Energy & Buildings 292 (2023) 113171

Contents lists available at ScienceDirect

Energy & Buildings

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

Building performance simulation in the brave new world of artificial intelligence and digital twins: A systematic review

Pieter de Wilde

University of Strathclyde, Department of Architecture, James Weir Building, 75 Montrose Street, Glasgow G1 1XJ, United Kingdom

ARTICLE INFO

Keywords:
Digital twin
Machine learning
Artificial intelligence
Cyber-physical system
Internet of things
Data mining
Building performance simulation

ABSTRACT

In an increasingly digital world, there are fast-paced developments in fields such as Artificial Intelligence, Machine Learning, Data Mining, Digital Twins, Cyber-Physical Systems and the Internet of Things. This paper reviews and discusses how these new emerging areas relate to the traditional domain of building performance simulation. It explores the boundaries between building simulation and these other fields in order to identify conceptual differences and similarities, strengths and limitations of each of these areas. The paper critiques common notions about these new domains and how they relate to building simulation, reviewing how the field of building performance may evolve and benefit from the new developments.

1. Introduction

Building performance simulation is the domain that replicates and predicts aspects of building performance, using a computer-based, mathematical model and applying fundamental physical principles and engineering techniques. Building performance simulation is a buoyant field, enjoying significant research, development, and an increased uptake in practice. However, building simulation does not exist in a vacuum. There are other digital developments in the wider building sector that are also gaining traction and interest, such as the emergence and progress of work on Digital Twins, Cyber-Physical Systems, Artificial Intelligence and Machine Learning, the Internet of Things, and Data Mining. These other fields partly overlap and partly compete with traditional views of building performance simulation.

Changes in the information technology and digital world are increasingly fast-paced. As a taster of generic developments, a quick dive in the technology briefs of leading IT consulting companies like Accenture and Gartner shows the rapid emergence of a wide range of interrelated digital concepts and themes like digital twinning, AI engineering, autonomous systems and others [1,2]. These IT topics typically permeate the building science domain with a delay. For instance, the term Digital Twin was coined in 2003 by Grieves [3], but it only appeared in the building performance literature around 2017 [4]. Similarly, the general concept of Cyber-Physical Systems emerged in 2006 [5] but the transition to the building performance domain took until 2015 [6].

The trends in numbers of peer-reviewed scientific publications within the general subject area of 'building' or building (performance) simulation, artificial intelligence, machine learning, digital twins, cyber-physical systems, internet of things and data mining are depicted in Fig. 1. The data in this figure was collected by searching for the corresponding keywords in the Primo (ExLibris) library search engine, which also accesses Web of Science and Scopus. A filter was applied to only show academic journal articles. Findings were then grouped in bins of 5 years, starting in 1965 and ending with 2020; the bin of 2020–2025 was excluded as this is still incomplete. Building (performance) simulation thus far is the dominant sub-domain, however artificial intelligence and machine learning also show a very similar trend. Papers on the internet of things and data mining have a later start but are seeing a steep increase. Digital twins only recently have started to emerge as a serious area of interest.

So, how do these emergent domains relate to the existing field of building performance simulation? How much of these new development is hyperbole and mainly new terminology for existing building simulation concepts? Are the new domains competitors, or are these new areas of work that allow building performance simulation to expand its sphere of influence? What are common notions and potential misconceptions about these fields in the building simulation area? When deciding on a career in building science, applying for research funding, or contemplating novel research, should one continue along the strong roots of building simulation or opt for one of the new 'hot topics'?

The key roles of scientific research are description, explanation/understanding, and prediction. As such it is important to understand and delineate these domains, and identify the underlying concepts and their

E-mail address: pieter.dewilde@strath.ac.uk.

https://doi.org/10.1016/j.enbuild.2023.113171
Received 4 March 2023; Received in revised form 28 April 2023; Accepted 10 May 2023
Available online 18 May 2023

properties, differences and similarities, strengths and limitations [7]. This article explores how building performance simulation relates to Artificial Intelligence, Machine Learning, Digital Twins, Cyber-Physical Systems, the Internet of Things and Data Mining. These terms were selected as the main ICT themes that appear in publications on building performance in academic journals such as the *Journal of Building Performance Simulation*, *Automation in Construction* and *Advanced Engineering Informatics*, as well as industry publications such as the *ASHRAE Journal* and *CIBSE Journal*. Note that there are further terms such as blockchain, deep learning, surrogate models, big data, explainable artificial intelligence and others that could feature in a wider exploration, however these are beyond the scope of this manuscript.

To achieve this aim, it has the following objectives:

1. review the generally accepted notions /definitions of the core concepts of Artificial Intelligence, Machine Learning, Digital Twins, Cyber-Physical Systems and Data Mining;
2. map out common notions about differences, similarities, strengths and limitations of each of these domains;
3. explore what work takes place on the interface of these new areas and building performance simulation, and what the main themes are;
4. critique what widespread notions and views about these new domains are held within the building performance simulation community;
5. review how the field of building performance simulation should position and develop itself vis-à-vis the new domains.

The approach followed in this study is dual: in part it follows a structured review, in other parts it is a position and discussion paper. The latter is the consequence of the start from inside the domain of building performance simulation, looking across to the novel digital domains.

The systematic literature review follows the roadmap and stages depicted in Fig. 2. To set the focus for review, a generic search in the literature was done in order to capture notions and definition of key concepts, and enable an initial conceptual comparison. This was followed by the main study, which took place between February and June 2022. The systematic literature review used the following keywords for scope: 'building (performance) simulation AND artificial intelligence' plus 'building (performance) simulation AND machine learning'; 'building (performance) simulation AND digital twin(s)', 'building (performance) simulation AND cyber-physical system'; 'building (performance) simulation AND internet of things', and 'building (performance) simulation AND data mining'. Note that these are more specific and limited searches than were used to find generic trends in Fig. 1. Only studies written and published in academic journals were included. A small number of books and reports that appeared as seminal texts in papers was added through backward searches. All material used is written in English. No date limitations were imposed on the search, as the emerging fields are mostly self-selective. The initial search was carried out using Primo (ExLibris), a common front-end discovery service that gives access to all resources available in most UK academic catalogues, including the British Library. Further dedicated searches were conducted using Scopus and Web of Science. For most search terms, an exhaustive search was carried out. For two searches the number of available documents was so large that it was decided to screen returns until saturation; in these cases the search was halted when in review of further literature only yielded the same topics and issues coming up. This concerns the search for papers using the search terms that (1) combine 'building (performance) simulation AND artificial intelligence' plus 'building (performance) simulation AND machine learning' and (2) combine 'building (performance) simulation AND data mining'. Retrieved articles were assessed for eligibility through screening of titles, abstract and methodology sections. Eligible publications were subjected to thematic analysis, whilst data was extracted on how the work relates to building simulation and the emerging field. Notions and views were also captured. The results represent perceptions as described in the literature that spans both building performance simulation and the adjacent domains. Results have been grouped in similar headings were possible; note that these groupings emerged from the thematic analysis and were not pre-defined search categories.

