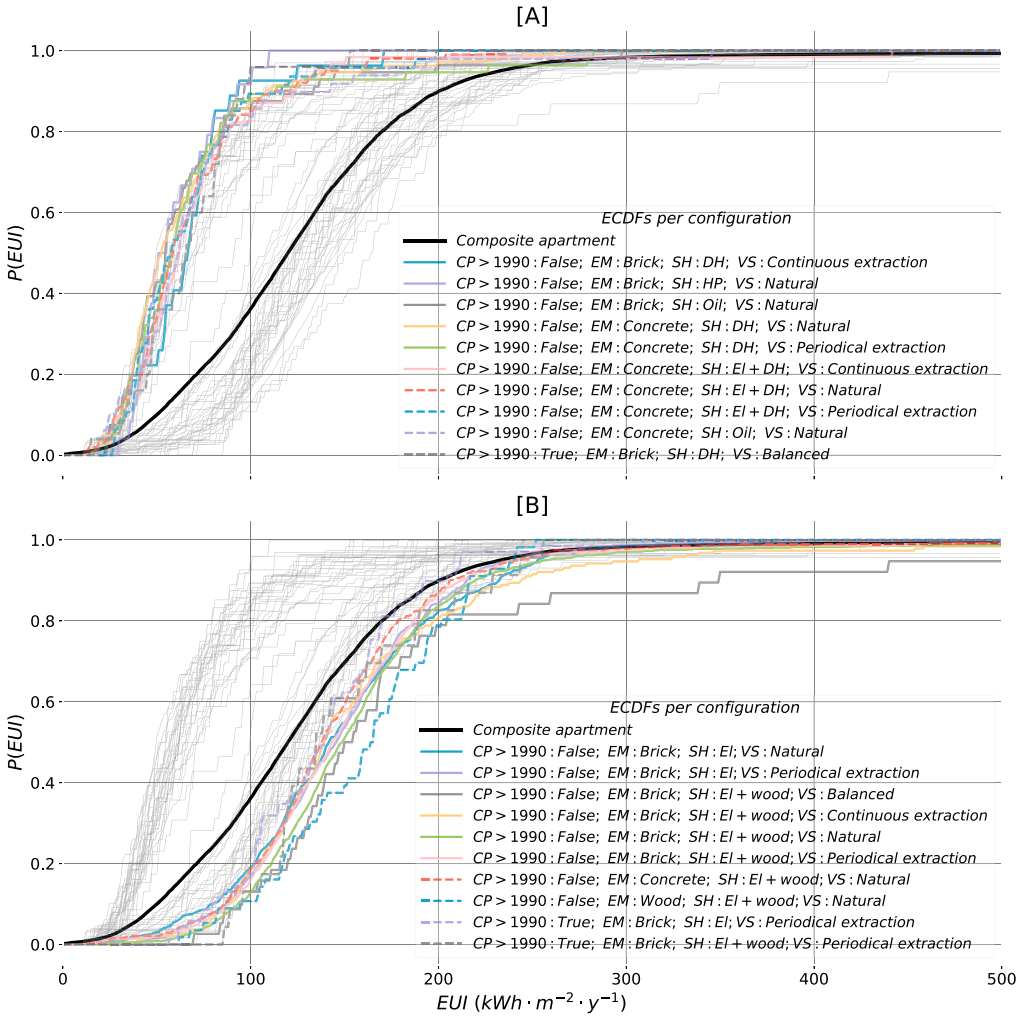


Figure 9. ECDFs for 62 building configurations within the scope of the study. Ten configurations that are found to perform considerably better (a) or worse (b) compared to the composite are highlighted.

effects are involved. Empirical evidence of a particular combination of attributes to perform better or worse compared to the alternatives is essential for energy policies that may promote or discourage certain architectural and technological tendencies in the built environment. A KS test statistic, therefore, is advocated in this study as the means for detecting the most and the least favourable combinations. An approach for structured sample analysis based on the hierarchical tree can be adapted to increasingly diversified combinations. Modelling of transformational processes in the built stock requires accommodating the underlying variability which is effectively addressed through the probabilistic framework. The inferred PDF for each combination of attributes is a parsimonious parametric approximate of the variability. The phenomenon can further be modelled as a random variable that follows its distribution. Figure 10, for example, illustrates the parameterized distributions for the most- (Figure 10a) and the least-favourable (Figure 10b) building configurations identified in the Results section. Energy policies that support the substitution of units from the least with the ones

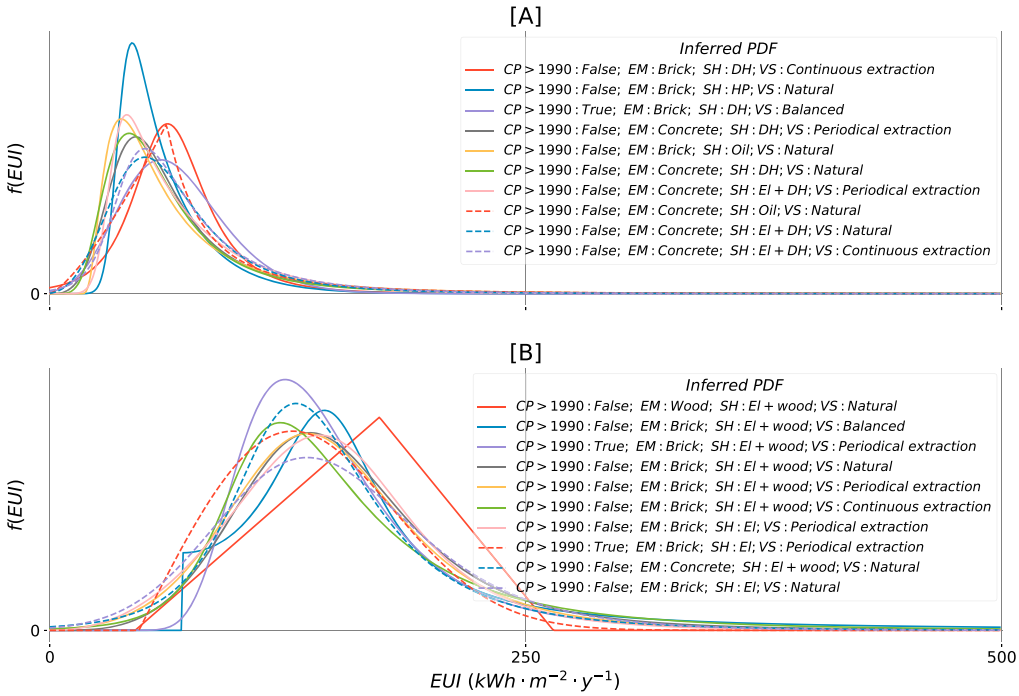


Figure 10. Theoretical distributions of the energy performance of apartments in Oslo featuring distinct building configurations: (a) with low and (b) high energy intensity, as described in Results.

from the most favourable groups is a rational step to increase the energy performance of the built stock within the shortest time. These conclusions, however, need to be justified and possibly corrected considering the expected targets, the size of populations targeted, the anticipated socioeconomic and technical constraints.

In this study, a list of 216 possible building configurations (Figure 1) is not examined exhaustively due to the absence of the sample or the limited available sample size for some of them. Because of scarce presence in the EPC registry, the anticipated size of their populations is small and therefore, of little significance to the current total energy use at the municipal level. Given the continuous transformation of the built stock, future analysis is likely to reveal other promising combinations of attributes in addition to or instead of those found in this study. Revising these analytical results systematically is also necessary for maintaining the knowledge about the actual state and the development of the built stock.

A hierarchical structure enables further up- or downscaling of the scope of the analysis carried out in this study. The level of architectural and technological detail may be extended by numerous attributes of interest. Upscaling the scope may further improve the understanding of the phenomenon across building types, geographical and national contexts. The presented case study, for example, is focused on apartments which is the largest residential building type in Oslo. A more comprehensive and complete inference for the municipality must involve other typologies, energy performance of which is known to exhibit distinct statistical properties (Figure 11).

