Missed Opportunities in Building Energy Performance Assessment

Journal Article

Author(s):
Agarwal, Minu; Cameron-Rastogi, Parag; Peronato, Giuseppe; Mavromatidis, Georgios

Publication date:
2024

Permanent link:
https://doi.org/https://doi.org/10.3929/ethz-b-000704549

Originally published in:
Journal of Sustainable Real Estate 16(1), https://doi.org/10.1080/19498276.2024.2387486

JOURNAL OF SUSTAINABLE REAL ESTATE
2024, VOL. 16, NO. 1, 2387486
https://doi.org/10.1080/19498276.2024.2387486

ESSAY

OPEN ACCESS Check for updates

Missed Opportunities in Building Energy Performance Assessment

Minu Agarwal[a], Parag Cameron-Rastogi[b], Giuseppe Peronato[c] and Georgios Mavromatidis[d,e]

[a]Faculty of Technology, CEPT University, Ahmedabad, India; [b]GRESB B.V., Glasgow, United Kingdom; [c]Energy Informatics Group, Idiap Research Institute, Martigny, Switzerland; [d]Group for Sustainability and Technology, ETH Zurich, Zurich, Switzerland; [e]Urban Energy Systems Laboratory, Empa Dübendorf, Dübendorf, Switzerland

ABSTRACT
Buildings continue to be important contributors to global greenhouse gas (GHG) emissions, despite the continued increase in stringency of energy and emissions codes and standards worldwide to address energy use in buildings. The stringency is limited by the available technology and design, and achieving even better performance may require fundamental changes in how the design and operation of buildings is regulated. In this paper we present limitations in current building energy performance assessment methods that, if addressed, could unlock unrealized energy saving potential in building projects: (a) split-incentives in multi-tenant buildings, (b) evaluation of performance using overly-long timescales, for example, yearly, (c) incorrect normalization of energy use over diverse uses and occupancy, and (d) ignoring robustness when rewarding efficiency. We discuss these issues through the lens of data availability and discuss ways of addressing them through better monitoring and a shift in accounting methods of building energy use. The solutions discussed complement the increasing use of smart meters and sensors in buildings.

POLICY RELEVANCE
Building energy standards are key policy instruments for reducing the environmental impact of buildings. The implementation of standards is often stymied by practical issues. This article highlights four technical issues and their relationship to two key non-technical issues: split incentives and ownership tenure. The article discusses possible paths to resolving these challenges.

ARTICLE HISTORY
Received 24 June 2024
Revised 23 July 2024
Accepted 25 July 2024

KEYWORDS
Building energy standards; building energy use evaluation; building energy use; robustness

1. Introduction

Buildings contribute a large portion of global anthropogenic greenhouse gas (GHG) emissions, significantly impact human health, and use substantial quantities of raw materials (Santamouris & Vasilakopoulou, 2021; Soares et al., 2017; Younger et al., 2008). This makes them important contributors as well as solutions to the linked issues of global health, pollution, and climate change. The “greening” of buildings, i.e., making them work for the planet and people, can be accomplished via voluntary uptake of systems and methods for rating and certifying sustainability (badges, ratings, recognition), legislation (policy, standards, codes), and market forces (Economidou et al., 2020; Lee & Yik, 2004; Marchi et al., 2021).

This paper examines issues with the implementation of the first two approaches – voluntary ratings and legislation – for energy efficiency and decarbonization. Specifically, the focus is on a shared aspect of these two approaches: how building performance is measured and assessed in regulatory or voluntary instruments, such as building energy standards, codes, or energy labels. In this work, “building energy standards”, or simply “standards”, will refer to any of these instruments, as part of a regulatory continuum (see Section 1.1). We discuss how shortcomings in data collection, processing, and reporting of operational energy use can hinder the achievement of desirable environmental outcomes in buildings such as decarbonization.

Multiple definitions of building energy performance assessment (BEPA) exist in the academic literature and beyond. Some definitions focus on the goal of the assessment, such as the one by Wang et al. (2012), which states that BEPA *“refers to schemes to judge the*

CONTACT Georgios Mavromatidis georgios.mavromatidis@empa.ch Urban Energy Systems Laboratory, Empa Dübendorf, Dübendorf, Switzerland.

energy performance of buildings by comparing with relevant benchmarks". Other definitions, however, adopt a wider scope. For instance, Gursel et al. (2009) define BEPA as "*the identification and quantification of the performative objective that a building is expected to satisfy, and the utilization of rigorous means for the evaluation of the building in order to assess its lifecycle performance quality*". With a similar wider scope, but more specific assessment types, Hu et al. (2021) state that BEPA "*analyses the level of the indoor thermal comfort and energy consumption of heating, cooling, ventilation, lighting, hot water, and accounts for the presence of people and domestic appliances*". BEPA, situated within some of the most used energy labels such as the Energy Performance Certificates (EPC, EU), Leadership in Energy and Environmental Design (LEED, originating in US, used worldwide) or EnergyStar (US), are based on the combined effect of energy-consuming, energy conversion systems such as boilers, electric lights and passive measures such as roof insulation. However, these measures have technological limits and may not be sufficient meet decarbonizations goals of all jurisdictions (Paoli & Cullen, 2020).

The focus of our essay is on problems and limitations connected with BEPA that can be improved with better data collection, processing, and reporting. Our essay begins with a discussion of two key non-technical issues with the implementation of effective building performance rating systems: split incentives and ownership tenure (Section 2). Especially how these two issues (split incentives and ownership tenure) change assessment of building energy use and enforcement of standards. The impact on standards (Section 3) is further broken down into three aspects: temporal granularity of assessment, robustness or resilience, and normalization of metrics (Section 4). The essay concludes with a discussion of the limitations that will affect the extent and impact of these proposals, and how these issues may be addressed together (Section 5). Sources and a brief description of each rating system or code discussed in this paper are provided in Table 1.

1.1. Scope

Policy decisions and the structure of voluntary or legislative instruments have an impact on how buildings are designed, built, and operated (see for example the study by Asdrubali et al., 2015). How performance is assessed can change the trajectory of performance outcomes, for example through the definition of scope (for example, the EU taxonomy for sustainable activities (European Commission, 2023)), which end-uses are included or excluded (for example, plug loads, space conditioning), how metrics are defined (for example, considering energy use per unit area versus per person), and how compliance is monitored and enforced (for example, automated versus manual). Standards, and the rating systems that use them can, therefore, structure and censor the information available from buildings.

Data collection and assessment practices in voluntary certification systems and legislative or regulatory

Table 1. List of building energy standards referenced in this paper.

Name	Type	Reference	Description
ASHRAE 90.1	Standard	(ASHRAE, 2022)	Normative standard on energy efficiency in non-domestic buildings
EDGE	Voluntary Certification	(IFC, 2022)	Rating scheme for evaluating the resource efficiency of buildings
EPC	Standard	(European Commission, 2018)	Energy Performance Certificates developed by the European Union and implemented by member states (incl. UK) in national legislation that rate the nominal performance of building designs
GRESB	Voluntary Certification	(GRESB, 2024)	Certification scheme for portfolios of real estate assets
KENAK	Building code	(YPEN, 2010)	Greek regulation outlining minimum performance requirements for buildings and building components, as well as other aspects like HVAC systems, use of renewable energy, lighting, and air quality, among others
LEED	Voluntary Certification	(USGBC, 2022)	A group of certification schemes originating in the USA that rate building designs, construction, and operation, owned and managed by US Green Building Council and Green Business Certification Inc. (the latter also administers WELL, EDGE, and other schemes)
NABERS	Voluntary Certification	(NABERS, 2022)	Certification scheme originating in Australia that rates buildings based solely on measured performance, regulated by Australian and international governmental and non-governmental bodies
RESET	Voluntary Certification	(GIGAbase, 2022)	Certification scheme originating in China that rates building based solely on measured performance, managed by GIGA, an independent non-governmental company
SIA 380/1	Standard	(SIA, 2016)	Swiss building norm that sets U-value requirements for components of new and retrofitted buildings and provides a methodology for building heat demand calculations, as well as target values for new and retrofitted building
WELL	Voluntary Certification	(IWBI, 2022)	Certification scheme that rates design, materials, and performance for their impact on human health and wellness

instruments (codes or standards) are examined together in this paper, since they form part of a continuum of instruments used to translate policy into engineering and environmental outcomes. Voluntary systems usually reference codes, and voluntary systems can sometimes be adopted into code, especially in the absence of centralized legislative action. For example, the cities of Atlanta and San Diego both require buildings above a certain size and value to be certified to Silver in the USGBC's LEED certification program (City of San Diego California, 2010; Holtermans & Kok, 2017). In addition, both voluntary and legislative instruments can use similar methods to rate, rank, and evaluate buildings.