The remainder of this paper is structured as follows. Section 2 provides a very brief summary of the field of building performance simulation, as basis for the comparison with the other areas. Section 3 introduces the general notions and definitions of digital developments in the specific areas of Artificial Intelligence / Machine Learning, Digital Twins, Cyber-Physical Systems and Data Mining. Section 4 then maps out the differences and overlaps between these definitions and notions with building simulation, and then presents (perceived) strengths and limitations of each of the domains that appear in literature. Section 5 presents work that takes place on the interface between building simulation and the novel digital domains, grouped in the main themes that emerge on this interface. Section 6 concludes the paper by summarizing findings, critiquing some of the notions and views that were found, and recommending how the field of building performance should position itself towards the new domains.

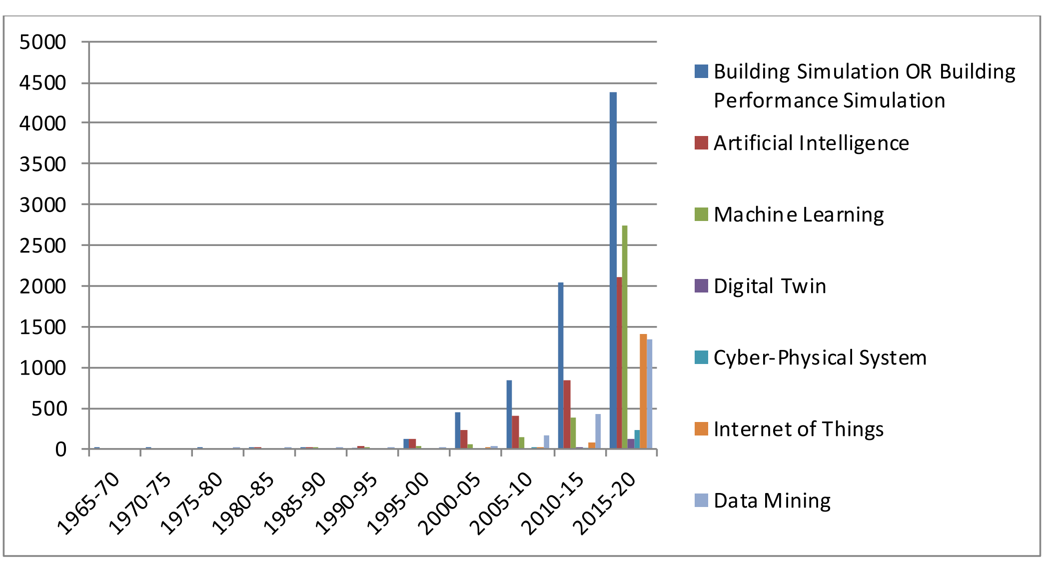

Fig. 1. Trends in number of scientific peer-reviewed publications on topics within the building domain.

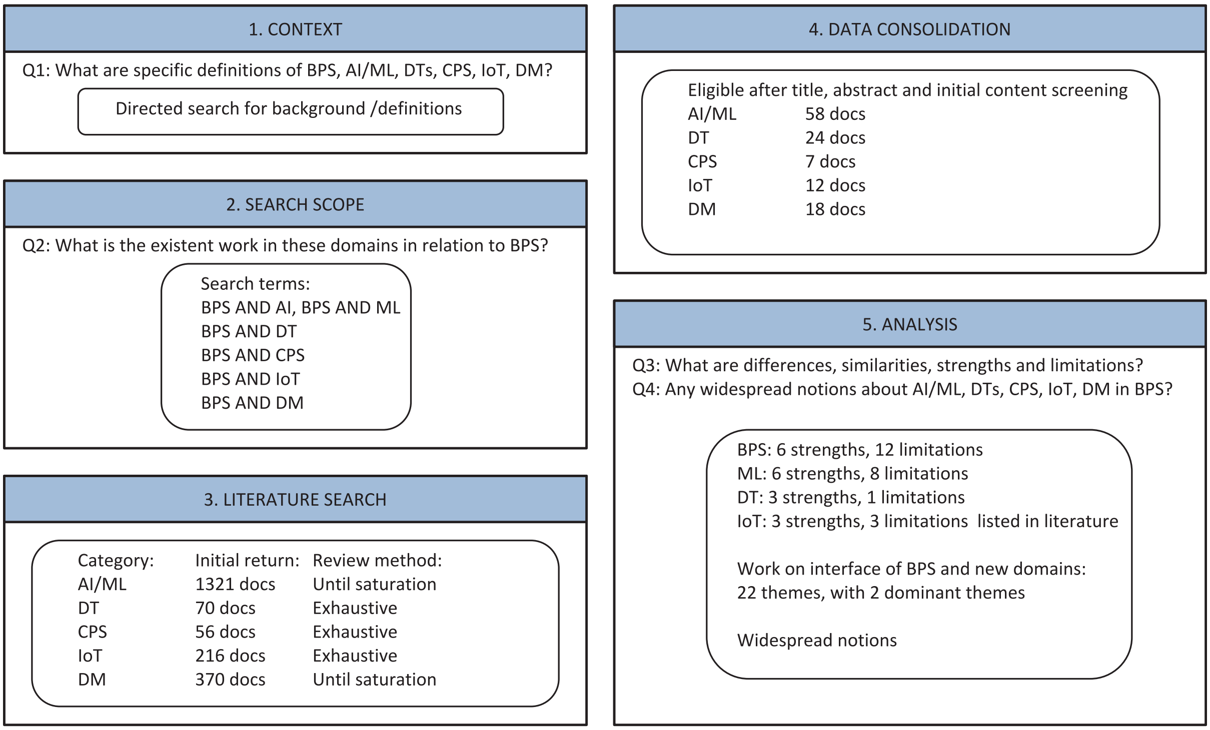

Fig. 2. Roadmap for systematic review.

2. Backgrounds, roots and state-of-the-art of building performance simulation

The essence of *building performance simulation* is the use of computer programs to imitate building reality. Building simulation involves the replication of building behaviour using a computer-based, mathematical model and the application of fundamental physical principles and engineering models. Modelling and simulation are part of the wider domain of scientific computing [8]. There are many different simulation models; for instance one may classify them as linear or nonlinear, static or dynamic, discrete or continuous, deterministic or stochastic. They may also represent different physical processes, such as heat and mass transfer, lighting, or acoustics. Seminal overviews of the essentials of building performance simulation are provided in the books by Beausoleil-Morrison [9] which explores the fundamental principles, Clarke [10] or Hensen and Lamberts [11] which cover fundamentals, application to design, operational optimization, system and urban simulation, and de Wilde [12] which emphasizes the underlying concept of building performance and its quantification. There are also many papers that review developments in the field, such as the overview by Wang and Zhai [13].