Figure 11 suggests the presence of two distinct groups of building types within the built stock. ECDFs of residential types are evidently shifted towards zero and steeper compared to the non-residential, which implies generally higher efficiency and smaller dispersion of the former. The variation among individual building types within both residential and non-residential groups is likewise evident. Comprehensive energy policies must consider the attributes that significantly affect the energy performance of all these typologies. A similar conclusion applies to the nation-wide energy efficiency programs.

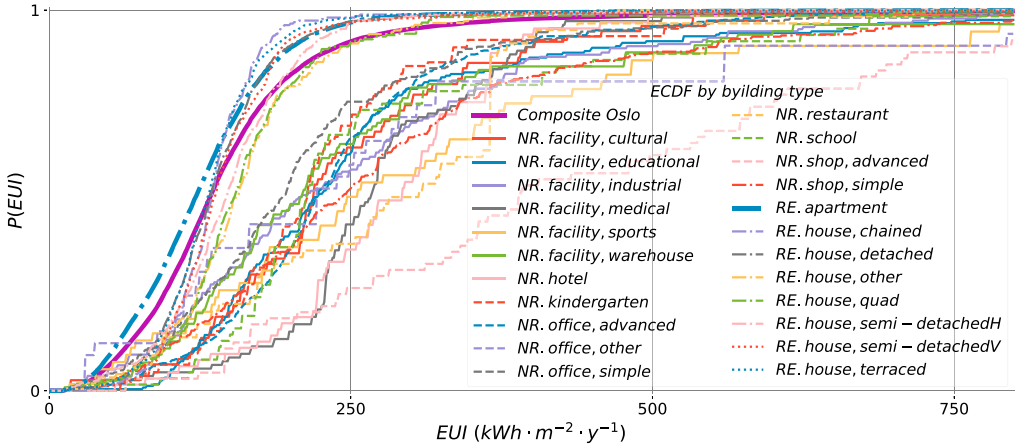


Figure 11. ECDFs for 21 building types in Oslo, illustrating the variability of energy performance in both residential (RE) and non-residential (NR) buildings.

The methods and procedures applied in this study are exemplary of a broader toolset offered by probability theory to tackle the problems alike. The analysis of skewness versus kurtosis proposed by Cullen and Frey (1999), for example, may support the choice of a theoretical continuous distribution. Probability–Probability (P–P) and Quantile–Quantile (Q–Q) plots may inform about the goodness-of-fit likewise. Alternatively to the KS test, Cramer–von Mises and Anderson–Darling methods (D'Agostino & Stephens, 1986) may be instrumental for evaluating the goodness-of-fit between the empirical sample and the theoretical continuous distribution. Examining the criteria related to penalties in the log-likelihood functions, such as Akaike and Bayesian information criteria is also a common practice to consider for this task. The present study applies MLE for parameter estimation which has alternatives, *e.g.* several variations of minimum distance estimation, moment matching estimation and quantile matching estimation. In some cases, a downhill simplex method for finding the maximum of the log-likelihood function is substituted by *e.g.* Broyden–Fletcher–Goldfarb–Shanno or conjugate gradient algorithms.

Aided by the inferential analysis, meaningful conclusions about the population based on the sample necessitate rather strict demands from experimental design and data quality (Breiman, 2001; Miller, 2014). Distortion of measurements, errors in readings, reporting and registering the data affect the reliability of conclusions further based on it. Preventing and/or mitigating the bias occurring within data collection/measurement procedures must be supplemented by responsible data management practices and the objective interpretation of findings. Concerning this study, the potential source of bias is in identifying, measuring and reporting the characteristics and the energy use in the buildings certified. Unless there is a systematic source of large error, the conclusions are expected to be valid. Minor and seldom inaccuracies in simulating the phenomenon are tolerated by the probabilistic programming.

Conclusion

High energy efficiency and flexibility are amongst the pivotal characteristics envisioned for sustainable cities and neighbourhoods. Modelling of such complex systems necessitates systematic identification and documenting of the causal relationships between building attributes and the phenomenon at the population level. This study suggests the means to obtain the empirical evidence of such relationships, if any, under an acute variability of the phenomenon. It elaborates on (i) statistical hypothesis testing to aid with concluding whether the buildings featuring certain attributes have a causal effect on the energy performance and (ii) methods for density estimation to obtain a parsimonious probabilistic representation of variability. The former is discussed and exemplified with the Kolmogorov–Smirnov test whereas the latter, in this study, is focused on maximum likelihood estimation and several metrics for goodness-of-fit. The proposed hierarchical structure enables hypothesis testing and density estimation for virtually any number of

attributes individually and in combinations. The same structure allows a more comprehensive inference with an extended list of attributes and an account for various building types, climate and administrative boundaries.

As exemplified through the case study, the combinations of building attributes already present in the built stock represent a wide spectrum, from high to poor energy performance. This leaves room for the policies to mediate them towards the achievement of energy-related and environmental targets. Practically, quantifying the variabilities given the architectural and technological configurations may provide the necessary support with setting realistic goals, identifying the bottlenecks/opportunities and screening the solutions for energy efficiency improvements.

Hence, several configurations are identified as capable of effectively improving the energy performance of the stock of apartments in Oslo. The case study likewise reveals the configurations that perform poorly and thus, considered to be the major barrier towards reducing the total energy use by apartments in the municipality. The former are typically featuring space heating solutions involving either district heating with and without electricity or oil or heat pumps and either concrete- or brick-based envelope. The latter rely on electricity alone or combined with wood to meet the demand for space heating and have either brick or concrete or wood as the main envelope material. The majority of apartments having distinctly high or low energy performance were constructed before 1990. Each configuration has unique statistical properties accommodated by the parameterized probability density function.

It is shown in this study that inferential statistics offers the essential means to improve the understanding of energy performance of the built stock, to advance the modelling approaches and thus, to safeguard the effectiveness of energy-related strategies based on these models.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The authors gratefully acknowledge the support from the Research Council of Norway (Norges Forskningsråd) through the research project Methods for Transparent Energy Planning of Urban Building Stocks – ExPOSe (project number 268248) under the EnergiX program.

ORCID

Ruslan Zhuravchak http://orcid.org/0000-0001-9462-3185
Natasa Nord http://orcid.org/0000-0003-1183-3561
Helge Brattebø http://orcid.org/0000-0001-8095-1663

References

Bayarri, M., & Berger, J. (2004). The interplay of Bayesian and frequentist analysis. *Statistical Science*, *19*(1), 58–80. https://doi.org/10.1214/088342304000000116

Bhattacharya, R., Lin, L., & Patrangenaru, V. (2016). *A course in mathematical statistics and large sample theory*. Springer.

Booth, A., Choudhary, R., & Spiegelhalter, D. (2012). Handling uncertainty in housing stock models. *Building and Environment*, *48*, 35–47. https://doi.org/10.1016/j.buildenv.2011.08.016

Braulio-Gonzalo, M., Juan, P., Bovea, M. D., & Ruá, M. J. (2016). Modelling energy efficiency performance of residential building stocks based on Bayesian statistical inference. *Environmental Modelling & Software*, *83*(4), 198–211. https://doi.org/10.1016/j.envsoft.2016.05.018

Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, *16*(3), 199–231. https://doi.org/10.1214/ss/1009213726

Brekke, T., Isachsen, O. K., & Strand, M. (2018). *EPBD implementation in Norway. Status in December 2016* (Tech. Rep.). Enova, Norwegian Water Resources and Energy Directorate (NVE), Norwegian Building Authority (DIBK).

van den Brom, P., Hansen, A. R., Gram-Hanssen, K., Meijer, A., & Visscher, H. (2019). Variances in residential heating consumption – Importance of building characteristics and occupants analysed by movers and stayers. *Applied Energy*, *250*(5–6), 713–728. https://doi.org/10.1016/j.apenergy.2019.05.078

van den Brom, P., Meijer, A., & Visscher, H. (2018). Performance gaps in energy consumption: Household groups and building characteristics. *Building Research & Information*, *46*(1), 54–70. https://doi.org/10.1080/09613218.2017.1312897

Choudhary, R. (2012). Energy analysis of the non-domestic building stock of Greater London. *Building and Environment*, *51*(8), 243–254. https://doi.org/10.1016/j.buildenv.2011.10.006

Choudhary, R., & Tian, W. (2014). Influence of district features on energy consumption in non-domestic buildings. *Building Research & Information*, *42*(1), 32–46. https://doi.org/10.1080/09613218.2014.832559

Cox, D. (2006). *Principles of statistical inference*. Cambridge University Press.