With this essay, we move the discussion towards: (i) expanding the scope of performance evaluation and generating new opportunities for energy savings and decarbonization through restructuring of methods to assess building performance, usually through standards, and (ii) restructuring standards to reward efficiency under varied uses, ownership, and operating conditions. For example, a shorter timescale of evaluation incentivizes greater ongoing vigilance and a recommendation in this paper below includes a variable weekly benchmark rather than a fixed annual benchmark. Higher granularity in evaluation also creates an opportunity to reward robustness: achieving energy compliance during a "tough" week with extreme weather can be more valuable than compliance during a "mild" week. If some of the real-world complexities in building operations are considered by building energy standards, greater resilience can be realized in an asset's ability to meet its decarbonization goals (Agarwal et al., 2021).

We do not address issues around conversion of design intent and construction requirements into a rating or label, whether due to human error or willful misrepresentation. Rather, we point to potential improvements in the method and scope of evaluation. The problem statements and corresponding suggestions are meant to encourage and incentivize greater investment in Measurement and Verification (M&V) once the building is complete. The objective is a fairer post-occupancy evaluation of performance for all parties. All references to "codes" and "standards" in this paper, therefore, refer only to those related to building energy performance.

2. Impact of Ownership Structures on BEPA

2.1. Split Incentives

Split incentives refer to *"any situation where the benefits of a transaction do not accrue to the actor who pays for the transaction"* (Castellazzi et al., 2017). In a building where the building owner and users are separate financial entities for operational expenses, split incentives can become a barrier to energy efficient building design and operation. For example, if the building owner pays for the operational costs, it disincentivizes energy conscious behavior in the occupants/tenants (Gillingham et al., 2012). On the other hand, the simple act of sub-metering and consumption-based billing can result in substantial savings in residential buildings (Burak Gunay et al., 2014; Felsmann et al., 2015). These two scenarios where tenants are explicitly billed for their consumption versus buildings where there is a fixed or averaged usage costs across tenants are easy to distinguish from the onset of many projects. When BEPA methods do not address the issue of split-incentives, they miss opportunities to incentivize investment in energy-efficiency measures and energy conscious behavior.

Several BEPA methods embedded in certifications, codes and standards circumvent the issue of split incentives by mandating simple efficiency and consumption goals. Either the scope of the evaluation is curtailed to easily measurable outcomes or complexity in operations and ownership tenure is left for the building owners to resolve, or both. For example, in the NABERS rating system[1] (Table 1), in-apartment tenant energy use is excluded in multi-family residential buildings. The suite of LEED rating systems[2] includes a separate rating system – LEED Core and Shell (LEED CS) – that is devised specifically for building owners and project developers who plan to rent or sell portions of a given project and are thus responsible only for those portions of the building that they will continue to own and operate. This leads to a fragmented view of the building systems both for the tenants and the owner, leaving little incentive for a partnership in exploring energy efficiency measures. Insufficient monitoring (for example, lack of sub-metering) and the problem of split incentives tend to reinforce each other: if the impact of changes cannot be correctly allocated between stakeholders, there is less incentive for any of them to act on it.

Some building energy codes encourage, and sometimes mandate, building operators to not take a simplistic view of the operational energy use by increasing requirements for spatial and temporal granularity of performance assessment through measurement. The ASHRAE 90.1 standard (Table 1) from 2013 onward, Building Regulations Part L2 and CIBSE TM68:2022 (CIBSE, 2022) in the UK, and the Energy Efficiency Directive 2012/27/EU (EED) in the

European Union (EU), all allow and encourage sub-metering in buildings as part of performance-based compliance pathways. This means that the methods for performance assessment contained in these standards, and others, are moving from mandating only compliance at design and construction to reporting, and eventually evaluation, of building performance in-use.

In the past, to mandate such a requirement would have meant substantial cost, as electrical supply to the various sub-systems would have to be completely independent. Such technological barriers to sub-metering have been lowered significantly through readily available "smart" meters and indirect metering systems (Celenza et al., 2016). With sub-metering, consumption-based billing is made possible that can incentivize more energy conscious behavior.

However, BEPA methods currently do not go beyond sub-metering and data collection and, most importantly, no performance criteria are linked to this new information that is gathered post-occupancy. Building owners and managers are not provided additional guidance on what to do with the data collected, how to use it to improve building operational performance, and subsequently to use this data to inform the design and renovation of other buildings.

In commercial buildings, even if tenant-specific, consumption-based billing is employed, the tenant may need to further invest in additional continuous training to realize operations energy savings. For example, a tenant's attempts to save energy by turning off computers and monitors at night-time may clash with common IT practices. For example, additional personnel may be needed to ensure occupant participation in energy savings programs based on the sub-metering (Talbot & Love, 2014). Without the right kind of incentives, training programs, and simplified directives to the building end-users, the additional information on user-specific consumption may be of little value. Nevertheless, BEPA methods need to include an essential next step where they not only require sub-metering of energy use, but also link this information to the building's energy performance evaluation.

2.2. Ownership Tenures and Predicted Energy Use

Buildings can undergo change in ownership multiple time though their lifetime. Design and construction choices made during the project's inception will impact the operating program and costs of multiple entities as and when the asset changes owners and occupiers during its functional lifetime. When a building is traded, some jurisdictions enforce the exchange of an energy-based asset rating or score, for example, the energy performance certificate (EPC) mandated by UK law (Government of the United Kingdom, 2023), apart from the usual physical and financial characteristics. While asset ratings are becoming more widespread worldwide, they are not necessarily correlated to the actual energy performance a given owner or occupier can expect (Van Dronkelaar et al., 2016). That is, information regarding in-use operational consumption or expenses of a building during the time that a single entity owns or operates a building are neither widely available at the time of purchase nor conventionally estimated/measured at a temporal resolution that is meaningful to most owners.

Conventional methods of evaluating building performance are based on annualized metrics, and usually estimated based on expected performance of the asset based on standardized typical conditions, e.g., ASHRAE Standard 90.1:2022 (pg. 215, Table 12.5.1) (ASHRAE, 2022), CIBSE TM59:2017 (Ch. 4) (Bonfigli et al., 2017). Evaluating a building's actual energy consumption on an annual basis is too coarse a resolution to determine whether it aligns with its nominal rating or expected performance based on its construction or design for a specific year and owner or occupier. In other words, under the operational conditions during a given tenancy or ownership, did the building perform as designed and rated? This is especially true for ownership or tenancy tenures that are much shorter than the expected functional lifetime of a building. For example, an analysis of ownership of budget hotels in the US found that 50% of them changed ownership in 5 years or fewer (Poretti & Das, 2020). This is too short to collect a statistically significant sample of a metric that is only calculated annually.

Say an EPC for a hotel rates it as very efficient, giving a nominal consumption of x kWh/m^2/year calculated using typical weather and occupancy. During a given year, usage is y kWh/m^2/year, where $y \geq x$. Was this higher-than-rated energy expenditure due to unusual weather or occupancy, is the building operating outside its design parameters, or was the rated performance based on incorrect assumptions in the first place? Even if a rating is expressed in probabilistic terms, i.e., the 95% Confidence Interval (CI) of average annual building energy consumption is [a, b] kWh/m^2/year, 3-5 years of annual values falling outside that CI are not "unexpected". The CI gives the range within which the values of an annual metric (for example, sum) should fall 95% of the time with a very large number of samples (years of operation), which is unlikely to be seen by a single owner.