Building simulation has evolved over several decades; see for instance the paper by Oh and Haberl [14] that traces back the origin of a number of whole-building energy simulation tools in the US to their foundations in the 1970 s. At the beginning of the new millennium, Augenbroe [15] described trends in building simulation, observing the maturation of the field over the three decennia spanning 1970–2000. He distinguished two key aspects as driving the maturation process: increased levels of quality assurance in building simulation, and increased integration of tools and expertise in the building process. Spitler [16] echoes these findings, but also points out some of the remaining challenges that prevent building simulation from fully realizing its promise: a need to more easily capture complex geometries and systems, to better integrate simulation with data-driven systems, to develop new models that better deal with large ranges of time scales and further types of physics, to create novel methods that allow optimization under uncertainty and risk, and to make further advances in verification and validation efforts. Similar issues are raised as focus points for the development of building performance simulation in the vision paper by Clarke [17]. Recently, significant research efforts in building simulation are directed to the 'performance gap': a mismatch between predicted and measured performance [18]. In essence this performance gap issue concerns the fundamental issues of verification, validation and calibration in scientific computing in the building domain, and asks how good the existing models are.

3. General background to the new digital developments

Whilst this paper talks about new digital developments, most of these subjects are already established and introduced in dedicated books; see for instance Zhang et al on Artificial Intelligence [19], Murphy on Machine Learning [20], Viola and Chen on Digital Twins [21], Alur on Cyber-Physical systems [22], Gyasi-Agyei on the Internet of Things [23] and Zaki and Meira on Data Mining [24]. The following paragraphs give a general introduction.

Artificial Intelligence (AI) is the concept where machines demonstrate intelligence. It is a broad study area that is concerned with systems that are able to perceive their environment and that can take actions towards achieving some objective. The Royal Society [25] defines AI is "*an umbrella term for the science of making machines smart*" and in general as efforts that create "*systems that think like humans, act like humans, think rationally, or act rationally*". Amongst others, it includes the areas of knowledge representation, planning, natural language processing, reasoning, and the planning and control of objects and systems. Artificial Intelligent methods are often grouped into symbolic AI and subsymbolic AI. Symbolic AI methods are based on high-level problem representations, logic and search approaches which are human-readable. Subsymbolic AI methods are approaches that learn directly from data; subsymbolic AI equates to Machine Learning. The application of artificial intelligence to buildings typically takes place in the fields of intelligent or smart buildings. Whilst there is a long-standing debate on what consist an intelligent/smart building [26] the emphasis is on creating buildings that respond to human and organizational needs rather than 'buildings that think like humans'. Efforts often take place at different scale levels, notably "smart district" and "smart city".

Machine Learning (ML) is the discipline that is concerned with computer algorithms that improve through the use of data and experience. Machine Learning is a subfield of Artificial Intelligence. The Royal Society [25] defines ML as "a *set of rules that allows systems to learn directly from examples, data and experience*". Machine Learning algorithms can be used for prediction and classification. Machine Learning techniques are typically classed as supervised, unsupervised or reinforcement learning [27]. Supervised learning uses training data to create a mathematical model that fits input and output data; this model can then predict output for input parameters that sit within the training range. It is typically used for regression (predicting/obtaining numeric values), data classification (sorting data into categories), and optimization and control [28]. Supervised learning typically employs t**raining data**, which is defined as "*data that can be used to train machine learning systems, having already been labelled or categorised into one or more groups*" [25]. Once the model has been created, other data stemming from the same source can be used to test and validate the model. Supervised machine learning is sometimes named 'predictive'. Sometimes different supervised machine learning algorithms can be combined; this is named ensemble learning [29]. Unsupervised learning searches for unknown patterns in a data set to predict outputs. It is typically used for clustering (identifying similarities amongst sets) or dimensional reduction (reducing the number of variables) [28]. Unsupervised learning is sometimes named 'descriptive'.

Reinforcement learning is an approach that employs algorithms or "agents" to navigate sequential decision making in an environment with limited feedback. This approach is especially suitable for defining policies and behaviours [30]. Reinforcement learning is sometimes seen as a method that belongs to semi-supervised techniques, which also can be used for optimization and control and generative models [28]. The main types of machine learning models used in artificial intelligence are linear regression models, linear classification models, neural networks, kernel methods, sparse kernel machines, graphical models, and mixture models and expectation–maximization algorithms [21]. The list of specific algorithms and approaches that supports Machine Learning is long and keeps growing; see Table 1 for a sample of some terms often seen in the building performance literature. A good overview of the main approaches of ML, discussed in the context of predicting daylighting in buildings, is provided by Ayoub [31]. Sun *et al* [32] provide another worthwhile overview in the context of structural analysis. Deng *et al.* [33] and Seyedzadeh *et al.* [34] discuss the application of some of these approaches to building energy prediction. Fathi *et al.* [35] review application at the urban energy prediction scale. In the realm of buildings, machine learning may feature wherever algorithms are used, such as in engineering design calculations or building control systems.

Whilst ML and subsymbolic AI are sometimes seen as new areas, the origins of these approaches have a long history that links back to statistical analysis. These 'data driven' approaches like (linear) regression have been around in building performance studies for a long time, as can be seen in Fig. 1. However recent advances in computing power allow to push the boundaries on possibilities, and have created a surge of interest in the approach. Reinforcement learning, as a method to guide sequential decision making in an environment with limited feedback, is highly relevant for the control of building systems.

A **Digital Twin** (DT) is generally defined as a real-time digital counterpart of a physical system or process. There are various definitions that expands on this core. For instance, the Centre for Digital Built Britain [36] defines Digital Twins as "*realistic digital representations of physical assets, for a example a digital representation of an aeroplane that can be used to monitor and predict performance, feeding out insights and interventions. These insights lead to better interventions and unlock real-world value from assets through financial savings, improved performance and services, and better outcomes for society*". Other authors and groups emphasize the link of a DT with real-world data, the need to represent the physical system across its lifecycle, and the enabling of analytical

Table 1
List of common Machine Learning algorithms and approaches, adapted/expanded after [28].

Supervised ML - classification	Support Vector Machine (SVM) Decision Tree (DT)Random Forest (RF)Neural Network (NN)Artificial Neural Network (ANN)Adaptive Neuro Fuzzy Inference System (ANFIS)Gradient Boosting (GB)
Supervised ML - regression	Linear RegressionMultiple Linear Regression (MLR)Polynomial Regression (PR)Gaussian Process (GP)AutoRegressive and Moving Average (ARMA)AutoRegressive Integrated Moving Average (ARIMA Multivariate Adaptive Regression Splines (MARS)Least Absolute Shrinkage and Selection Operator regression (LASSO)
Supervised ML – optimization and control	Linear Control Genetic AlgorithmDeep Model Predictive Control (DMPC)
Semi-supervised ML - reinforcement learning	Q-learning Markov Decision Processes Deep Reinforcement Learning
Semi-supervised ML - generative models	Generative Adversarial Network (GAN)
Unsupervised ML – clustering	K-meansK-Nearest Neighbours (KNN) Spectral Clustering
Unsupervised ML - dimensionality reduction	Principal Component Analysis (PCA)Self Organizing Map (SOM)