Cullen, A., & Frey, H. (1999). *Probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs*. Springer US.

D'Agostino, R., & Stephens, M. (1986). *Goodness-of-fit-techniques*. Marcel Dekker Inc.

Feldman, R., & Valdez-Flores, C. (2010). *Applied probability and stochastic processes* (2nd ed.). Springer-Verlag.

Fonseca, J., & Panão, M. J. O. (2017). Monte Carlo housing stock model to predict the energy performance indicators. *Energy and Buildings*, *152*(2), 503–515. https://doi.org/10.1016/j.enbuild.2017.07.059

Gangolells, M., Casals, M., Forcada, N., Macarulla, M., & Cuerva, E. (2016). Energy mapping of existing building stock in Spain. *Journal of Cleaner Production*, *112*(10), 3895–3904. https://doi.org/10.1016/j.jclepro.2015.05.105

Gao, F., & Han, L. (2012). Implementing the Nelder–Mead simplex algorithm with adaptive parameters.

Computational Optimization and Applications, *51*(1), 259–277. https://doi.org/10.1007/s10589-010-9329-3

Guerra Santin, O. (2013). Occupant behaviour in energy efficient dwellings: Evidence of a rebound effect. *Journal of Housing and the Built Environment*, *28*(2), 311–327. https://doi.org/10.1007/s10901-012-9297-2

Hamilton, I. G. (2016). Balancing theory with practice: Studying the rebound effect. *Building Research & Information*, *44*(8), 935–938. https://doi.org/10.1080/09613218.2016.1174909

Hamilton, I. G., Summerfield, A. J., Lowe, R., Ruyssevelt, P., Elwell, C. A., & Oreszczyn, T. (2013). Energy epidemiology: A new approach to end-use energy demand research. *Building Research & Information*, *41*(4), 482–497. https://doi.org/10.1080/09613218.2013.798142

Härdle, W., Okhrin, O., & Okhrin, Y. (2017). *Basic elements of computational statistics*. Springer International Publishing.

Heo, Y., Augenbroe, G., Graziano, D., Muehleisen, R. T., & Guzowski, L. (2015). Scalable methodology for large scale building energy improvement: Relevance of calibration in model-based retrofit analysis. *Building and Environment*, *87*(5–6), 342–350. https://doi.org/10.1016/j.buildenv.2014.12.016

Hjortling, C., Björk, F., Berg, M., & af Klintberg, T. (2017). Energy mapping of existing building stock in Sweden – analysis of data from energy performance certificates. *Energy and Buildings*, *153*(October (5)), 341–355. https://doi.org/10.1016/j.enbuild.2017.06.073

Johari, F., Peronato, G., Sadeghian, P., Zhao, X., & Widén, J. (2020). Urban building energy modeling: State of the art and future prospects. *Renewable and Sustainable Energy Reviews*, *128*(1), 109902. https://doi.org/10.1016/j.rser.2020.109902

Jones, P., Lannon, S., & Patterson, J. (2013). Retrofitting existing housing: How far, how much? *Building Research & Information*, *41*(5), 532–550. https://doi.org/10.1080/09613218.2013.807064

Kennedy, C. A., Ibrahim, N., & Hoornweg, D. (2014). Low-carbon infrastructure strategies for cities. *Nature Climate Change*, *4*(5), 343–346. https://doi.org/10.1038/nclimate2160

Langevin, J., Reyna, J., Ebrahimigharehbaghi, S., Sandberg, N., Fennell, P., Nägeli, C., & Camarasa, C. (2020). Developing a common approach for classifying building stock energy models. *Renewable and Sustainable Energy Reviews*, *133*(1), 110276. https://doi.org/10.1016/j.rser.2020.110276

Marsaglia, G., Tsang, W. W., & Wang, J. (2003). Evaluating Kolmogorov's distribution. *Journal of Statistical Software, Articles*, *8*(18), 1–4. https://doi.org/10.18637/jss.v008.i18

Menezes, A. C., Cripps, A., Bouchlaghem, D., & Buswell, R. (2012). Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap. *Applied Energy*, *97*(2), 355–364. https://doi.org/10.1016/j.apenergy.2011.11.075. Energy Solutions for a Sustainable World -- Proceedings of the Third International Conference on Applied Energy, May 16–18, 2011 -- Perugia, Italy.

Miller, T. (2014). *Modeling techniques in predictive analytics with Python and R: A guide to data science*. Pearson Education.

Mizrahi, M. (2020). Hypothesis testing in scientific practice: An empirical study. *International Studies in the Philosophy of Science*, *33*(1), 1–21. https://doi.org/10.1080/02698595.2020.1788348

Moghadam, S. T., Delmastro, C., Corgnati, S. P., & Lombardi, P. (2017). Urban energy planning procedure for sustainable development in the built environment: A review of available spatial approaches. *Journal of Cleaner Production*, *165*(S), 811–827. https://doi.org/10.1016/j.jclepro.2017.07.142

Nelder, J. A., & Mead, R. (196501). A simplex method for function minimization. *The Computer Journal*, *7*(4), 308–313. https://doi.org/10.1093/comjnl/7.4.308

Park, R. (2019). Practical teaching strategies for hypothesis testing. *The American Statistician*, *73*(3), 282–287. https://doi.org/10.1080/00031305.2018.1424034

Raue, A., Kreutz, C., Theis, F. J., & Timmer, J. (2013). Joining forces of Bayesian and frequentist methodology: A study for inference in the presence of non-identifiability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*(1984), 20110544. https://doi.org/10.1098/rsta.2011.0544

Reinhart, C. F., & Davila, C. C. (2016). Urban building energy modeling – A review of a nascent field. *Building and Environment*, *97*, 196–202. https://doi.org/10.1016/j.buildenv.2015.12.001

Rigby, A. S. (1998). Statistical methods in epidemiology: I. Statistical errors in hypothesis testing. *Disability and Rehabilitation*, *20*(4), 121–126. https://doi.org/10.3109/09638289809166071PMID: 9571378.

Robert, C., & Casella, G. (2013). *Monte Carlo statistical methods*. Springer.

Sandberg, N. H., Sartori, I., Vestrum, M. I., & Brattebø, H. (2016). Explaining the historical energy use in dwelling stocks with a segmented dynamic model: Case study of Norway 1960–2015. *Energy and Buildings*, *132*(6), 141–153. https://doi.org/10.1016/j.enbuild.2016.05.099

Sartori, I., Wachenfeldt, B. J., & Hestnes, A. G. (2009). Energy demand in the Norwegian building stock: Scenarios on potential reduction. *Energy Policy*, *37*(5), 1614–1627. https://doi.org/10.1016/j.enpol.2008.12.031

Silva, I. R. (2018). On the correspondence between frequentist and Bayesian tests. *Communications in Statistics -- Theory and Methods*, *47*(14), 3477–3487. https://doi.org/10.1080/03610926.2017.1359296

Swan, L. G., & Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, *13*(8), 1819–1835. https://doi.org/10.1016/j.rser.2008.09.033

Thomopoulos, N. T. (2017). *Statistical distributions. Applications and parameter estimates* (1st ed.). Springer International Publishing.