The evaluation of performance over shorter timescales is only practical if continuous post-occupancy evaluation systems, for example, the sub-circuit electrical energy use measurement required by ASHRAE Standard 90.1-2022 (Section 8.4.3), and a mechanism to evaluate the data generated are in place. Sub-metering is often written as a design requirement rather than a performance criterion, i.e., a performance assessment method may mandate sub-metering but to use that data to assess the building requires additional guidance or rules on how that data must be evaluated. However, using the collected data would provide valuable information in post-occupancy energy management, and is a necessary first step towards finer-grained evaluation of performance relevant to how buildings are often operated and owned. Like the problem of insufficient monitoring and how it disincentivizes improvements by making it more difficult to allocate costs and benefits, a lack of continuous post-occupancy evaluation hinders the evaluation of a building at different stages of its life cycle continuously. Conventional one-off evaluations at design, handover, or sale cannot be used to evaluate continuous performance or its improvement.

3. Standards – State of the Art

3.1. Prescriptive versus Performance-Based Standards

Building energy standards and codes have been successful in reducing building energy consumption globally (IEA, 2013). The first iterations of many standards and codes (and many codes still in force today) adopted a primarily prescriptive approach, focusing mainly or solely on setting minimum performance requirements for individual building components, for example, maximum acceptable U-values for building walls and windows (IEA, 2013; Rosenberg et al., 2016; Schwarz et al., 2020). While the prescriptive approach allowed for a simpler compliance path by not requiring any detailed calculations, it could not reward innovative or appropriate design or synergies from the interaction of different building components and technologies, leading to diminishing returns with increasing stringency.

As a result, newer versions of building energy standards and codes adopted performance-based rating methods by codifying one or more computational methods to rate or evaluate a building. Many standards and codes, and several certification schemes, offer a choice between prescriptive and performance-based compliance. For example, SIA 380/1 (Table 1) gives a choice between specifying compliant building components or demonstrating that the calculated overall building energy needs are below a threshold. While the achieved performance is meant to be similar for both tracks, the latter provides a greater level of flexibility to the designer. Whether wholly or partially, the rating methods used in standards, codes, or certification rely on the notion that improved operational outcomes, for example, energy efficiency, will be delivered by assets with higher nominal performance at the design stage. These rating methods typically estimate energy performance by specifying how as-designed components and systems must be evaluated under idealized conditions, to provide a combined score to the building envelope and the installed energy-consuming systems. A theoretical baseline is often used to provide a minimum performance cut-off and scores are given relative to this baseline. Overall, performance-based building energy rating is deemed more flexible than prescriptive norms (Hui, 2002; Rosenberg & Hart, 2014). However, codes have also been criticized for not being linked to the building's actual energy performance in operation – the "performance gap" (Altomonte et al., 2019; Rosenberg et al., 2016; Rosenberg & Hart, 2014) – also because of the difficulty of putting post-occupancy evaluation schemes in place.

It is worth mentioning an additional approach that combines prescriptive and performance-based elements and is proposed by Rosenberg and Hart (2014). In their study, to avoid complex simulations that would be required even for simple buildings in performance-based assessments, they propose that standards could include a list of prescriptive packages[3]. These packages would provide turnkey solutions that combine different options for both the building's envelope and its energy systems and guarantee a minimum performance level for the building. Such packages could be developed and evaluated by modeling a series of different building types and they could include options that range from cost-effective to more ambitious, albeit expensive, solutions.

3.2. The Design of Standards

Schwarz et al. (2020) point out that current energy standards often neglect building embodied energy, allow fossil-fuel heating systems, do not address low retrofitting rates, and overlook the performance gap.

While the first two criticisms question the scope and stringency of standards or codes and are outside the scope of this paper which focuses on operational energy needs, the latter two criticisms (low retrofit rates and performance gap) are related to operational energy needs and point towards structural weaknesses in the design of building energy standards. More specifically, standards and codes do not include any provisions to accelerate retrofitting rates, which currently stand at less than 1% in many countries. This rate is significant considering the existing stock of buildings worldwide. For example, it is estimated that over 90% of the existing European building stock will remain standing in 2050 (Sandberg et al., 2016).

With a more specific focus on building occupants, Gram-Hanssen et al. (2018) argue that building regulations can fail to deliver the intended energy reductions if the building occupant perspective is excluded from them. Similarly, O'Brien et al. (2020) analyzed building energy codes from 23 countries and highlighted considerable variations in the ways that occupants are represented in them. Most importantly, though, the authors have shown that in most cases occupant representation, especially in energy codes that include modeling-based performance compliance, is rather simplistic and lags scientific literature.

The effectiveness of a code does not only depend on its design, but also on its implementation. Evans et al. (2017) argue that the emphasis should be shifted from adopting more stringent requirements to supporting the implementation of existing ones instead. By examining the energy codes of 22 countries, the authors also highlight that key factors for successful implementation, compliance and enforcement are the availability of human and financial resources, as well as training programs, especially given the increasing complexity of building energy codes.

The proposals that follow highlight further issues in contemporary certification schemes, codes and standards pertaining to building energy performance measurements, evaluation, and reporting, and consider how these issues may be overcome with better data collection, processing, and reporting.

4. Addressing Technical Issues in Implementing Standards

4.1. Temporal Issues

Net zero and other performance targets enshrined in codes and standards are usually calculated for a year at a time during design, and sometimes averaged over the expected lifetime of a building (e.g., the performance targets assessed by NABERS, GRESB, LEED, and others). This has the advantage of balancing out expected variations in performance, for example, energy production and consumption, due to seasonal effects, varying occupancy, and other factors operating at sub-annual timescales. It also means that buildings can be compared using typical and uniform operating conditions. Later sections discuss how codes may be updated to account for this, for example, by checking robustness at design. In this section, we discuss how continuous evaluation for building energy standards during the lifetime of the building must change to accelerate efficiency and decarbonization outcomes. Potential changes to standards are discussed in the conclusion.

The lifetime of a building is a temporal scale that has little relevance to day-to-day building operation and needs like comfort, ventilation, and expense-management or grid carbon intensity. Without meaningful continuous feedback an operator or owner would not know that they just had a "bad" day or week of over-consumption or used particularly carbon-intensive electricity. Yet ultimately it is this short-term performance that makes up the annual metric. Finding out that a building passed or failed an annual performance target weeks or months after the fact is of little use; buildings cannot simply be "turned off" for the remainder of the year or generate more renewable energy than system capacity or weather conditions allow. Recommendations and capital improvements are helpful over the long term, but annualized metrics and infrequent feedback do not create the possibility of rapid correction and remedial action using tools like predictive maintenance and demand side management or response.

If a building fails to meet a performance target for a given year, a desirable environmental outcome over the long term would be to ask the operators to "over-perform" to compensate for the difference. The opportunities to make-up for bad periods during a year are constrained by the renewable production capacity of a building and the tradable portion of the energy demand of its occupants and processes. The option to compensate for over-consumption based on some design baseline by buying renewable credits or carbon offsets, assuming these are properly policed to have produced additional renewable electricity or removed additional units of carbon, could make buildings effective participants in decarbonization. This is not currently encouraged in standards, which instead aim to drive down overall consumption.

Before considering carbon offsetting, it is more desirable for buildings to address their over-consumption on a given day through active demand management strategies. This means that buildings should adjust their energy usage patterns in real-time, such as raising cooling set-points or implementing aggressive zone-wise load management. These measures help to balance the building's energy demand and minimize strain on the grid, as the carbon intensity and building load profiles of consecutive days are more likely to be correlated than those of similar days over several years. Waiting for the building to compensate over a subsequent year is not reliable; emissions due to building over-consumption during a daily peak one year cannot reliably be compensated by reducing consumption next year around the same day of the year because of varying and unpredictable demands and grid carbon intensity. A building may also have a different owner in the next year, who may not subscribe to the same demand management program or the same rating scheme.

Rapid assessment and feedback through performance assessment daily or at a similar timescale can, therefore, encourage owners and operators to try for better performance and compensatory actions regularly. Smaller deviations from targets can be recovered more easily through small adjustments without substantial demand-curtailment or similar drastic action detrimental to comfort or health outcomes.

While the effort to install widespread continuous monitoring is not trivial, and will be difficult to enforce, if buildings are to become meaningful actors in a more dynamic grid, they need to be able to respond to grid conditions at more granular timescales (for example, through voluntary demand-response schemes). Rating schemes and targets that require more data or analysis may impose a larger compliance burden if the data must be collected or analyzed manually. Thus, a standard that requires reporting and compliance more frequently must include mechanisms for automated evaluation, scoring, and feedback. The cost of compliance with more comprehensive, high-resolution rating frameworks can be lowered if those frameworks are able to evaluate performance without expensive specialist review.