processes. On the surface, a Digital Twin is a software representation of some physical entity and thus might be considered to be the same as any building performance simulation or machine learning model. A Digital Twin may also be perceived as something very similar to a Building Information Model (BIM). However, in the engineering domain the concept of Digital Twins has already seen significant development and gained deeper meanings. The notion of a Digital Twin was first suggested in 2003 in the context of manufacturing and product lifecycle management [3] and was defined as a concept that combines a physical product, a virtual model or that product, and connections between the physical and virtual parts. As with any novel concept there are still different definitions and interpretations of Digital Twins [37,38]. Yet a common notion is that there should be an information flow from the physical part to the digital replica. This requires that the digital part of the twin is regularly updated so that it represents the actual state of the physical part. Another common notion is that a Digital Twin must make a contribution to the operation and management of the physical counterpart, for instance through predictions that impact control, maintenance and replacement of components. Ganguli and Adhikari [39] summarize this as a Digital Twin requiring four key aspects: modelling and simulation, data fusion, interaction and collaboration, and service. The manufacturing view of Digital Twins suggests differentiation along the project life cycle. A Digital Twin Prototype (DTP) is used during the design phase. Once a product is actually made, each single product instance is coupled to a Digital Twin Instance (DTI). Ultimately, different DTIs can be combined into a Digital Twin Aggregate (DTA) which captures the overall product behaviour and which can be used for learning and product improvement [38]. Whilst Digital Twins are generally seen as a promising new technology, they have also attracted some critique in terms of opaque objectives towards optimization and efficiency [40]. A solid review of Digital Twins in general is the paper by Tao *et al.* [41]; for a good discussion of Digital Twins in the built environment see Brilakis *et al.* [37] and Delgado and Oyedele [42]. The latter also discuss the differences between the concepts of Digital Twins and BIM in further detail. Applications to buildings have just started to emerge over the past years; often these relate to building control and operation.

A **Cyber-Physical System** (CPS) is a system that combines physical and cyber or software components. The National Science Foundation defines Cyber-Physical systems as "*engineered systems that are built from, and depend upon, the seamless integration of computation and physical components*" and goes on to list control, data analytics, machine learning including real-time learning for control, system autonomy, system design, Internet of Things (IoT), mixed initiatives including human-in-the-loop, networking, real-time systems, system safety and security, and verification as key contributors to CPS [43]. Cyber-Physical Systems thus integrate two systems: a computational or 'cyber' system and physical system. However, unlike Digital Twins, the computational system does not need to represent the physical one. Instead, here the computational system can be considered the 'smart', 'intelligent' or control system that operates and steers the physical system. An example would be a robotic arm, or indeed a building with a digital building management system. A seminal work on Cyber-Physical Systems in the built environment is the book by Anumba and Roofigari-Esfahan [6]. Given their definition, Cyber-Physical Systems in the built environment have a significant overlap with the existent body of knowledge on building control and operation systems. However, the term still offers a new viewpoint to re-evaluate technological efforts in the field.

The **Internet of Things** (IoT) has been defined as the concept of connecting any device to the internet and to other connected devices, thus creating a network of connected devices and people. This network then collects and shares data about the status of all these devices, the way they are used, and the environment in which they sit [44]. In buildings, the concept is widely used for wireless data collection and sensing devices. The idea of the Internet of Things relates to the connection of any type of object to other objects using the internet, building on the ubiquitous sensors, software and distributed computing opportunities. In a common example it may allow the fridge to communicate with a supplier to ensure stocks remain at required levels; in buildings it may allow the window to communicate with the heating system to notify that it is in an 'open' position and hence preventing wasting energy. Jia *et al.* [45] discuss the application of the Internet of Things to (smart) buildings. Tang *et al.* [46] explore the integration of the IoT with BIM.

Data Mining (DM) is the extraction and discovery of patterns in large data sets, using machine learning, statistics and database technology. As such, the concept of Data Mining employs Machine Learning methods. IBM defines data mining as "*the process of uncovering patterns and other valuable information from large data sets*", and considers it a synonym for knowledge discovery in data (KDD). IBM divides the techniques that underpin data mining into two classes: one that describes the data itself, and one that predicts outcomes on the basis of machine learning [47]. Data Mining typically involves the analysis of large data sets (which may be Big Data or not) using traditional statistical, mathematical and computer science approaches, often using machine learning. In the context of building performance, data mining is mostly applied to energy efficiency; see for instance Pena *et al.* [48], Fan *et al.* [49] or Zhao *et al.* [50]. Efforts in machine learning, big data analytics and sensing through the Internet of Things often overlap and combine; see for instance the paper by Qolomany *et al.* [51] for a discussion on how these combine in the work towards smart buildings. Another term associated with data mining is **big data**. Big data is defined as data that is characterized by a large variety, increased volume and velocity of arrival, thus yielding larger and more complex data sets. Big data sets may be too large or complex for traditional data analysis approaches in terms of volume, variety and velocity [52].

4. Conceptual overlaps, differences, strengths and limitations

The definitions and discussions in the previous section can now be compared and contrasted with the notion of building performance simulation. The following observations are made on conceptual differences and overlaps:

- Artificial Intelligence is a generic concept that may be applied to "smart" and "intelligent" systems, buildings and cities. Whilst building performance simulation may be used in achieving the smart and intelligent aspect, there is no conceptual overlap leading to misunderstandings.
- Cyber-Physical Systems combine and integrate computational systems with physical systems. The computational side may involve building performance simulation. However, this always will be part of the larger CPS, which must also include physical parts. Conceptually, this allows to delineate where simulation fits within the CPS.
- The Internet of Things is a network of connected devices and people. Whilst the IoT may be of high interest in gathering data from a physical building and its content, there is no conceptual overlap with building performance simulation.
- Conceptual overlaps does exist between four domains that all involve models: building performance simulation, machine learning, digital twins and building information modelling. To differentiate between these four categories, the following shows the main emphasis in each category:
  - o Building performance simulation models are scientific computational models, based on physics and engineering principles. Where physical processes are sufficiently understood and can be fully described, building simulation models can be considered 'white box models' or 'first principle models'. However it must be noted that building simulation also employs approximations and empirical correlations.
  - o A term often used in computing and engineering is 'black box model'. This refers to a model that established a relation of inputs and outputs of a system, without having knowledge of underlying mechanisms or working. The concept of a black box model overlaps with sub-symbolic AI and supervised ML.
  - o Digital Twin models are a real-time digital counterpart of an existing physical system (building or building system).
  - o Building information models are a comprehensive digital representation of a building, and typically capture information about 3D geometry of the building and systems, spaces and zones, and the project structure/schedule.

All four types of models represent buildings or building systems, and have some sort of data or information flow that connects them to these buildings. However, the content and direction of these data flows differs. For BIM, the content is typically building attributes and the flow is only one-way towards the model. For building performance simulation, data is needed to convey building attributes but may also involve performance data that is used for validation and calibration of the model; performance prediction might be sent back to building to support things like control and building management. For machine learning models the input data flow changes to training data, where attributes are no longer included. Again there is flow back to the building to support control and management. A tight two-way interflow is required for digital twins, as this concept requires a model that reflects the current state of the building and data flowing back to the actual building to provide a service. See Fig. 3.