Tian, W., & Choudhary, R. (2012). A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in greater London. *Energy and Buildings*, *54*, 1–11. https://doi.org/10.1016/j.enbuild.2012.06.031

Tian, W., Heo, Y., [de Wilde], P., Li, Z., Yan, D., Park, C. S., & Augenbroe, G. (2018). A review of uncertainty analysis in building energy assessment. *Renewable and Sustainable Energy Reviews*, *93*, 285–301. https://doi.org/10.1016/j.rser.2018.05.029

Tozer, L. (2020). Catalyzing political momentum for the effective implementation of decarbonization for urban buildings. *Energy Policy*, *136*, 111042. https://doi.org/10.1016/j.enpol.2019.111042

Zhao, F., Lee, S. H., & Augenbroe, G. (2016). Reconstructing building stock to replicate energy consumption data. *Energy and Buildings*, *117*, 301–312. https://doi.org/10.1016/j.enbuild.2015.10.001

Zhuravchak, R., Pedrero, R. A., Nord, N., & Brattebø, H. (2021). Top-down spatially-explicit probabilistic estimation of building energy performance at a scale. *Energy and Buildings*, *238*(17), 110786. https://doi.org/10.1016/j.enbuild.2021.110786

Appendix. Sample density estimation results

Sample information (Min/Max values and sample size (SS)) per building configuration (construction period (CP), primary envelope material (EM), space heating (SH) solution and ventilation system (VS) type), best-fit distribution with its parameters and the goodness-of-fit metrics (D-statistic, p-value, sum of squared errors (SSE) and R^2).

	CP> 1990	EM	SH	VS	Min	Max	SS	Distribution	Parameters	*D*	*p*-value	SSE	R^2
1	False	Concrete	El	Natural	13.196	396.667	709	Generalized logistic	[7.09, 16.69, 42.862]	0.016	0.993	0.000075	0.998
2	False	Concrete	El	Periodical extraction	13.649	381.733	461	Power normal	[0.016, 28.979, 9.901]	0.024	0.942	0.000123	0.994
3	False	Concrete	El	Continuous extraction	10.276	453.396	819	Mielke Beta-Kappa	[2.827, 5.671, −0.618, 133.578]	0.016	0.977	0.000064	0.997
4	False	Concrete	El	Balanced	21.951	464.646	110	Von Mises (non-circular)	[1.342, 109.521, 38.188]	0.053	0.899	0.000611	0.821
5	False	Concrete	El+wood	Natural	10.314	772.000	478	Johnson SU	[−0.55, 1.261, 115.888, 48.979]	0.020	0.987	0.000034	0.922
6	False	Concrete	El+wood	Periodical extraction	24.417	638.489	247	Johnson SU	[−1.09, 1.5, 83.738, 63.089]	0.027	0.993	0.000063	0.967
7	False	Concrete	El+wood	Continuous extraction	21.739	768.116	120	Log-Laplace	[2.573, −0.344, 126.966]	0.045	0.961	0.000100	0.982
8	False	Concrete	DH	Natural	14.667	267.463	187	Alpha	[3.149, −29.211, 261.422]	0.034	0.978	0.000621	0.904
9	False	Concrete	DH	Periodical extraction	18.319	282.169	56	Burr (Type XII)	[4.163, 0.587, −0.333, 46.343]	0.065	0.961	0.002146	0.935
10	False	Concrete	DH	Continuous extraction	19.315	178.056	72	Generalized gamma	[7.79, 0.429, 17.017, 0.325]	0.057	0.961	0.005247	0.971
11	False	Concrete	DH	Balanced	20.505	154.833	22	Exponentially modified Normal	[2.848, 33.594, 12.517]	0.099	0.967	0.024961	0.974
12	False	Concrete	El+DH	Natural	10.497	236.294	101	Fisk	[3.587, −4.152, 63.456]	0.045	0.980	0.001602	0.995
13	False	Concrete	El+DH	Periodical extraction	27.000	198.630	47	Exponentially modified Normal	[6.099, 31.173, 5.896]	0.056	0.997	0.007868	0.975
14	False	Concrete	El+DH	Continuous extraction	18.304	441.778	62	Johnson SU	[−1.153, 1.217, 35.521, 23.058]	0.040	1.000	0.000598	0.811
15	False	Concrete	El+oil	Natural	14.020	212.101	64	Alpha	[3.56, −54.072, 398.62]	0.050	0.995	0.003993	0.963
16	False	Concrete	El+oil	Periodical extraction	16.250	159.746	23	Johnson SB	[1.604, 1.342, 3.66, 254.167]	0.091	0.982	0.021860	0.983
17	False	Concrete	Oil	Natural	18.519	345.041	49	Log-laplace	[2.337, −0.16, 61.442]	0.053	0.998	0.001350	0.985
18	False	Concrete	El+HP	Continuous extraction	30.469	186.339	28	Exponential power	[0.839, 30.469, 113.378]	0.095	0.943	11.111212	0.958
19	False	Brick	El	Natural	11.080	424.174	691	Exponentially modified Normal	[1.059, 105.06, 41.627]	0.019	0.964	0.000111	0.997
20	False	Brick	El	Periodical extraction	13.275	612.959	594	Mielke Beta-Kappa	[3.251, 6.387, −0.753, 171.718]	0.018	0.987	0.000039	0.984
21	False	Brick	El	Continuous extraction	11.174	550.523	349	Exponentially modified Normal	[1.362, 95.813, 38.307]	0.022	0.996	0.000085	0.982
22	False	Brick	El	Balanced	14.925	288.089	81	Johnson SU	[−0.615, 1.51, 99.518, 68.032]	0.059	0.927	0.002206	0.963
23	False	Brick	El+wood	Natural	13.322	944.698	1326	Mielke Beta-Kappa	[4.407, 5.067, −0.64, 153.924]	0.025	0.393	0.000009	0.892
24	False	Brick	El+wood	Periodical extraction	10.795	902.246	911	Burr (Type III)	[6.388, 0.824, −38.001, 189.015]	0.024	0.674	0.000012	0.880
25	False	Brick	El+wood	Continuous extraction	30.887	588.111	189	Johnson SU	[−1.065, 1.19, 96.667, 42.581]	0.034	0.977	0.000197	0.984
26	False	Brick	El+wood	Balanced	69.907	700.461	38	Folded Cauchy	[2.093, 69.907, 35.699]	0.064	0.995	0.000452	0.813
27	False	Brick	DH	Natural	10.320	258.693	136	Burr (Type XII)	[3.872, 0.606, −0.336, 50.827]	0.044	0.947	0.001036	0.872
28	False	Brick	DH	Periodical extraction	25.352	183.103	46	Pearson type III	[2.011, 70.922, 45.827]	0.063	0.988	0.006619	0.983
29	False	Brick	DH	Continuous extraction	22.785	170.886	27	Johnson SU	[−0.258, 1.27, 58.496, 26.287]	0.093	0.955	0.016185	0.928
30	False	Brick	El+DH	Natural	14.647	483.946	108	Folded Cauchy	[2.119, 14.647, 20.844]	0.059	0.823	0.000366	0.717
31	False	Brick	El+DH	Periodical extraction	25.571	368.613	27	Alpha	[2.539, −12.58, 178.077]	0.088	0.972	0.002373	0.925
32	False	Brick	El+DH	Continuous extraction	16.746	215.385	26	Folded Cauchy	[2.471, 16.746, 21.791]	0.075	0.996	0.010785	0.761
33	False	Brick	El+oil	Natural	18.271	200.000	38	Exponentially modified Normal	[5.144, 27.44, 7.419]	0.052	1.000	0.006414	0.990
34	False	Brick	El+oil	Periodical extraction	27.529	996.513	23	Johnson SU	[−0.584, 0.542, 53.14, 5.768]	0.072	0.999	0.000570	0.832
35	False	Brick	Oil	Natural	23.583	297.562	56	Exponentiated Weibull	[73.823, 0.335, 15.053, 0.387]	0.061	0.979	0.002784	0.969
36	False	Brick	HP	Natural	29.885	109.311	24	Exponentially modified Normal	[4.544, 35.848, 5.196]	0.079	0.995	0.056493	0.972
37	False	Wood	El	Natural	15.152	317.857	53	Right-skewed Gumbel	[108.094, 53.304]	0.073	0.919	0.001995	0.972
38	False	Wood	El	Periodical extraction	55.825	265.080	63	Exponentially modified Normal	[1.23, 100.668, 29.206]	0.046	0.998	0.002800	0.992
39	False	Wood	El	Continuous extraction	40.349	881.172	30	Cauchy	[144.382, 35.279]	0.111	0.817	0.000250	0.647
40	False	Wood	El+wood	Natural	61.652	251.801	56	Triangular	[0.584, 44.787, 219.684]	0.064	0.964	0.004923	0.995
41	False	Wood	El+wood	Periodical extraction	42.913	494.388	72	Burr (Type III)	[3.974, 1.063, −0.41, 132.314]	0.059	0.952	0.000774	0.968
42	True	Concrete	El	Natural	11.111	244.605	87	Generalized logistic	[0.859, 128.423, 23.045]	0.048	0.981	0.002116	0.994
43	True	Concrete	El	Periodical extraction	46.796	345.558	291	Mielke Beta-Kappa	[3.958, 6.61, −0.406, 139.505]	0.025	0.989	0.000261	0.995
44	True	Concrete	El	Continuous extraction	10.198	325.896	521	Generalized gamma	[31.658, 1.789, −302.169, 61.842]	0.022	0.959	0.000229	0.984
45	True	Concrete	El	Balanced	10.255	342.037	179	Von Mises (non-circular)	[3.589, 111.255, 73.46]	0.033	0.986	0.000513	0.966
46	True	Concrete	El+wood	Periodical extraction	54.217	320.251	26	Alpha	[5.460, −107.623, 1278.951]	0.093	0.964	0.005331	0.941
47	True	Concrete	El+wood	Continuous extraction	52.909	400.000	26	Hyperbolic secant	[128.552, 31.785]	0.094	0.958	0.002606	0.722
48	True	Concrete	DH	Periodical extraction	20.667	188.679	34	Exponentially modified Normal	[7.555, 29.519, 7.11]	0.069	0.993	0.008710	0.945
49	True	Concrete	DH	Continuous extraction	28.302	239.958	83	Weibull minimum	[1.04, 28.269, 55.724]	0.046	0.992	0.002223	0.989
50	True	Concrete	DH	Balanced	10.526	701.518	213	Exponentially modified Normal	[3.897, 32.868, 13.146]	0.038	0.899	0.000133	0.839
51	True	Concrete	El+DH	Periodical extraction	23.616	185.185	33	Kappa (3 parameters)	[5.649, 23.616, 63.862]	0.077	0.981	0.013423	0.954