4.2. Normalization and Categorization

Normalization and categorization of performance metrics are a common practice in building energy codes and certifications, so that values from specific buildings can be compared to benchmarks, such as building energy targets. Normalization by floor area is by far the most common way to account for the different space use in different buildings (Bracke et al., 2018). Similarly, categorization by building end-use and climate zone is typically used to provide acceptable energy targets that account for factors other than building design. Research has pointed out the flaws of current practices in normalization and suggested instead metrics based on occupancy for building performance evaluation (Li et al., 2021; O'Brien et al., 2017). Normalization and categorization cannot account for the whole range of diversity that is typically found in building occupancy patterns and design features. These diversity factors might significantly influence the final energy consumption of a given building, while they are not typically included in the benchmark metrics. In particular, some energy-efficient behaviors or design features might not be captured or even discouraged by these metrics. Hereafter we review some typical factors that are typically neglected by normalization and categorization methods.

4.2.1. Ignoring Diversity in Occupancy and Occupants

There is a large range of variety in occupancy and energy use patterns among buildings with the same end-use (Evins et al., 2015). As pointed out elsewhere in this paper, the influence of real occupancy and weather conditions can be accounted for by post-occupancy evaluations, while they are often neglected in annualized metrics used for benchmarking. This is especially true for residential buildings, whose end-use is highly dependent on the occupant behavior (Haas et al., 1998; Hong et al., 2017). Occupancy and related energy use in office spaces are also prone to variation (Chang & Hong, 2013), especially due to differences between working schedules and habits depending on the specific company. To this end, propositions have been made to use full-time equivalent occupancy for office spaces (Selvacanabady & Judd, 2017).

Another challenging category of buildings is industrial facilities, including warehouses and factories since they are highly unlikely to fit in a specific benchmark category because of the strong dependence of energy use on the type of manufacturing product and/or process. In most cases, the problem of a variety of occupancy profiles could be addressed by specifying benchmarks as ranges. However, using these ranges will likely not reward design features that contribute to small energy efficiency gains within the range, and, consequently, discourage the application or use of these measures in design or operational optimization.

The use of square meter floor area normalization can also be considered a factor contributing to hiding the fact that there is a great difference in space use (notably for housing) between different users, depending on, but not limited to, available income, age, family situation. Similarly the increase of floor area in the building stock can significantly offset the improvement of building energy performance, which might not be visible due to normalization (O'Brien et al., 2017). In this sense, what should really matter is the actual absolute energy that is used to perform a given action or meet a human need, for example, shelter. By using space-normalized metrics, the efficiency of performing this action is not considered, penalizing for example, inhabitants of a small studio in comparison to large single-family houses, in the so-called small-house penalty (O'Brien et al., 2017; Ueno, 2010).

4.2.2. Ignoring Volume

Different floor-to-ceiling heights among buildings depending on site-specific needs are also a problem for benchmarks using floor-area normalization (Bracke et al., 2018). Although minimum floor height requirements usually exist in many contexts, and these could become the standard floor height for many real estate projects striving to maximize the floor area within a plot area, there is usually no requirement of maximum floor height in a design project. This results in variation in the final building volume, which is the actual physical quantity that needs to be conditioned (i.e., cooled and/or ventilated) by the energy systems. This issue is especially true for warehouses and industrial buildings, where the specific manufacturing process highly influences the design requirements of the floor height. As a consequence, projects needing floor heights higher than the standard (minimum) values are penalized, when using floor area normalization. On the contrary, normalization by conditioned volume would provide a fairer comparison with energy targets.

4.2.3. Ignoring Ownership

The use of shared appliances among residents (for example, tumble dryers, washing machines), directly in the building or with external laundromats, can reduce energy consumption and emissions (Amasawa et al., 2018; Wasserbaur et al., 2020). This happens because their use is regulated by a schedule, which encourages the use of the machine with a full load, and/or by a pay-per-use model, which is a more transparent billing mechanism than the overall apartment electricity bill, and hence encourages a more energy-conscious attitude. Shared equipment also saves embodied energy and carbon compared to multiple individual ones. On the contrary, heat gains provided by shared equipment likely do not benefit individual units, unless the installation of such equipment is done in a living communal area or waste heat is explicitly harvested. Neither the benefits nor the costs of shared equipment are accounted for in typical benchmarks.

By enlarging the perspective of shared equipment for all living needs (for example, access to job/education and food production/retail), we might find even larger differences in operational energy (notably for mobility) between residential units, which are built within a walkable, mixed-used urban development and those in a suburban, fully residential area. In some cases, such equipment (for example, for decentralized sewage) is even enforced by local by-laws (Dewalkar & Shastri, 2021), resulting in increased energy use.

In all cases, as it was already pointed out for occupancy diversity, it would be reasonable to normalize by number of users rather than by space, to better account for all these building programming features. While it can be argued that by enlarging the analysis perimeter, building energy codes would be able to better grasp the whole energy use induced by buildings, this would go beyond the scope of this paper, which is focused on building-scale activities. Similar considerations could be conducted about the metric. In this essay, we focus on building energy needs as directly influenced by building design and occupancy, while enlarging the perimeter of the analysis to energy systems and energy sources in terms for example of final energy, primary energy, or $CO_{2eq.}$ emissions would allow for a more comprehensive analysis, yet it is out the scope of this work.

4.3. Resilience and Robustness

As discussed in Section 3.1, most performance-based building standards involve an evaluation of a building's performance under nominal conditions that pertain to typical occupant presence, activities, and preferences (for example, thermostat settings, ventilation needs, etc.) as well as to the typical weather conditions that the building will encounter over its decades-long lifetime (Cecconi et al., 2017; Crawley & Barnaby, 2019; Tahmasebi and Mahdavi, 2017). The evaluation most commonly involves a calculation of building performance that can range from the use of simplified methods and templates to advanced dynamic building simulation tools (Pan et al., 2023; Wang et al., 2012). The results of this calculation, commonly summarized in metrics like floor area-normalized energy demand, are then compared to benchmark values to test for compliance or to provide an asset rating. Simplicity is a key motivation behind this practice, as it allows for the evaluation and comparison of alternative building and system designs with minimal computational effort.

This practice, however, entails two important limitations. First, the nominal evaluation conditions, that are meant to represent "average" operating conditions for all buildings of a specific type, correspond to stylized information. For inherently stochastic elements, such as weather, and occupant presence and activities, these nominal conditions in most cases fail to accurately represent any one building's actual operating conditions (Mavromatidis et al., 2018). Second, by evaluating a building's performance only under a single set of conditions, insufficient data is available to assess the building's performance over a lifetime of environmental variability or assess whether it will remain efficient, comfortable, and healthy when its operating conditions inevitably deviate from the nominal ones. These limitations can be overcome by incorporating two concepts into building design evaluation - resilience and robustness - that improve the estimated environmental performance and impact of buildings.

Building resilience has been defined as the ability of a building to withstand during and after an extreme event (Moazami et al., 2019). Thus, in the context of this study, we define resilience as the ability of a building to maintain a comfortable and healthy environment for its occupants under most operating conditions. Some standards already address this aspect. For instance, some standards (for example, ASHRAE design days or conditions) and codes (for example, the Greek Regulation for the Energy Performance of Buildings (KENAK), Table 1) require the use of defined heating and cooling design days in order to dimension a building's Heating Ventilation and Air Conditioning (HVAC) system. This practice aims to ensure that HVAC systems can maintain the desired indoor temperature and humidity under more extreme conditions than the nominal ones. These conditions are typically defined as percentiles of the coldest and warmest possible times in the building's geographical region. Nevertheless, this practice only partly addresses the issue of resilience, as it overlooks the influence of building occupants. As a result, even if a building is assumed to be resilient to extreme weather according to a norm, depending on the occupants' behaviors, preferences, and activities, the building's resilience might be reduced.