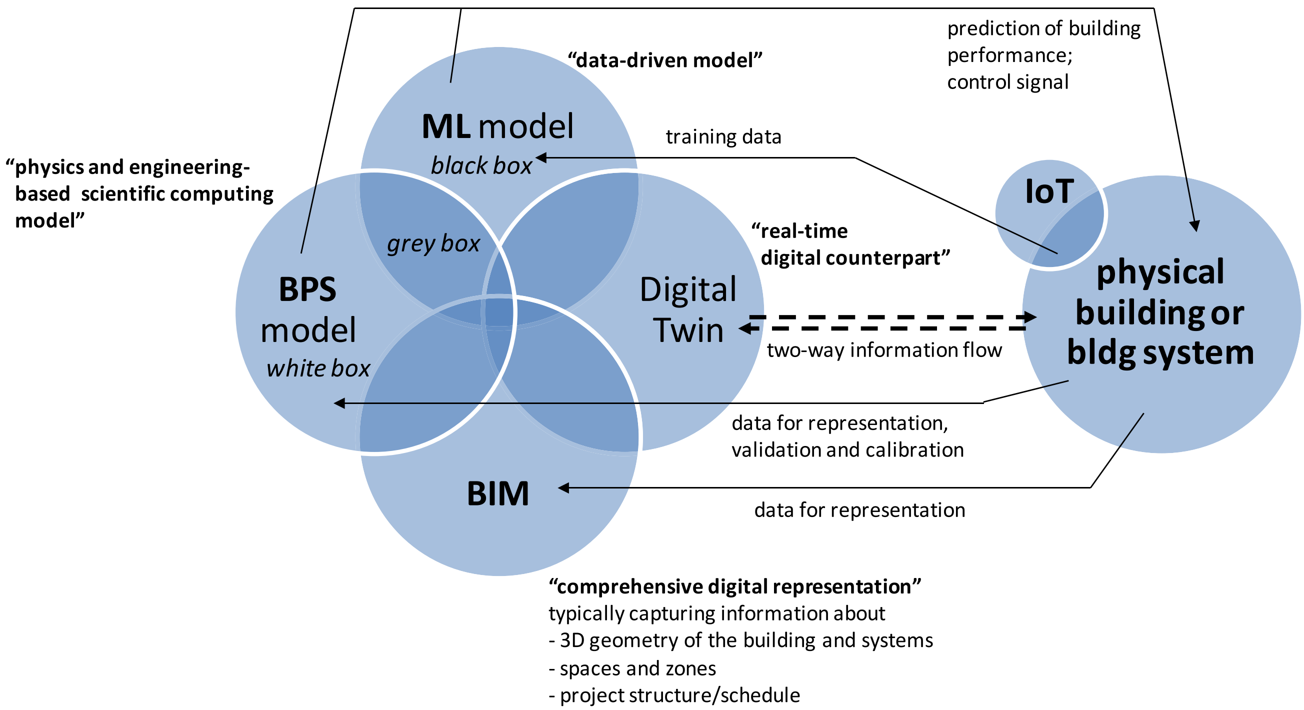

Fig. 3. Conceptual overlaps and differences between building performance simulation (bps)models, machine learning (ML) models, Digital Twins and building information models (BIM), in relation to physical buildings and systems.

- Regarding the differences between BPS models and ML models, it must be noted that many classical building simulation tools, even when they explicitly describe the physics of heat and mass transfer, may still have black box components included – for instance many tools include fan or pump performance curves that are in fact based on empirical measurement and regression analysis. Another phenomenon on the overlap between BPS and ML is the use of data generated by BPS for training of (supervised) ML models, which then become meta- or surrogate models that emulate the simulation models. Such *meta*-models have been around for a while already and have been demonstrated to replicate simulation results well within their training domain, as for instance in the work by Eisenhower *et al.* [53] or de Wilde *et al.* [54]. By their very nature a surrogate model that has been trained on simulation data only approximates the results from the underlying BPS output on which it is trained; it will never be a better predictor than the BPS model itself. However, a significant advantage of surrogate models is that they allow real-time interaction, which is highly advantageous for building control and design applications [55,56].
- Whilst concepts such as Machine Learning, Digital Twins, Cyber-Physical Systems and the Internet of Things enjoy a high uptake in both industry and academia, they are partly new 'buzzwords' that overlap with existing concepts and terminology. This is particularly true in the area of building control and management systems, which for a long time have been based on observing states of the physical building, processing data through some form of model, and following up with control actions. The notions of Digital Twins and Cyber-Physical System thus appear to in part new terms for existing approaches such as feed forward controllers and embedded systems. Similarly, supervised Machine Learning models overlap with the existing term of 'black box models'.

In terms of strengths and limitations, literature yields a range of observations and beliefs. Note that the sample studied pertains to work that spans both building performance simulation and the adjacent domains so all observations and beliefs belong to papers on the interface of these areas. Perceived strengths and limitations of Building Performance Simulation are presented in Table 2a and Table 2b; those for Machine Learning are listed in Table 3a and Table 3b, for Digitals Twins in Table 4a and Table 4b, and for the Internet of Things Table 5a and Table 5b. No literature was found showing perceived strengths or limitations of Cyber Physical Systems, whereas any discussion on Data Mining reflected the same issues that are already discussed for machine learning. The perceived strengths and limitations reported in Tables 2 to 5 reflect the comments and views emerging from in literature, grouped in themes as arising from the search. The findings, which emerge from literature, have been loosely grouped in main themes; however these were not pre-conceived units of assessment. The terms used to name these themes have been taken from the literature that was suveryed.

The content of Tables 2 to 5 can be summarized as follows:

- Perceived strengths of building performance simulation include the ability to deliver precise results, prediction of performance where there is no historical data, excellent comparison of design alternatives, the ability to propagate uncertainties in models, the capacity to gain deep understanding and insights, and the ability to work from information that is readily available. Perceived limitations of

Table 2a
Perceived strengths of building performance simulation in the building simulation literature.

BPS strengths	References
precise results	Chakrabarty *et al* [57];Geyer and Singravel [58]; Westermann *et al* [59]
predicts performance where there is no historical data	Deb and Schlueter [60]; Foucquier *et al* [61]; de Wilde *et al* [54]
comparison of design alternatives	Papadopoulos *et al* [62]; Singh *et al* [63]; Ward *et al* [64];
propagation of uncertainties	Singh *et al* [63]
good for gaining understanding, insight	Li and Yao [65]; Sanyal *et al* [66]; Yezioro *et al* [67]
based on information that is readily available	Chakrabarty *et al* [57]

Table 2b
Perceived limitations of building performance simulation in the building simulation literature.