52	True	Concrete	EI+DH	Continuous extraction	12.308	208.979	69	Johnson SU	[−1.535, 0.97, 34.314, 10.698]	0.073	0.826	0.002763	0.921
53	True	Concrete	EI+DH	Balanced	16.724	245.614	107	Power log-normal	[0.058, 0.101, −55.969, 82.579]	0.039	0.996	0.001665	0.997
54	True	Concrete	EI+gas	Balanced	11.774	295.484	32	Generalized normal	[0.69, 99.873, 18.118]	0.091	0.931	0.004564	0.913
55	True	Brick	EI	Natural	69.561	309.462	31	Exponentially modified Normal	[3.042, 93.153, 16.074]	0.052	1.000	0.004743	0.989
56	True	Brick	EI	Periodical extraction	56.310	277.778	69	Rice	[1.18, 46.749, 55.728]	0.059	0.956	0.003439	0.988
57	True	Brick	EI	Continuous extraction	16.947	244.903	66	Alpha	[16.943, −606.3, 12439.065]	0.052	0.989	0.002600	0.985
58	True	Brick	EI	Balanced	45.000	282.923	37	Beta prime	[15.106, 19.677, −0.303, 168.949]	0.083	0.944	0.004783	0.975
59	True	Brick	EI+wood	Periodical extraction	85.940	255.528	23	Alpha	[4.366, −39.66, 780.872]	0.086	0.990	0.013691	0.967
60	True	Brick	DH	Balanced	14.524	152.174	25	Inverted gamma	[44.633, −117.146, 8007.042]	0.083	0.989	0.023475	0.954
61	True	Wood	EI	Periodical extraction	71.930	222.796	33	Folded normal	[1.004, 71.93, 49.677]	0.062	0.999	0.013860	0.993
62	True	Wood	EI	Continuous extraction	12.802	275.000	28	Johnson SU	[−0.587, 0.899, 107.293, 21.818]	0.065	0.999	0.003854	0.953

Appendix A

Built Stock Explorer: documentation page

Built Stock Explorer 4X

Release 0.1

Jul 16, 2021

Contents

This documentation page provides the information relevant to the ongoing development of Built Stock Explorer (https://buildingstockexplorer.indecol.no) - a web application for interactive analysis and modelling of urban building energy use. Distinct sections explain the rationale behind the development, application's functional capabilities, data characteristics, methods and instruments used. The information about software licenses, developers and contributors is provided likewise.

CHAPTER 1

About

1.1 Summary

Built Stock Explorer (https://buildingstockexplorer.indecol.no) is a research software developed for interactive analysis and modelling of large-scale building energy use. Through an intuitive web interface, it makes accessible i) a comprehensive *dataset** on the energy performance of buildings and ii) several *components/instruments* for the statistical information retrieval and data-enabled modelling, both explanatory and predictive. The software is designed to support a rather heuristic (hence the Explorer) learning about the energy-related properties of the built stock and about the methods of predicting it. The functional capabilities, therefore, are featuring an immediate illustration of any analytical operation.

*** NB:** The scope of application, so as the underlying dataset, are limited to the Norwegian cities and municipalities.

1.2 Motivation

Examining and predicting building energy performance at a stock level is motivated by the need for mediating the built environment towards more sustainable. This may be done through political and practical mechanisms at all levels, from local to state-wide. The awareness about the energy performance of buildings currently, in the nearest and/or in the distant future is necessary to develop, implement and ensure the effectiveness of these mechanisms. Because of the associated complexities, scale and pace with which built stock energy use evolves, documenting its status quo and predicting the future state is a rather challenging task. The examples of relevant questions are:

1. How much energy do buildings actually use?
2. What is the relationship between building characteristics (e.g. type, age, size) and their energy use?
3. Does the energy use of the buildings differ given the local climate, architectural and socio-political variations across the communities they belong to?
4. And finally, can we accurately predict energy use in buildings, by means of either or both, classical statistical modelling or more novel machine learning techniques?

Built Stock Explorer accommodates a set of instruments to answer some of these and the related questions.

1.3 Users

The application's design choices allow anyone to use it. This could be a curious user interested in the discipline or an experienced analyst involved in research or policy development. Although the illustrated results are meant for intuitive interpretation, the understanding of basic statistical concepts and terminology is beneficial.

CHAPTER 2

Data and metadata

2.1 Source

Background data (the Norwegian Energy Performance Certificates (EPC) dataset) is a component of the Norwegian Energy Labelling System for Houses and Dwellings. Certification carried out by Enova and managed by the Norwegian Water Resources and Energy Directorate.

Currently, the available dataset contains ≈ 79000 EPC records that meet the following criteria:

1. The EPC has been issued before September 2020;
2. The EPC specifies the total (across all energy sources) annual measured/reported energy use of the certified unit;
3. The certified unit belongs to one of 74 municipalities with the total EPC records count ≥ 100;
4. Energy use/intensity is in range of sensible values for the units of given type/age/size/administrative affiliation. This means that the dataset has been subject to standard outlier removal procedures. These procedures are intended to mitigate the implications of errors made during certification on the analytical conclusions.

2.2 Metadata

The dataset contains variables/features of categorical and numerical types:

- Categorical
    1. City;
    2. Building type.
- Numerical:
    1. Construction year [CY];
    2. Heated floor area [HFA] (m^2);
    3. Total annual energy use [EU] ($kWh \cdot y^{-1}$);

4. Energy intensity [EI] ($kWh \cdot y^{-1} \cdot m^{-2}$).

2.3 Data slicing

Built Stock Explorer supports examining a comprehensive dataset in smaller subsets of interest by controlling the variables. The control of categorical variables is made available through the dropdown menus where the selection of multiple values is supported. Numerical variables are controlled using the range sliders that define the upper and the lower limits per variable. These sliders and menus are nested under the "Dataset" tab (Fig. 1). Dropdown menus support searching while typing with Latin or Norwegian alphanumeric characters. "City" menu contains a list of 74 municipalities sorted by the number of EPC records they are associated with. "Type" menu contains a list of building types that are available given the constraints specified by the "City" menu and all three sliders. A prefix "NR."/"RE." is used as a convention to denote Non-Residential and Residential building types. Range sliders have either linear (Construction year) or logarithmic ("Heated floor area (sq.m.)" and "Total energy use (kWh/year)") scales. The "Subset totals" section displays the summary of the dataset that is active given the user-defined constraints: total number of records, the sum of heated floor area and the sum of total energy use for these records.