The concept of robustness (Chung, 2011) is built upon and extends the concept of resilience. More specifically, we define a building as robust when it is not only capable of maintaining a comfortable environment for its occupants across a large range of operating conditions (resilient building), but also of achieving that while maintaining high levels of energy efficiency. In order to properly evaluate a building's robustness, its performance would have to be examined across multiple moderate and extreme scenarios for both occupancy and weather. By collecting information on how a building fares across a wide range of scenarios, its mean energy performance could still be evaluated against a benchmark, similarly to today's practices; however, the building's range of performance, quantified, for example, in terms of standard deviation, would also provide information regarding the building's robustness. Due to the complexity of such an investigation, the concept of robustness is not included in most of today's codes and certification systems.

Even though most building standards overlook the concepts of resilience and robustness, researchers have investigated them extensively. On the one hand, many studies have presented methods and tools to represent stochastic occupant behaviors (for example, Carlucci et al., 2020; Happle et al., 2018; Hong et al., 2017; Yan et al., 2015) and to generate stochastic weather data that also reflect the impacts of climate change (Goffart et al., 2017; Guan, 2009; Herrera et al., 2017; Moazami et al., 2019; Pernigotto et al., 2020; Rastogi, 2016). With regards to the latter, methods are able to provide future weather data that reflect the "average" changes to the climate, but no in the literature is yet able to generate realistic extreme weather time series for building energy analysis. On the other hand, multiple studies have investigated how building energy performance varies due to both stochastic occupant behaviors and weather. For instance, Picard et al. (2020) investigated the impact of weather variability, climate change, and occupant behavior on Zero-Net-Energy (ZNE) homes and identified that, first, a ZNE-designed house may not achieve ZNE targets in real operation, and, second, that occupant behavior exerts the highest influence among the aspects examined. Kotireddy et al. (2019) also investigated the impact of occupant behavior and weather on building performance and compared different methods to assess building robustness. Finally, extensive projects (Annexes) by the International Energy Agency's (IEA) "Energy in Buildings and Communities (EBC)" program like Annex 66 (IEA, 2018) and Annex 79 (IEA, 2022) have been dedicated to studying occupant behavior in buildings and developing mathematical approaches to represent them in models and performance standards.

By integrating resilience and robustness for improved building performance evaluation during the design phase, modern building standards can encourage design features that improve a building's performance under varying operating conditions. By evaluating a building's performance across a wide range of conditions that reflect typical and extreme real-world building operation, building owners and operators can gain access to statistical information regarding the distribution of the building's energy performance, which they can then use to benchmark the building's actual energy use. This development can function in conjunction with the concepts of post-occupancy evaluation and building performance under real occupancy and weather conditions discussed earlier in this paper.

5. Conclusions

This paper has described four issues with how building energy standards are implemented. These issues relate to the metrics, methods, and timelines that make up the rating system in these codes. The issues are discussed in terms of their impact on desirable environmental outcomes starting from the design stage through operation. We end with a discussion of how codes, and the rating systems that they use, could evolve to incorporate the use of more and different data to address the challenges outlined above. While the issue of gathering more and better data can be delegated to private actors such as building owners, technology providers, and system integrators, most of the issues highlighted in this paper rely on action by regulatory and standard-setting bodies.

Ongoing work by the authors aims to demonstrate the application of these proposals to realistic case studies. Future research into these metrics should include collaboration with building owners and policy specialists to examine the practicality of implementing these proposals in future policy instruments for building energy use management. Finally, a systematic review and comparative analysis of studies that explored the various issues highlighted in this paper would complement this paper and offer a more comprehensive perspective on these issues.

5.1. Recommendations

5.1.1. Split-Incentives

Split incentives between building owners and occupants could be addressed by restructuring performance metrics to make them more occupant-centric. Tenants may not always want to participate in owners' initiatives towards energy efficient operation schemes because the building gets certified and not the tenant. The tenant may feel that paying a price premium for a certified green property is sufficient participation in the collective societal goal of reducing energy and carbon emissions. The narrative around building energy codes could therefore be shifted such that even transient occupants feel the need to participate in control of energy consuming systems in the spaces they use. For example, the 2000-Watt Society (2000-Watt Society, 2024) concept supports the idea that it is not the buildings and other infrastructure that have an energy allowance but rather the occupants/users. The tenants do not rent a building with a certain energy use intensity (EUI), but rather increase their chances of remaining under their specific limit of consumption by leasing a more efficient space. This shifting of narrative must be done carefully to not also shift the burden of responsibility to fall entirely on individual tenants: while individual consumer choices can make a difference by sending signals to producers, consumers are able to choose only the least bad option on offer from building owners and developers. Even so, concepts like the 2000-Watt Society are more appropriate for large scale developments such as neighborhoods or cities to comprehensively cover each person's demand for energy. At the building level, sub-metering could still unlock avenues for greater participation from tenants through, for example, helping to allocate baseline consumption and quantify or explain the impact of efficiency measures on usage.

Some other ways to address split incentives include: (i) more stringent requirements in energy efficiency measures for buildings where the tenants do not pay for the energy use of space directly (Maruejols & Young, 2011); (ii) mandatory disclosure of efficiency measures, which have been shown to partially alleviate the agency problem and improve uptake of green properties by tenants (Melvin, 2018); (iii) mandating energy efficiency features in codes instead of leaving it to the discretion of individual owners and operators (Melvin, 2018); and (iv) introducing end-use specific performance benchmarks (in addition to total building energy based benchmarks).

5.1.2. Temporal Issues

To address the mismatch between annual targets and daily operational demands, standards that provide feedback, metrics, and guidance on shorter timescales are required. A building that is responsive to carbon and pricing signals from the grid and able to adjust loads flexibly is more valuable as a decarbonization actor. The gross efficiency gains promoted by current standards are helpful in lowering overall demand but a building that can adjust its demands in response to grid signals is more valuable. An early manifestation of this is the mismatch in grids with high renewable capacity today, for example, California, USA. If solar production is high in the daytime, but demand peaks just after sunset, a building that can shift its demand from that peak will avoid using electricity generated by gas peaker plants or neighboring states' grids with fossil fuels.

5.1.3. Normalization

The concept of normalization is perhaps intrinsically linked to that of benchmarking and cannot be avoided when creating widely applicable benchmarks and rating systems, although better efforts should be made to account for the diversity of the ownership, usage, space requirements, etc. Updating the normalization to use volume would solve some issues due to floor-area normalization only, while using number of users/ inhabitants as a denominator, as already proposed in the literature (Li et al., 2021; O'Brien et al., 2017), would better highlight the actual building energy efficiency, which is also a space-use efficiency. Evaluating it more often, and for shorter time periods, could enable better feedback and remediation when performance goes out of bounds.

5.1.4. Resilience and Robustness

Resilience and robustness can be integrated in today's codes and standards in non-disruptive ways to the current processes. We argue that a resilience check should be included as a minimum requirement wherever building energy standards are used. This would include the evaluation of a building's performance across a range of operating conditions pertaining to occupancy and weather/climate. This would ensure that the building can maintain comfort for its occupants and allow for better sizing of the increasingly complex building energy systems that combine conversion and storage technologies. Beyond that, robustness criteria that ensure a building can maintain a certain level of efficiency across the whole range of possible operating conditions could be integrated in more ambitious certification schemes that are awarded to buildings that meet the highest ecological standards.

Taking the same ideas one step further would include using the power of computing and online platforms, such as those proposed by Clarke (2019). Given that most norms and schemes involve some type of computational model in the evaluation process, a designer could easily upload a candidate design into an online evaluation platform, which could test the building's operation across a standardized range of predefined operating conditions or even perform Monte Carlo simulations. Results could be then returned within a matter of minutes or hours (depending on the complexity of the simulation model), which the designer can use to assess the range of the building's performance and promote the more robust options.

5.2. Limitations

Providing more technical information, at greater granularity, and/or restructuring information, as proposed in this paper, cannot solve all aspects of problems related to incentives, agency, and stakeholder conflict. While ease of collecting and accessing technical information can make compliance easier, the cost will always be more than collecting no information. Additionally, issues that cannot necessarily be addressed by better monitoring, normalization, or checking for robustness include social, economic, and political issues like the cost of compliance, cost of enforcement, unwillingness to be transparent about performance, and split incentives.