BPS limitations	References
computational cost: high information needs, long computing times, large numbers of parameters	Brownlee and Wright [68]; Deb and Schlueter [60]; Edwards *et al* [69]; Ekici *et al* [70]; Foucquier *et al* [61]; Geyer and Singravel [58]; Kim *et al* [71]; Mazuroski *et al* [72]; Nourkojouri *et al* [73]; Papadopoulos *et al* [62]; Sanyal *et al* [66]; Sha *et al* [74]; Sha *et al* [55]; Sharif and Hammad [75]; Singh *et al* [63]; Tang *et al* [76]; Thrampodoulis *et al* [77]; Westermann *et al* [59]
complexity, expertise required	Chakraborty and Elzarka [29]; Nutkiewicz *et al* [78]; Sha *et al* [74]; Sharif and Hammad [75]; Thrampodoulis *et al* [77]
performance gap predicted/measured	Carstens *et al* [79]; Causone *et al* [80]; Deb and Schlueter [60]; Deng and Chen [81]; Edwards *et al* [69]; Mo *et al* [82]; Papadopoulos *et al* [62]; Singh *et al* [63]; Vollmer *et al* [83]; Ward *et al* [64]
calibration being expensive	Carstens *et al* [79]; Kim *et al* [71]
limited attention to inter-building dynamics and district/urban scale	Nutkiewicz *et al* [78]; Sanchez *et al* [84]
requires many assumptions and simplifications	Martinez-Comesaña *et al* [85]; Mazuroski *et al* [72]; Vollmer *et al* [83]
non-linearity and numerical stiffness can be challenging	Chakrabarty *et al* [57]
often based on simplistic control	Le *et al* [86]
not suitable for fast controller response	Dey *et al* [87]
less suitable for existing buildings	Chakraborty and Elzarka [29]
not suitable for architects	Nourkojouri *et al* [73]
impact of moisture and latent heat on thermal predictions sometimes problematic	Foucquier *et al* [61]

Table 3a
Perceived strengths of machine learning in the building simulation literature.

ML strengths	References
fast computation times, rapid feedback, suitable for optimization processes	Chakraborty and Elzarka [29]; Deb and Schlueter, [60]; Ekici *et al* [70]; Foucquier *et al* [61]; Geyer and Singravel [58]; Papadopoulos *et al* [62]; Sha *et al* [74]; Singh *et al* [63]; Westermann *et al* [59]
good for anomaly detection	de Wilde *et al* [54]
accuracy and realism	Ekici *et al* [70]; Fu and Miller [88]; Papadopoulos *et al* [62]; Vollmer *et al* [83]
aggregation at city scale	Roth *et al* [89]
growing data set for training available	Liguori *et al* [90]
good for benchmarking	Veiga *et al* [91]
suitability for architects	Singh *et al* [92]
no statements found on: comparison of design alternatives; propagation of uncertainties; understandability and insights	not applicable

building performance simulation are first and foremost a high computational costs, which is linked to high information needs, long computation times, and the use of a large number of model parameters. Further limitations are complexity and high expertise requirements, the challenge of the performance gap between predicted and measured performance, the expense of calibration, some issues with limited cover of inter-building dynamics and the district/urban scale, the need for many assumptions and simplifications, challenges where models are non-linear and numerically stiff, a reliance on simple control, unsuitability to model fast controllers, being less suitable for existing buildings, issues with use by architects, and issues in terms of handling moisture and latent heat in thermal predictions.

- Perceived strengths of Machine Learning are the fast computation times, which enable rapid feedback and support for optimization, the ability to be used for anomaly detection, accuracy and realism, the

Table 3b
Perceived limitations of machine learning in the building simulation literature.

ML limitations	References
limited to range of training data; scalability and generalization issues	Deng and Chen [81]; Geyer and Singravel [58]; Westermann *et al* [59]
ability to identify anomaly source	de Wilde *et al* [54]
additional performance gap	Singh *et al* [63]
ignores uncertainties in measured data	Carstens *et al* [79]
often overlooks occupant behaviour	Fu and Miller [88]
lack of building retrofit data	Thrampodoulis *et al* [77]
data sets often incomplete or with errors	Chakrabarty *et al* [57]; Liguori *et al* [90]
large training data sets required	Foucquier *et al* [61]
no statements found on: computational cost; complexity, expertise required; representation of advanced control; suitability for architects	not applicable

Table 4a
Perceived strengths of digital twins in the building simulation literature.

DT strengths	References
allows virtual testing	De Gaetani *et al* [93]
may represent a set of BPS models	Bass *et al* [94]; Brennenstuhl *et al* [95]; Buckley *et al* [96]; Marchione and Ruperto [97]
enables predictive maintenance	Hosamo *et al* [98]
suitability for architects	Kalantari *et al* [99]
no statements found on: comparison of design alternatives; propagation of uncertainties; understandability and insights	not applicable

Table 4b
Perceived limitations of digital twins in the building simulation literature.

DT limitations	References
emerging technology, 'imperfect'	Cai *et al* [100]
no statements found on: computational cost; complexity, expertise required; performance gap issues; calibration needs; representation of advanced control; suitability for architects	not applicable

Table 5a
Perceived strengths of the internet of things in the building simulation literature.

IoT strengths	References
may use data from mobile systems	Wu *et al* [101]
input for smart systems	Aste *et al* [102]; Tagliabue *et al* [103]
more efficient and reliable than traditional monitoring systems	Gilani and O'Brien [104]
no statements found on: precision of results; comparison of design alternatives; propagation of uncertainties; understandability and insights	not applicable

Table 5b
Perceived limitations of the internet of things in the building simulation literature.

IoT limitations	References
requires link with BIM data	Tagliabue *et al* [103]
requires link with occupant behaviour	Delinchant *et al* [105]
security and privacy issues	Noye *et al* [106]; Wang *et al* [107]
no statements found on:	not applicable
computational cost	
complexity, expertise required	
performance gap issues	
representation of advanced control	
suitability for architects	
applicability to existing buildings	

capacity to aggregate building performance on a city scale, the ability to capitalize on a growing set of training data, and support for benchmarking. Perceived limitations of Machine Learning are the limitation to the range of the training data, and the related scalability and generalization issues, some weakness in identifying the source of anomalies, the creation of an additional performance gap where surrogate models are used, the ignorance of uncertainties in training data, a tendency to overlook occupant behaviour, a general lack to retrofit training data, and dependency on training sets that need to be large but which also often are incomplete or contain errors.

- Perceived strengths of Digital Twins are, beyond a general attractiveness of the concept to the building industry, the ability to incorporate digital testing, the representation of a set of simulation models, and prospect of predictive maintenance, and some ability to support architectural design. Perceived limitations of Digital Twins are that this is an 'imperfect', emerging technology.
- Perceived strengths of the Internet of Things, in the context of building performance studies, are the ability to use data from mobile systems, the link to smart systems, and especially the higher efficiency and reliability when compared to traditional monitoring systems. Perceived limitations of the Internet of Things are a need to link with BIM data, the mapping to occupant behaviour, and security and privacy issues.

It must be stressed that these strengths and limitations stem from literature, and are *perceived* strengths and limitations only. They reflect views of authors that have been found in literature, and not by the author of this article. These strengths and limitations may be subjective, and subject to criticism. Terms such as precise, good, expensive, and suitable have an inherent value judgement. Some of the perceived strengths and limitations run into clear issues. Table 2a lists as a strength of building performance simulation that it is typically "based information that is readily available"; however many modellers will disagree with this view and will have experienced long and deep searches for specific info. Other views even go against common understanding. For instance, Table 2b lists the views that building performance simulation is "less suitable for existing buildings" and "not suitable for architects". Again, these views are not uncontested. See for instance the work of Kramer *et al.* [108] who apply thermal and moisture simulation to a museum that resides in a building from the 17th century; or Attia *et al.* [109] and Alsaadani and Bleil de Souza [110] who present a much more nuanced view on the use of simulation by architects. Similar critiques may apply to other views listed in these tables. The emergence of contestable notions show the dangers inherent in discursive comparison of building performance simulation to the new domains and the need for these to be corrected by empirical work. Note that a certain notion may be voiced by a number of authors, but that this not necessarily shows that this notion is correct. Even so, the results represented in Tables 2 to 5 show the notions about building simulation and the emerging digital domains.