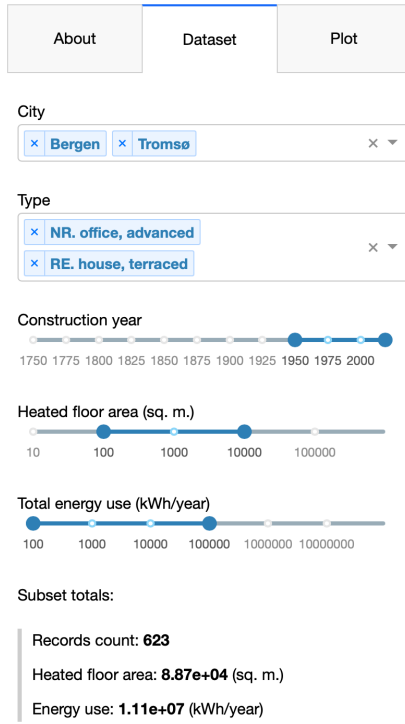


Fig. 1: Fig. 1: The components of the "Dataset" tab in Built Stock Explorer

Fig. 1 illustrates the selection of all certified terraced houses and advanced offices located in Bergen and Tromsø, constructed after 1950, having heated floor area within [100…10 000] sq.m., and using no more than 100 MWh per annum. In this example, 623 units with the total $8.87 \times 10^4\ m^2$ and $1.11 \times 10^7\ kWh \cdot y^{-1}$ match the criteria.

CHAPTER 3

Graphical components

Currently, Built Stock Explorer has two graphical components that are nested under the "Plot" tab: "Datacube" and "Distplot". They contain a collection of instruments for multi- and univariate analysis accordingly.

3.1 Datacube

Datacube (Fig. 2) is a 3-dimensional plot used to study the relationship between the six available variables. By default, three numerical variables are used as the dimensions of the cube. Adding one or two categorical variable(s) will project the data on the hyperplane: 2-dimensional (plane) or 1-dimensional (line) respectively. The same numerical variable selected for both X- and Y-axis, creates a 2-dimensional plane, passing diagonally through the X-Y space. The choice of the variable for Z-dimension is limited to either "Energy use" or "Energy intensity". Log scale, if active, applies to all numerical variables except "Construction year". Datacube has the instruments for clustering and regression model building as explained below.

3.1.1 *Clustering*

Datacube displays cluster centroids obtained using k-means clustering on the selected data subset. This unsupervised machine learning method seeks an optimal division of the subset into k clusters, with each building assumed to belong to the closest one. Cluster centroids, therefore, generalise the properties of larger groups of buildings. In Built Stock Explorer, k-means clustering is carried out per city per building type based on three numerical variables (construction year, heated floor area and energy use) for the number of clusters $1 \leq k \leq 25$. Standard feature scaling (to zero-mean and unit-variance) is applied automatically prior to clustering, and the resulting centroids are automatically reverse-scaled. The size of the cluster centroid is proportional to the number of certified units assigned to the cluster. A hover-box shows the values of the selected variables for the cluster centroid and the number of units that are assigned to it (as illustrated in Fig. 2).

Fig. 2 illustrates cluster centroids found for four distinct building types (highlighted with distinct colours) in Bergen. The advanced offices are clustered around four visible centroids (illustrated in red) that generalise/represent the size, age and energy use of the units belonging to these clusters. The coordinates of the centroids define archetype/representative units and may be used to model advanced offices in Bergen. The cluster that the cursor

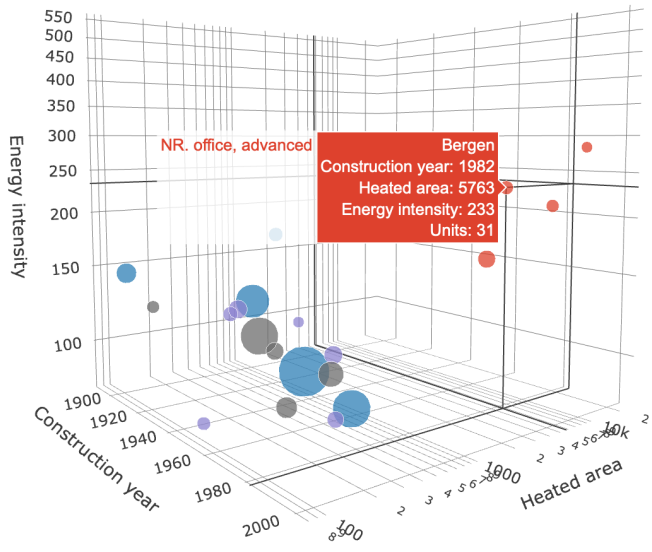


Fig. 1: Fig. 2: Cluster centroids shown in the "Datacube" component of Built Stock Explorer

points to has 31 units assigned to it, with the mean construction year: 1982, mean heated floor area of 5763 m^2 and mean energy intensity of 233 $kWh \cdot y^{-1} \cdot m^{-2}$.

NB:

The centroids of the clusters that have 3 or fewer units are not shown (a data confidentiality measure).

A new set of clusters is computed every time the user updates the subset or the number of clusters.

3.1.2 *Regression modelling*

Datacube can fit and display multiple linear regression models per city per building type within the selected subset. These models have a form: $[Z] = a_0 + a_1 \cdot [CY] + a_2 \cdot [HFA]$ where Z is the selected variable for the Z-axis, either total annual energy use (EU, $kWh \cdot y^{-1}$) or energy intensity (EI, $kWh \cdot y^{-1} \cdot m^{-2}$). CY and HFA are the construction year and heated floor area (m^2) accordingly. The intercept a_0 and slope coefficients a_1, a_2 are found through the ordinary least squares method. For each point within the input space of construction year and heated floor area, the output can be predicted using this model. A set of these outputs forms a surface that is displayed (Fig. 3). A hover-box shows the values of the selected independent variables, the model and the model's prediction through the entire input space.

Fig. 3 illustrates three regression models that may be used to predict energy intensity given the construction year and the heated floor area (m^2) for three distinct building types in Trondheim. The surface that the cursor is pointing to is associated with detached houses, for which energy intensity is governed by the model:

$$[EI] = 8.485 \times 10^2 - 3.418 \times 10^{-1}[CY] - 2.515 \times 10^{-1}[HFA]$$

Slope coefficient $a_1 = -3.418 \times 10^{-1}$ has a negative value, thus suggesting that the energy intensity inversely depends on construction year, i.e. the newer the building - the smaller the energy intensity. One of the reasons for observing this tendency is the improvement of energy efficiency standards during recent years. Also, larger houses typically utilise energy more efficiently compared to smaller ones. This explains why the heated floor area also has a negative slope coefficient ($a_2 = -2.515 \times 10^{-1}$). Using this model, a detached house of 430 m^2 HFA constructed in 1892 in Trondheim is expected to have an energy intensity of 93.72 $kWh \cdot y^{-1} \cdot m^{-2}$ (as shown in the hover-box in Fig. 3).