The requirement for more information from buildings is likely to create resistance at first, as building owners, operators, and even designers are apprehensive about the value of continuous measurement and "what the data will show". While from a technical perspective, these proposals look like incremental improvements and are based on existing literature, enforcing rules such as compulsory monitoring or novel normalizations could be revolutionary for most buildings in a world where even energy ratings based on steady-state models built from asset surveys are only enforced in a handful of jurisdictions (for example, Energy Performance Certificates (EPC) used in the EU and UK).

It is prudent for policy to favor simple, enforceable rules, even at the cost of reduced impact or comprehensiveness. Simplicity can improve policy acceptance, uptake, and enforcement. Some of the proposals discussed in this paper will increase the complexity of evaluation, such as by requiring sub-metering or the assessment of robustness and resilience. Rules governing resource or material efficiency will usually favor the reduction of a single measurable or calculable metric such as Energy Use Intensity (energy used per square meter of building area), since standards targeting broader goals such as healthy or climate-positive buildings are complex to design and costly to comply with. A system like EDGE (Table 1), with its few criteria and simplified calculations, can produce comparable (but possibly simplistic) ratings for buildings across the world without much input, whereas a rating system like WELL (Table 1), with its comprehensive scores addressing multiple aspects of building design and operation, requires considerably more effort, time, and money (but analyzes building performance in detail).

Complex standards and rating systems may also hinder comparison and broader benchmarking, as multiple options and pathways are available to buildings. This is not necessarily a drawback, since establishing desired outcomes while allowing flexibility in meeting those outcomes may encourage innovation and wider applicability. Complexity also opens up the possibility for gaming systems, though this is not a unique weakness of the proposals discussed here since most systems are open to cheating and gaming (Pollitt, 2013). Complex compliance or rating systems with several innovation pathways are also suitable for multilateral and international institutions where the interests and needs of several stakeholders must be balanced at once. While unnecessary complexity is to be avoided, simplicity is not a guarantee of scalability since data must still be collected to feed even the simplest metrics.

5.3. Other Instruments

While they are the focus of this paper, codes, legislation, or norms are not the only way to fix the issues described here and achieve desirable environmental outcomes. Tools like market pressures, for example, demand- or peak-based pricing by utilities, demand from investors to decarbonize, and voluntary or reputational measures, for example, demonstrating leadership, public boycotts or pressure from civil society, can also shape the built environment. Better data collection is a tool that can be used to improve and enhance the effect of social or political decisions, whether they are legal or voluntary.

Acknowledgments

GP acknowledges funding from EU Horizon 2020 research and innovation program under grant agreement N°884161, in the framework of the EnerMaps project (www.enermaps.eu).

Author Contributions

MA wrote the sections dealing with split incentives, PCR with temporal issues, GM with resilience and robustness, and GP with normalization. All authors reviewed the overall paper and contributed to common sections like introduction, state of the art, recommendations, and limitations. The authors jointly conceived the paper.

Disclosure statement

The authors declare that there are no relevant financial or non-financial competing interests to report.

Notes

1. NABERS (National Australian Built Environment Rating System) is a performance-based rating system for buildings, which measures, among others, their environmental impact, including energy efficiency, water usage, waste management, and indoor environment quality (NABERS, 2022).
2. LEED (Leadership in Energy and Environmental Design) is a globally recognized green building certification system that provides a framework for healthy, efficient, and sustainable building design, construction, and operation (USGBC, 2022).
3. The packages referred to predefined sets of building retrofit solutions that combine various elements related to the building's envelope (such as insulation, windows, and walls) and its energy systems (like HVAC, lighting, and renewable energy sources). These packages are designed to offer turnkey solutions that ensure a minimum performance level for buildings. They are created to simplify the retrofit process by avoiding the need for complex simulations, instead providing a range of options from cost-effective to more ambitious and expensive solutions.

ORCID

Parag Cameron-Rastogi http://orcid.org/0000-0002-9026-8080
Giuseppe Peronato http://orcid.org/0000-0002-2992-355X
Georgios Mavromatidis http://orcid.org/0000-0003-0227-4518

References

2000-Watt Society. (2024). 2000-Watt society and 2000-Watt Site. Retrieved from https://www.2000watt.swiss/

Agarwal, M., Rastogi, P., & Mavromatidis, G. (2021). Reducing reliance on deterministic asset ratings to improve the operational energy performance of buildings.

In *Proceedings of building simulation 2021: 17th conference of IBPSA* (pp. 3342–3349). IBPSA.

Altomonte, S., Schiavon, S., Kent, M. G., & Brager, G. (2019). Indoor environmental quality and occupant satisfaction in green-certified buildings. *Building Research & Information*, *47*(3), 255–274. https://doi.org/10.1080/09613218.2018.1383715

Amasawa, E., Suzuki, Y., Moon, D., Nakatani, J., Sugiyama, H., & Hirao, M. (2018). Designing interventions for behavioral shifts toward product sharing: The case of laundry activities in Japan. *Sustainability*, *10*(8), 2687. https://doi.org/10.3390/su10082687

Asdrubali, F., Baldinelli, G., Bianchi, F., & Sambuco, S. (2015). A comparison between environmental sustainability rating systems LEED and ITACA for residential buildings. *Building and Environment*, *86*, 98–108. https://doi.org/10.1016/j.buildenv.2015.01.001

ASHRAE. (2022). Standard 90.1-2022 – Energy Standard for Buildings Except Low-Rise Residential Buildings. American Society of Heating, Refrigerating and Air-conditioning Engineers (ASHRAE). Retrieved from https://www.ashrae.org/technical-resources/bookstore/standard-90-1

Bonfigli, C., Chorafa, M., Diamond, S., Eliades, C., Mylona, A., Taylor, B., & Virk, D. (2017). *TM59: Design methodology for the assessment of overheating risk in homes. TM 59*. The Chartered Institution of Building Service Engineers.

Bracke, W., Delghust, M., Laverge, J., & Janssens, A. (2018). Building energy performance: Sphere area as a fair normalization concept. *Building Research & Information*, *47*(5), 549–566. https://doi.org/10.1080/09613218.2018.1467703

Burak Gunay, H., O'Brien, W., Beausoleil-Morrison, I., & Perna, A. (2014). On the behavioral effects of residential electricity submetering in a heating season. *Building and Environment*, *81*, 396–403. https://doi.org/10.1016/j.buildenv.2014.07.020

Carlucci, S., De Simone, M., Firth, S. K., Kjærgaard, M. B., Markovic, R., Saiedur Rahaman, M., Khalid Annaqeeb, M., Biandrate, S., Das, A., Dziedzic, J. W., Fajilla, G., Favero, M., Ferrando, M., Hahn, J., Han, M., Peng, Y., Salim, F., Schlüter, A., & van Treeck, A. (2020). Modeling occupant behavior in buildings. *Building and Environment*, *174*(May), 106768. https://doi.org/10.1016/j.buildenv.2020.106768

Castellazzi, L., Bertoldi, P., & Economidou, M. (2017). *Overcoming the split incentive barrier in the building sector: Unlocking the energy efficiency potential in the rental & multifamily sectors. EUR 28058 EN JRC101251*. Joint Research Office, European Commission. https://data.europa.eu/doi/10.2790/912494

Cecconi, F. R., Manfren, M., Tagliabue, L. C., Ciribini, A. L. C., & De Angelis, E. (2017). Probabilistic behavioral modeling in building performance simulation: A Monte Carlo approach. *Energy and Buildings*, *148*, 128–141. https://doi.org/10.1016/j.enbuild.2017.05.013

Celenza, L., Dell'Isola, M., Ficco, G., Greco, M., & Grimaldi, M. (2016). Economic and technical feasibility of metering and sub-metering systems for heat accounting. *International Journal of Energy Economics and Policy*, *6*(3), 581–587.