Literature shows some tensions and competition between different concepts, with some authors such as Geyer and Singravel [58] and Kathirgamanathan *et al* [111] voicing the belief that Machine Learning models will replace traditional physics-based BPS models. Others take a different view; for instance Roth *et al.* [89] are of the opinion that "*supervised machine learning models (…) are not a silver bullet for energy prediction, despite the recent rapid advancements in these tools. These models are often difficult to interpret and provide little physical explanatory power*". The choice between ML and BPS models may less straightforward than sometimes claimed; an interesting observation is made by Arroyo *et al.* [112] in the context of building management and control, where they note that often "*authors use either one approach or another based on their expertise and justify their choice by highlighting the strengths of their approach and stressing the weaknesses of the other*". Deb and Schlueter [60] make the point that there is an opportunity to integrate black box, grey box and white box models in a common frame that exploits the benefits of each approach.

Another observation is that some of the views and notions fail to take into account developments outside the building performance sphere. For instance, Frank *et al* [113] discuss computational speed of Machine Learning in the wider computational science and engineering realm, cautioning that '*in spite of recent success, ML is still at its infancy.*' Giacomini *et al* [114] give further depth to the concepts of computational credibility, high-fidelity and reduced computational costs in the wider context of numerical methods for computational engineering. Gómez-Carmona *et al* [115] have done deeper work on the computational cost of machine learning within the context of the Internet of Things. Classic works like Oberkampf and Roy [8] also deal with matters such as the philosophy of accuracy and validation. However, even if the notions and views reported here may be subject to critique, these views still represent important notions in literature and help to understand how the relation between building performance simulation and the emerging digital domains is perceived. Such notions show the dangers inherent in discursive comparison of building performance simulation to emerging domains; they ought to be corrected by further empirical work.

5. Work at the interface of the emerging domains and building performance simulation

With the emergence of the new digital areas of work, it is interesting to observe what work takes place on the interface of these new areas and building performance simulation. For each of the different areas, main themes have been drawn from literature. Table 6 lists the main themes that are explored on the interface between building simulation and Machine learning; Table 7 the main themes on the interface with Digital Twins, Table 8 with Cyber-Physical Systems, Table 9 with the Internet of Things, and Table 10 with Data Mining.

In general, there are 22 themes. Two of these show up across the different areas: urban/district modelling, and control systems. Urban and district models clearly represent a higher complexity than individual buildings and hence benefit from *meta*-models, new data collection methods, and novel ways of constructing models. However, it is noted that collecting data remains challenging, and that often traditional simulation is needed to fill in missing information; see for instance Veiga *et al.* [91]. Work on building control systems looks at the building (sub) system and is driven by the need to operate buildings efficiently. The traditional approaches here are rule-based control (RBC) and model predictive control (MPC), with Machine Learning opening up a new route of data-driven control (DDC); see for instance Kathirgamanathan *et al* [111]. This theme is closely related to data sensing and therefore has logical links to work on the Internet of Things, Data Mining and Machine Learning, whereas by nature a digital controller and physical system form a Cyber-Physical System. Within this theme, Arroyo *et al.* [112] observe that "*the control and machine learning communities keep evolving independently with a radically different notation natively adopted to formulate the same problem*" in reference to MPC and DDC approaches. Surrogate models are often suggested as a solution, however these seem

Table 6
Research themes on the interface of Building Performance Simulation and Machine Learning.

Theme	Related articles
meta and surrogate models	Brownlee and Wright [68]; Edwards *et al* [69]; Ekici *et al* [70]; Geyer and Singaravel [58]; Lin *et al* [116]; Mazuroski *et al* [72]; Sharif and Hammad [75]; Singh *et al* [63]; Thrampoulidis *et al* [77]; Veiga *et al* [91]; Verma *et al* [117]; Westermann *et al* [59]
urban and district models	Ang *et al* [118]; Li and Yao [65]; Nutkiewicz *et al* [78]; Roth *et al* [89]; Sánchez and Marijuan [84]; Vázquez-Canteli *et al* [119]
occupant behaviour	Amayri *et al* [120]; Arslan *et al* [121]; Causone *et al* [80]; Dai *et al* [122]; Deng and Chen [81]; Elkhoukhi *et al* [123]; Fu and Miller [88]; Huchuk *et al* [124]; Mo *et al* [82]; Ryu and Moon [125]; Silva *et al* [126]; Ward *et al* [64]; Yi [127]
MPC and other building control systems	Aftab *et al* [128]; Arroyo *et al* [112]; Drgoňa *et al* [129]; Le *et al* [86]; Noye *et al* [106]; Sha *et al* [55]; Taheri and Razban [130]; Vollmer *et al* [83]
model validation and calibration	Carstens *et al* [79]; Chakrabarty *et al* [57]; Grieu *et al* [131]; Kim *et al* [71]; Martinez-Comesaña *et al* [85]; Sanyal *et al* [66]; Yezioro *et al* [67]
data quality	Liguori *et al* [90]
retrofit	Deb and Schlueter [60]; Sharif and Hammad [75]; Thrampoulidis *et al* [77]
comfort models	Wu *et al* [132]; Zhang *et al* [133]
fault detection and diagnosis	Dey *et al* [87]; de Wilde *et al* [54]
ensemble models	Chakraborty and Elzarka [29]; Papadopoulos *et al* [62]
weather and climate data	Hosseini *et al* [134]; Kalyanam and Hoffmann [135]
daylight	Ayoub [136]; Lee *et al* [137]; Nourkojouri *et al* [73]
seismic performance	Tang *et al* [76]

Table 7
Research themes on the interface of Building Performance Simulation and Digital Twins.

Theme	Related articles
urban and district models	Bass *et al* [94]; Buckley *et al* [96]; Chaturvedi and Kolbe [138]; Garrison and New [139]; Marchione and Ruperto [97]; Sibilla and Abanda [140]; Simonsson *et al* [141]
interoperability	De Gaetani *et al* [93]; Porsani *et al* [142]; Zhao *et al* [143]
MPC and other building control systems	Brennenstuhl *et al* [95]; Kontes *et al* [144]; Narayanan *et al* [145]
model validation and calibration	Chakrabarty *et al* [57]; Chong *et al* [146]
facility management	Cai *et al* [100]; Hosamo *et al* [98]; Jafari *et al* [147]; Jafari *et al* [148]; Garrison and New [139]
retrofit and renovation	Duch-Zebrowska and Zielonko-Jung [149]; Massafra *et al* [150]
smart and intelligent buildings	Pavon *et al* [151]
design and optimization	Lydon *et al* [152]; Togashi *et al* [153]
education and training	Johra *et al* [154]

Table 8
Research themes on the interface of Building Performance Simulation and Cyber-Physical Systems.