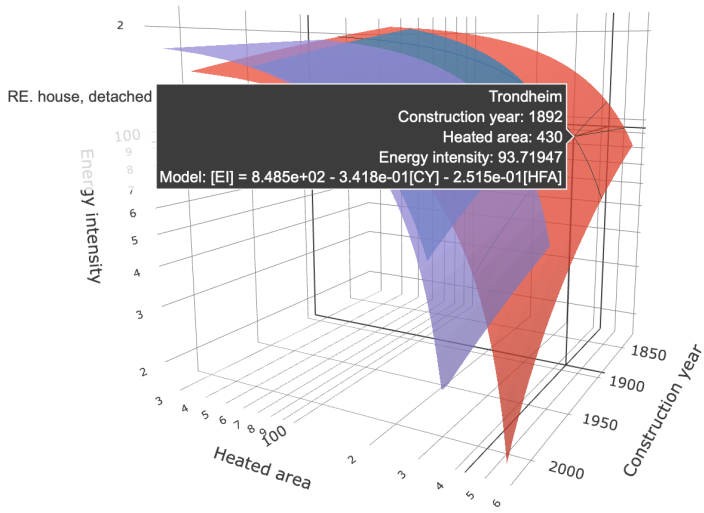


Fig. 2: Fig. 3: Linear models displayed as surfaces in the "Datacube" component of Built Stock Explorer

The fitted model can be used in the computational environment of choice to obtain the prediction given any pair (construction year : heated floor area). With python, for example, the energy intensity of a detached house constructed in 1975 in Trondheim having 200 m^2 can be predicted as follows:

```
>>> 848.5 - 0.3418*1975 - 0.2515*200
123.145
>>>
```

Or, through a python function:

```
>>> def predict_ei(cy,hfa):

        # Predict EI based on CY and HFA of a detaches house in Trondheim

        return 848.5 - 0.3418*cy - 0.2515*hfa

# Call a function with CY and HFA as arguments
>>> predict_ei(1975, 200)
123.145
# Or use numpy arrays as CY and HFA arguments to do the same for many units at once
>>> import numpy as np
>>> predict_ei(np.array([1975,2002]), np.array([200, 401]))
array([123.145,  63.421])
>>>
```

NB:

On the linear scale, a linear model always yields a rectangular plane. This plane, however, becomes curved if some variables are on a log scale (e.g. Fig. 3).

3.2 Distplot

Distplot is a histogram used to examine the selected variable in detail and consists in finding/documenting how likely are certain values to occur through the entire range of possible values. This is done based on the available subset and may be further projected on the buildings for which there is no information available. Built Stock Explorer supports these tasks through several features elaborated below.

3.2.1 *Density and cumulative density histograms*

In Distplot, the distribution of values taken by the variable can be visualised as a histogram of either density (Fig. 4) or cumulative density (Fig. 5). A density histogram displays the relative likelihood of a certain range of values to occur, whereas the cumulative density histogram indicates the probability of obtaining a value below a certain threshold. The number ($1 \leq n \leq 100$) of bins evenly spaced through the entire range of values is controlled through a slider. For the histogram to be shown - a "Histogram" checkbox must be active. The histograms are shown per city per building type. A hover-box shows the value that the variable has in the selected bin, the density (or cumulative density) and the total number of records in the selected subset.

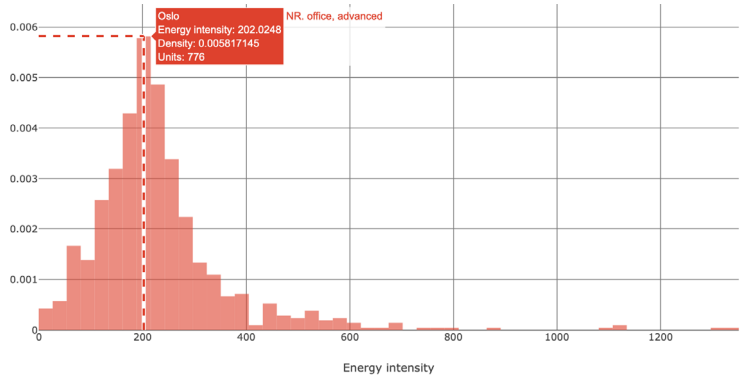


Fig. 3: Fig. 4: Density histogram in the "Distplot" component of Built Stock Explorer

Fig. 4 illustrates a density histogram of energy intensity ($kWh \cdot y^{-1} \cdot m^{-2}$) of advanced offices in Oslo. The subset consists of 776 records in the range [0...1300] $kWh \cdot y^{-1} \cdot m^{-2}$. A peak of density is at 202 $kWh \cdot y^{-1} \cdot m^{-2}$, which is the most common value. Values above 600 $kWh \cdot y^{-1} \cdot m^{-2}$ have low density, thus suggesting that the advanced offices with such high energy intensity are rather uncommon in Oslo.

Fig. 5 illustrates a cumulative density histogram for the subset shown in Fig. 4. A hover-box suggests that there is 0.381 (or 38.1%) chance that the randomly picked advanced office in Oslo will have energy intensity of 175 $kWh \cdot y^{-1} \cdot m^{-2}$ or less.

NB:

Five or more records are needed for the histogram to be shown.

The number of bins must be smaller than the number of records in the subset.

3.2.2 *Sample statistics*

Several sample statstics (Fig. 6) can be computed and shown on top of either density or cumulative density histogram. These are:

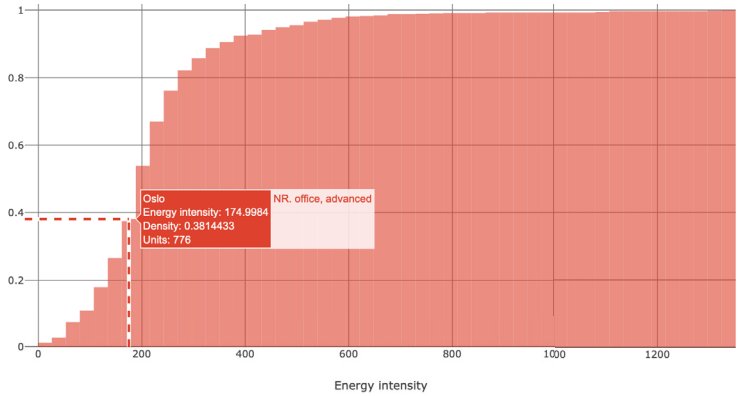


Fig. 4: Fig. 5: Cumulative density histogram in the "Distplot" component of Built Stock Explorer

- mean (average) of values;
- mode (peak of density);
- 1st, 2nd (median) and 3rd quartiles.

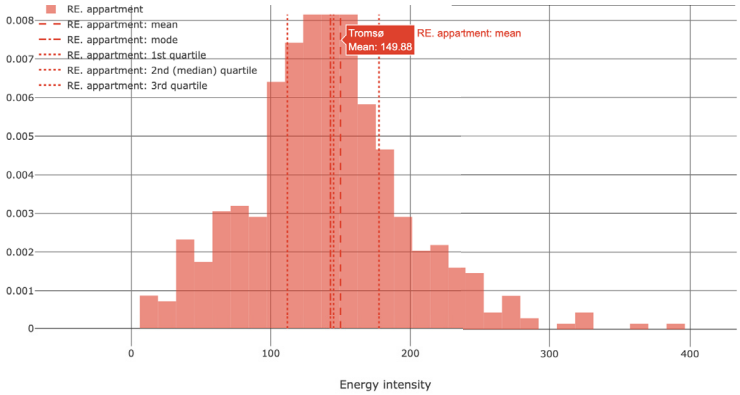


Fig. 5: Fig. 6: Sample statistics in the "Distplot" component of Built Stock Explorer

Fig. 6 shows the sample statistics computed for the energy intensity ($kWh \cdot y^{-1} \cdot m^{-2}$) for apartments constructed in Tromsø between 1976 and 2020. This subset is associated with a rather symmetric distribution where the mean, mode and median are close to 150 $kWh \cdot y^{-1} \cdot m^{-2}$. 50% of records are being observed within a rather narrow range between the 1st and the 3rd quartile (112 and 177 $kWh \cdot y^{-1} \cdot m^{-2}$ accordingly).