Chang, W.-K., & Hong, T. (2013). Statistical analysis and modeling of occupancy patterns in open-plan offices using measured lighting-switch data. *Building Simulation*, *6*(1), 23–32. https://doi.org/10.1007/s12273-013-0106-y

Chung, W. (2011). Review of building energy-use performance benchmarking methodologies. *Applied Energy*, *88*(5), 1470–1479. https://doi.org/10.1016/j.apenergy.2010.11.022

CIBSE. (2022). TM68: Monitoring Indoor Environmental Quality. Technical Manual. London, UK: The Chartered Institution of Building Service Engineers. Retrieved from https://www.cibse.org/knowledge-research/knowledge-portal/tm68-indoor-environmental-quality-2022

City of San Diego California. (2010). Sustainable building policy. https://docs.sandiego.gov/councilpolicies/cpd_900-14.pdf

Clarke, J. A. (2019). A simulation-based procedure for the holistic resilience testing of building performance. *IOP Conference Series: Earth and Environmental Science*, *329*(October), 012027. https://doi.org/10.1088/1755-1315/329/1/012027

Crawley, D. B., & Barnaby, C. S. (2019). Weather and climate in building performance simulation. In *Building performance simulation for design and operation* (pp. 191–220). Routledge.

Dewalkar, S. V., & Shastri, S. S. (2021). An innovative approach for wastewater management using staged onsite wastewater treatment system positioned on environmental floors of multi-storey building. In H. Yu (Ed.), *Challenging issues on environment and earth science* (Vol. 4, pp. 19–31). Book Publisher International (a part of SCIENCEDOMAIN International). https://doi.org/10.9734/bpi/ciees/v4/2221F

Economidou, M., Todeschi, V., Bertoldi, P., D'Agostino, D., Zangheri, P., & Castellazzi, L. (2020). Review of 50 years of EU energy efficiency policies for buildings. *Energy and Buildings*, *225*, 110322. https://doi.org/10.1016/j.enbuild.2020.110322

European Commission. (2018). Energy performance of buildings directive. Retrieved from https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en

European Commission. (2023). EU taxonomy for sustainable activities. Retrieved from https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance/eu-taxonomy-sustainable-activities_en

Evans, M., Roshchanka, V., & Graham, P. (2017). An international survey of building energy codes and their implementation. *Journal of Cleaner Production*, *158*(August), 382–389. https://doi.org/10.1016/j.jclepro.2017.01.007

Evins, R., Orehounig, K., & Dorer, V. (2015). Variability between domestic buildings: The impact on energy use. *Journal of Building Performance Simulation*, *9*(2), 162–175. https://doi.org/10.1080/19401493.2015.1006526

Felsmann, C., Schmidt, J., & Mróz, T. (2015). *Effects of consumption-based billing depending on the energy qualities of buildings in the EU* (pp. 1–44). Technische Universität.

GIGAbase. (2022). RESET standard. GIGAbase Canada. Retrieved from https://reset.build/

Gillingham, K., Harding, M., & Rapson, D. (2012). Split incentives in residential energy consumption. *The Energy Journal*, *33*(2), 37–62. https://doi.org/10.5547/01956574.33.2.3

Goffart, J., Mara, T., & Wurtz, E. (2017). Generation of stochastic weather data for uncertainty and sensitivity analysis of a low-energy building. *Journal of Building Physics*, *41*(1), 41–57. https://doi.org/10.1177/1744259116668598

Government of the United Kingdom. (2023). Guidance: Non-domestic private rented property: Minimum energy efficiency standard - landlord guidance. Retrieved from https://www.gov.uk/government/publications/non-domestic-private-rented-property-minimum-energy-efficiency-standard-landlord-guidance

Gram-Hanssen, K., Georg, S., Christiansen, E., & Heiselberg, P. (2018). What next for energy-related building regulations?: The occupancy phase. *Building Research & Information*, *46*(7), 790–803. https://doi.org/10.1080/09613218.2018.1426810

GRESB. (2024). GRESB | Global ESG benchmark for real assets. Retrieved from https://gresb.com/

Guan, L. (2009). Preparation of future weather data to study the impact of climate change on buildings. *Building and Environment*, *44*(4), 793–800. https://doi.org/10.1016/j.buildenv.2008.05.021

Gursel, I., Sariyildiz, S., Akin, Ö., & Stouffs, R. (2009). Modeling and visualization of lifecycle building performance assessment. *Advanced Engineering Informatics*, *23*(4), 396–417. https://doi.org/10.1016/j.aei.2009.06.010

Haas, R., Auer, H., & Biermayr, P. (1998). The impact of consumer behavior on residential energy demand for space heating. *Energy and Buildings*, *27*(2), 195–205. https://doi.org/10.1016/S0378-7788(97)00034-0

Happle, G., Fonseca, J. A., & Schlueter, A. (2018). A review on occupant behavior in urban building energy models. *Energy and Buildings*, *174*(September), 276–292. https://doi.org/10.1016/j.enbuild.2018.06.030

Herrera, M., Natarajan, S., Coley, D. A., Kershaw, T., Ramallo-González, A. P., Eames, M., Fosas, D., & Wood, M. (2017). A review of current and future weather data for building simulation. *Building Services Engineering Research and Technology*, *38*(5), 602–627. https://doi.org/10.1177/0143624417705937

Holtermans, R., & Kok, N. (2017). National green building adoption index 2017. University of Maastricht, CBRE, EnergyStar. Retrieved from https://www.energystar.gov/sites/default/files/tools/Green-Building-Adoption-Index-2014_Final.pdf

Hong, T., Yan, D., D'Oca, S., & Chen, C.-f. (2017). Ten questions concerning occupant behavior in buildings: The big picture. *Building and Environment*, *114*(March), 518–530. https://doi.org/10.1016/j.buildenv.2016.12.006

Hu, S., Wang, J., Hoare, C., Li, Y., Pauwels, P., & O'Donnell, J. (2021). Building energy performance assessment using linked data and cross-domain semantic reasoning. *Automation in Construction*, *124*, 103580. https://doi.org/10.1016/j.autcon.2021.103580

Hui, S. C. M. (2002). Using performance-based approach in building energy standards and codes. *Proc. of the Chonqing-Hong Kong Joint Symposium 2002*, 8–10 July 2002, pp. A52–61.

IEA. (2013). "Policy pathway - Modernising building energy codes 2013, IEA, Paris. Retrieved from https://www.iea.org/reports/policy-pathway-modernising-building-energy-codes-2013

IEA. (2018). Annex 66 - Definition and simulation of occupant behavior in buildings. Retrieved from http://www.annex66.org/

IEA. (2022). Annex 79 - Occupant-centric building design and operation. Retrieved from https://annex79.iea-ebc.org/

IFC. (2022). EDGE Certification. International Finance Corporation. World Bank Group. Retrieved from https://edgebuildings.com/

IWBI. (2022). The WELL Building Standard (WELL). International WELL Building Institute. Retrieved from https://v2.wellcertified.com/en

Kotireddy, R., Loonen, R., Hoes, P.-J., & Hensen, J. L. M. (2019). Building performance robustness assessment: Comparative study and demonstration using scenario analysis. *Energy and Buildings*, *202*(November), 109362. https://doi.org/10.1016/j.enbuild.2019.109362

Lee, W. L., & Yik, F. W. H. (2004). Regulatory and voluntary approaches for enhancing building energy efficiency. *Progress in Energy and Combustion Science*, *30*(5), 477–499. https://doi.org/10.1016/j.pecs.2004.03.002

Li, H., Wang, Z., & Hong, T. (2021). Occupant-centric key performance indicators to inform building design and operations. *Journal of Building Performance Simulation*, *14*(6), 814–842. https://doi.org/10.1080/19401493.2021.1876771

Marchi, L., Antonini, E., & Politi, S. (2021). Green building rating systems (GBRSs). *Encyclopedia*, *1*(4), 998–1009. https://doi.org/10.3390/encyclopedia1040076

Maruejols, L., & Young, D. (2011). Split incentives and energy efficiency in Canadian multi-family dwellings. *Energy Policy*, *39*(6), 3655–3668. https://doi.org/10.1016/j.enpol.2011.03.072

Matschoss, K., Heiskanen, E., Atanasiu, B., & Kranzl, L. (2013, June). Energy renovations of EU multifamily buildings: Do current policies target the real problems? In *Rethink, renew, restart. Proceedings of the eceee 2013 Summer Study* (pp. 1485–1496).

Mavromatidis, G., Orehounig, K., & Carmeliet, J. (2018). A review of uncertainty characterisation approaches for the optimal design of distributed energy systems. *Renewable and Sustainable Energy Reviews*, *88*, 258–277. https://doi.org/10.1016/j.apenergy.2018.01.062

Melvin, J. (2018). The split incentives energy efficiency problem: Evidence of underinvestment by landlords. *Energy Policy*, *115*(April), 342–352. https://doi.org/10.1016/j.enpol.2017.11.069

Moazami, A., Carlucci, S., & Geving, S. (2019). Robust and resilient buildings: A framework for defining the protection against climate uncertainty. *IOP Conference Series: Materials Science and Engineering*, *609*, 072068. https://doi.org/10.1088/1757-899X/609/7/072068

Moazami, A., Nik, V. M., Carlucci, S., & Geving, S. (2019). Impacts of future weather data typology on building energy performance investigating long-term patterns of climate change and extreme weather conditions. *Applied Energy*, *238*(March), 696–720. https://doi.org/10.1016/j.apenergy.2019.01.085

NABERS. (2022). National Australian built environment rating system - NABERS. Federal, State and Territory governments of Australia. Retrieved from https://www.nabers.gov.au/

O'Brien, W., Gaetani, I., Carlucci, S., Hoes, P., J, &., & Hensen, J. L. (2017). On occupant-centric building performance metrics. *Building and Environment*, *122*, 373–385. https://doi.org/10.1016/j.buildenv.2017.06.028

O'Brien, W., Tahmasebi, F., Andersen, R. K., Azar, E., Barthelmes, V., Belafi, Z. D., Berger, C., Chen, D., De Simone, M., d'Oca, S., Hong, T., Jin, Q., Khovalyg, D., Lamberts, R., Novakovic, V., Park, J. Y., Plagmann, M., Rajus, V. S., Vellei, M., … Zhou, J. (2020). An international review of occupant-related aspects of building energy codes and standards. *Building and Environment*, *179* (July), 106906. https://doi.org/10.1016/j.buildenv.2020.106906

Pan, Y., Zhu, M., Lv, Y., Yang, Y., Liang, Y., Yin, R., Yang, Y., Jia, X., Wang, X., Zeng, F., Huang, S., Hou, D., Xu, L., Yin, R., & Yuan, X. (2023). Building energy simulation and its application for building performance optimization: A review of methods, tools, and case studies. *Advances in Applied Energy*, *10*, 100135. https://doi.org/10.1016/j.adapen.2023.100135

Paoli, L., & Cullen, J. (2020). Technical limits for energy conversion efficiency. *Energy*, *192*, 116228. https://doi.org/10.1016/j.energy.2019.116228

Pernigotto, G., Prada, A., & Gasparella, A. (2020). Extreme reference years for building energy performance simulation. *Journal of Building Performance Simulation*, *13*(2), 152–166. https://doi.org/10.1080/19401493.2019.1585477

Picard, T., Hong, T., Luo, N., Lee, S. H., & Sun, K. (2020). Robustness of energy performance of zero-net-energy (ZNE) homes. *Energy and Buildings*, *224*(October), 110251. https://doi.org/10.1016/j.enbuild.2020.110251

Pollitt, C. (2013). The logics of performance management. *Evaluation*, *19*(4), 346–363. https://doi.org/10.1177/1356389013505040

Poretti, C., & Das, P. (2020). What explains differing holding periods across hotel investments? A hazard rate framework. *International Journal of Hospitality Management*, *89*(August), 102564. https://doi.org/10.1016/j.ijhm.2020.102564

Rastogi, P. (2016). *On the sensitivity of buildings to climate: The interaction of weather and building envelopes in determining future building energy consumption* (PhD), Lausanne, Switzerland: Ecole polytechnique fédérale de Lausanne. Retrieved from https://infoscience.epfl.ch/record/220971?ln=en

Rosenberg, M., & Hart, R. (2014). Roadmap toward a predictive performance-based commercial energy code. In ACEEE *summer study on energy efficiency in buildings*.

Rosenberg, M., Jonlin, D., & Nadel, S. (2016). A perspective of energy codes and regulations for the buildings of the future. *Journal of Solar Energy Engineering*, *139*(1), 010801. https://doi.org/10.1115/1.4034825

Sandberg, N. H., Heidrich, O., Dawson, R., Dimitriou, S., Vimm-R, T., Filippidou, F., Stegnar, G., Zavrl, M. Š., & Brattebø, H. (2016). "Dynamic building stock modelling: Application to 11 European countries to support the energy efficiency and retrofit ambitions of the EU. *Energy and Buildings*, *132*(November), 26–38. https://doi.org/10.1016/j.enbuild.2016.05.100

Santamouris, M., & Vasilakopoulou, K. (2021). Present and future energy consumption of buildings: Challenges and opportunities towards decarbonisation. *e-Prime*, *1*, 100002. https://doi.org/10.1016/j.prime.2021.100002

Schwarz, M., Nakhle, C., & Knoeri, C. (2020). Innovative designs of building energy codes for building decarbonization and their implementation challenges. *Journal of Cleaner Production*, *248*(March), 119260. https://doi.org/10.1016/j.jclepro.2019.119260

Selvacanabady, A., & Judd, K. (2017). *The influence of occupancy on building energy use intensity and the utility of an occupancy-adjusted performance metric*. PNNL-26019. Pacific Northwest National Laboratory.

SIA. (2016). "380/1: Heizwärmebedarf". Swiss Society of Engineers and Architects. shop.sia.ch/normenwerk/architekt/380-1_2016_d/D/Product

Soares, N., Bastos, J., Pereira, L. D., Soares, A., Amaral, A. R., Asadi, E., Rodrigues, E., Lamas, F. B., Monteiro, H., Lopes, M. A. R., & Gaspar, A. R. (2017). A review on current advances in the energy and environmental performance of buildings towards a more sustainable built environment. *Renewable and Sustainable Energy Reviews*, *77*, 845–860. https://doi.org/10.1016/j.rser.2017.04.027

Tahmasebi, F., & Mahdavi, A. (2017). The sensitivity of building performance simulation results to the choice of occupants' presence models: A case study. *Journal of Building Performance Simulation*, *10*(5-6), 625–635. https://doi.org/10.1080/19401493.2015.1117528

Talbot, J., & Love, J. (2014). I will if you will: A replicable strategy for tenant behavior change. *ACEEE Summer Study on Energy Efficiency in Buildings*, *13*.

Ueno, K. (2010). "BSD-152: Building energy performance metrics." Building Science Digest. Retrieved from https://www.buildingscience.com/documents/digests/bsd152-building-energy-performance-metrics

USGBC. (2022). "LEED rating system". U.S. Green Building Council. Retrieved from https://www.usgbc.org/leed

Van Dronkelaar, C., Dowson, M., Spataru, C., & Mumovic, D. (2016). A review of the regulatory energy performance gap and its underlying causes in non-domestic buildings. *Frontiers in Mechanical Engineering*, *1*(January). https://doi.org/10.3389/fmech.2015.00017

Wang, S., Yan, C., & Xiao, F. (2012). Quantitative energy performance assessment methods for existing buildings. *Energy and Buildings*, *55*, 873–888. https://doi.org/10.1016/j.enbuild.2012.08.037

Wasserbaur, R., Sakao, T., Ljunggren Söderman, M., Plepys, A., & Dalhammar, C. (2020). What if everyone becomes a sharer? A quantification of the environmental impact of access-based consumption for household laundry activities. *Resources, Conservation and Recycling*, *158*(July), 104780. https://doi.org/10.1016/j.resconrec.2020.104780

Yan, D., O'Brien, W., Hong, T., Feng, X., Burak Gunay, H., Tahmasebi, F. & Mahdavi, A. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings*, *107*(November), 264–278. https://doi.org/10.1016/j.enbuild.2015.08.032

Younger, M., Morrow-Almeida, H. R., Vindigni, S. M., & Dannenberg, A. L. (2008). The built environment, climate change, and health: Opportunities for co-benefits. *American Journal of Preventive Medicine*, *35*(5), 517–526. https://doi.org/10.1016/j.amepre.2008.08.017

YPEN. (2010). Regulation on the energy performance of buildings – KENAK. *Greek Ministry of Environment and Energy*, YPEN. Retrieved from ypen.gov.gr/energeia/energeiaki-exoikonomisi/ktiria/kenak/