3.2.3 *Univariate density estimation*

The objective of parametric univariate density estimation is to find which type and parameters of the theoretical distribution describe the data. This is necessary to simulate a larger number of buildings distributed similarly to what is seen in the available data. Built Stock Explorer enables to fit, evaluate the goodness-of-fit and display the theoretical distributions for any subset (e.g. Fig. 7). A list of 95 theoretical distributions is available:

```
['alpha', 'anglit', 'arcsine', 'argus', 'beta', 'betaprime', 'bradford', 'burr',
↪'burr12', 'cauchy', 'chi', 'chi2', 'cosine', 'crystalball', 'dgamma', 'dweibull',
↪'erlang', 'expon', 'exponnorm', 'exponweib', 'exponpow', 'f', 'fatiguelife', 'fisk',
↪ 'foldcauchy', 'foldnorm', 'genlogistic', 'gennorm', 'genpareto', 'genexpon',
↪'genextreme', 'gausshyper', 'gamma', 'gengamma', 'genhalflogistic', 'gilbrat',
↪'gompertz', 'gumbel_r', 'gumbel_l', 'halfcauchy', 'halflogistic', 'halfnorm',
↪'halfgennorm', 'hypsecant', 'invgamma', 'invgauss', 'invweibull', 'johnsonsb',
↪'johnsonsu', 'kappa4', 'kappa3', 'ksone', 'kstwobign', 'laplace', 'levy', 'levy_l',
↪'logistic', 'loggamma', 'loglaplace', 'lognorm', 'lomax', 'maxwell', 'mielke',
↪'moyal', 'nakagami', 'ncx2', 'ncf', 'nct', 'norm', 'norminvgauss', 'pareto',
↪'pearson3', 'powerlaw', 'powerlognorm', 'powernorm', 'rdist', 'reciprocal',
↪'rayleigh', 'rice', 'recipinvgauss', 'semicircular', 'skewnorm', 't', 'trapz',
↪'triang', 'truncexpon', 'truncnorm', 'tukeylambda', 'uniform', 'vonmises',
↪'vonmises_line', 'wald', 'weibull_min', 'weibull_max', 'wrapcauchy']
```

This list of distributions and their naming conventions is consistent with and relies upon scipy.stats. Fitting is carried out using the maximum likelihood estimation (MLE) method. Goodness-of-fit of the theoretical distribution may be judged by examining how close its probability density function (PDF) or the cumulative distribution function (CDF) follows the density or cumulative density histograms accordingly of the subset. Also, a quantitative metric for goodness-of-fit based on the (two-sided) Kolmogorov–Smirnov test is implemented in Built Stock Explorer. The test computes $D-$ statistic and $p-$ value. $D-$ statistic is the largest absolute difference between the CDF of the fitted theoretical distribution and the empirical cumulative distribution function of the subset. Smaller $D-$ statistic suggests a better fit. The $p-$ value is a measure of how likely is it to observe $D-$ statistic that large if the data does in fact follow the fitted distribution. Practically, a $p-$ value larger than 0.05 suggests a good fit. The "Function" checkbox must be active for the distribution(s) to be shown on the histogram. A hover-box shows the value of the variable, the corresponding density, name and parameters of the theoretical distribution, $D-$ statistic and $p-$ value associated with the fit. Parameters of the distribution are listed in the format $[p_1, p_2, ...p_n, loc, scale]$, where $[p_1, p_2, ...p_n]$ are the shape parameters, if any. Location and scale parameters are always placed in the two last positions.

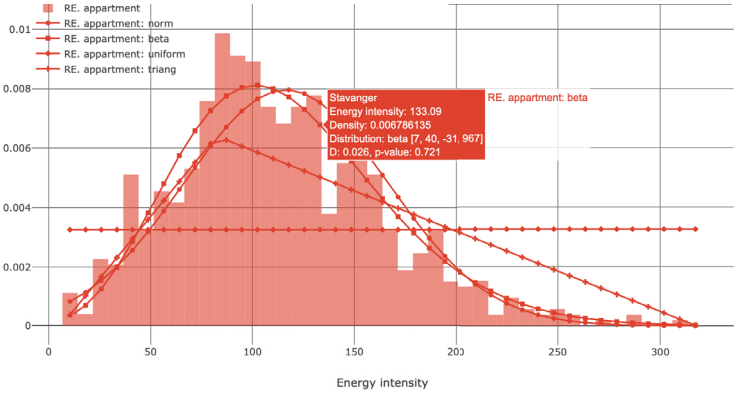


Fig. 6: Fig. 7: Fitted theoretical distributions in the "Distplot" component of Built Stock Explorer

Fig. 7 illustrates the density histogram and the PDFs of several distributions fitted to the subset of apartments in Stavanger. It is shown that beta and normal distributions are associated with a better fit compared to triangular and uniform distributions. The hover-box suggest that beta distribution with parameters [7, 40, -31, 967] has a small $D-$ statistic and a large $p-$ value. The energy intensity of a larger number of apartments in Stavanger can be simulated as a random variable that follows beta [7, 40, -31, 967] distribution. With python and scipy, this may be done as follows:

```
# Simulate 10000 apartments given the parameterised theoretical distribution
```

(continues on next page)

(continued from previous page)

```
>>> from scipy.stats import beta
>>> r = beta.rvs(7, 40, -31, 967, size=10000)
>>> r
array([ 84.3377058, 93.90328628, 153.86908053, ..., 45.10635533, 88.0250997, 194.
↪9165163 ])
# Illustrate the simulated results using a density histogram in matplotlib
>>> import matplotlib.pyplot as plt
>>> plt.figure()
>>> plt.hist(r, bins=50, density=True)
>>> plt.show()
```

which the generates the histogram shown in Fig. 8.

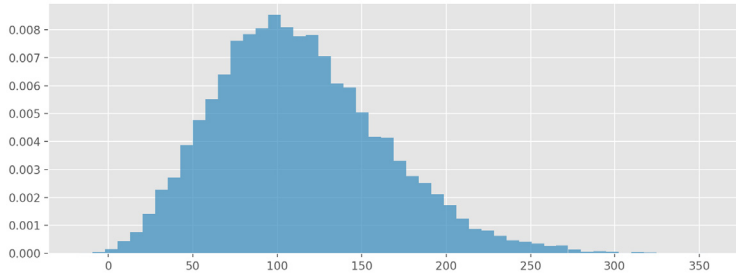


Fig. 7: Fig. 8: Density histogram of energy intensities sampled randomly from beta [7, 40, -31, 967] distribution for 10000 apartments in Stavanger

As expected, the histogram in Fig. 8 looks similar to the one that the inference is done upon (Fig. 7). Accurately replicating the distribution of the data through pseudo-random numbers is the basis of modelling/simulating built stock by means of probabilistic programming.

CHAPTER 4

Technical information

A comprehensive list of technologies, libraries and tools enabled Built Stock Explorer to be brought to the users.

It is being developed primarily in Python with some pieces of HTML and CSS scripts and the extensive use of TeX and Markdown.

The components at the interface level are produced with Dash, Plotly and Matplotlib.

Functional capabilities of the application rely on a number of Python libraries:

- Numpy for numerous algebraic operations;
- Pandas for data manipulations;
- Scipy for density estimation and computing sample statistics in Distplot;
- Sklearn for the implementation of clustering and regression modelling in Datacube.

Additionally:

- IPython and Jupyter are used as the environments for interactive development and testing of solutions;
- GitLab infrastructure made the collaborative workflow and version control possible;
- Docker facilitated many procedures related to the development and deployment;
- Sphinx and Readthedocs enabled to build and host this documentation page.

Many thanks to all the talented teams who contributed to make them available. And enabled to develop Built Stock Explorer.

CHAPTER 5

General information

5.1 Credits

Built Stock Explorer is being developed and maintained by Ruslan Zhuravchak within his doctoral work at NTNU, The Department of Energy and Process Engineering. The development is overseen by Helge Brattebø and Natasa Nord.

5.2 Contributors

Web infrastructure by Industrial Ecology Digital Lab.
Data access and contextualization by Enova.

5.3 Contacts

Contact the developer at rus.zhuravchak@gmail.com

5.4 License

Built Stock Explorer is under the BSD 3-Clause license:

```
Copyright 2021, Ruslan Zhuravchak

Redistribution and use in source and binary forms, with or without
modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this
list of conditions and the following disclaimer.
```