Theme	Related articles
MPC and other building control systems	Böke *et al* [155]; Li *et al* [156]; Karbasi and Farhadi [157]; Schmidt *et al* [158]; Ye *et al* [159]
monitoring	Bonci *et al* [160]
policy and decision-making	Meuer *et al* [161]

to also have their limitations; for instance Kontes *et al.* [144] state that *"from our experience with real-building control experiments, the assumption that the original simulation model (and thus also the surrogate model derived from it) can actually predict all states of the real building, under all different combinations of weather conditions and occupant actions, usually does not hold in practice. Instead, what is required is a continuous calibration process, for example as defined in the digital twin paradigm"*. In the literature on the interface between building simulation and Digital Twin, it has been noted that many papers just drop the term 'Digital Twin' without discussing the concept in detail, or engaging with the notion of Digital Twins as discussed in section 3. Often the term only appears only once or twice; such papers have been excluded from the search for research themes.

Table 9
Research themes on the interface of Building Performance Simulation and the Internet of Things.

Theme	Related articles
occupant behaviour	Brambilla *et al* [162]; Gilani and O'Brien [104]; Mataloto *et al* [163]; Wu *et al* [101]
MPC and other building control systems	Aste *et al* [102]; Kathirgamanathan *et al* [111]; Lee and Yeo [164]; Tagliabue *et al* [103]
urban and district models	Chagnon-Lessard *et al* [165]; Wang *et al* [107]
policy and decision-making	Delinchant *et al* [105]

Table 10
Research themes on the interface of Building Performance Simulation and Data Mining.

Theme	Related articles
urban and district models	Capozzoli *et al* [166]
occupant behaviour	Amasyali and El-Gohary [167]; Amasyali and El-Gohary [168]; Chen and Soh [169]; Darakdjian *et al* [170]; Krishnan *et al* [171]; Ouf *et al* [172]; Sun *et al* [173]
MPC and other building control systems	May-Ostendorp *et al* [174]; Sun *et al* [173]
model validation and calibration	Hiyama [175]; Rouchier [176]
retrofit	Simpson *et al* [177]
categorization and clustering	Abdelrahman *et al* [178]; Bhatia *et al* [179]; Capozzoli *et al* [166]; Gunay *et al* [180]; Morbitzer *et al* [181]; Yang *et al* [182]; Zhan *et al* [183]

6. Discussion and conclusions

This article explores how the traditional domain of building performance simulation relates to the emerging digital areas of Artificial Intelligence / Machine Learning, Digital Twins, Cyber-Physical Systems, the Internet of Things and Data Mining. A study of definitions and concepts shows that some of these novel fields are clearly distinct from simulation and operate on quite different concepts: Artificial Intelligence can be linked to smart and intelligent buildings but leaves open how the smartness or intelligence is achieved. Cyber-Physical Systems can be distinguished by the obvious interaction between digital and physical counterparts, and the Internet of Things is a about connecting devices and people and enabling data collection. There is however conceptual overlap and risk of confusion where the novel trends involve digital models. This concerns building performance simulation models, Machine Learning models, Digital Twins and to some extent Building Information Models. Important conceptual differences manifest themselves in the data flow and how tight such models are connected to a physical building or building system. The conceptual difference between building simulation model and (supervised) Machine Learning model aligns with the more traditional view of white box versus black box models, with hybrid (grey box) models occupying a middle ground.

The literature that straddles the domain of building (performance) simulation and emerging areas of Machine Learning, Digital Twins, and Internet of Things gives a wide range of views on the benefits and challenges for these various concepts, as detailed in Tables 2 to 5:

There is significant research activity on the interfaces between building performance simulation and the novel digital domains, across a range of themes; see Tables 6 to 10 for details. Two areas that gain attention across the spectrum are urban/district models, and building management and control systems.

The literature about perceived benefits and limitations of building performance simulation and Machine Learning appears to have some bias towards the advantages of the latter. More limitations are listed for building simulation, however this may also reflect the maturity of the concept. There are some authors who claim that Machine Learning will replace building simulation. However, some critical comments are in order:

- Benefits and limitations voiced in the literature are often beliefs. There is a sparsity of hard empirical tests that quantify the performance of one type of model against the other. The use case for any empirical tests is crucial, and needs to include aspects of uncertainty such as inherent in design. Further research is needed in order to move beyond subjective positions.
- Fast computation times and rapid feedback are often cited as advantages of Machine Learning approaches over building performance simulation. Whilst this may be true for strict running time of the algorithms involved, the situation becomes much more complex and unclear when the modelling effort is included. There are few, if any, accounts that include the preparatory work needed before one is able to run a simulation or applicable machine learning algorithm. No views or information has been found on the efforts and time needed to deploy Digital Twins, Cyber-Physical Systems, Internet of Things or Data Mining.
- The engineering and physics knowledge and routines used in traditional building performance simulation models remains highly valuable and should not be neglected. First-principle models remain the tool of choice to predict the future performance of buildings that are still at the design stage, since at that point there is no data yet to train ML algorithms. Also, first-principle models retain a high instructional value for educational and upskilling purposes.
- Efforts in the area of system control and management are starting to explore hybrid solutions, where the benefits of particular algorithms are combined. This may be a path forwards for building performance prediction in general. Whilst the data-driven Machine Learning models have a high promise, their weakness is the identification of their application range which needs to correlate to their training domain. It may be that these models could play a stronger role if applied to typical building components such as boilers, chillers and AHUs. Such component models might then be integrated with traditional BPS models in the same way that components are combined in for instance the Modelica language, and indeed have been integrated in other tools such as TRNSYS, EnergyPlus and ESP-r in the past.

Whilst the literature review shows the many views, overlaps and intersection between building performance simulation and the emergent domains, a general attempt to order them according to applicability is as follows:

- Building performance simulation is an appropriate approach to predict future building performance, especially at a design stage of a building, or where use conditions are changing. It also is appropriate in an educational setting, where it can be used to operationalize physical and engineering knowledge.
- Supervised Machine Learning is an appropriate technique where data is available to train a mathematical black box model. This is especially true for existing buildings, and is highly useful for fault detection. The speed of Machine Learning algorithms also have distinct advantages in contexts where real-time interaction is required, such as in design support or building control systems. Reinforcement learning is an important approach for the control of buildings.
- Digital Twins become useful where a model (digital counterpart) can be used to manage a real building (physical counterpart), for instance for optimal control or predictive maintenance.
- The Internet of Things is an approach that supports data collection and monitoring, and can support Machine Learning, Digital Twins as well as traditional building simulation model calibration.

Whilst the new emergent digital domains enjoy significant traction and help to energize efforts in the building research field, in part they offer new terminology for existing technology. Machine Learning has long been around as 'black box' approaches, with the main difference between the past and present being the increased availability of training data. Buildings and building systems with any form of digital control are by nature Cyber-Physical Systems; here the new term mainly provides a fresh world view rather than a fundamental shift. Other terms such as Digital Twins indeed offer new concepts to link building simulation with actual buildings and other digital technology. However it is important that the definition of concepts like Digital Twins is properly understood; if the building simulation community simply considers any simulation model to be a digital twin then the intended technological progress will be missed. With a good conceptual understanding and differentiation in place, the new digital domains provide new impetus for the building research community, which is to be welcomed.

Disclosure statement

No potential conflict of interest was reported by the author.

Data availability statement

Data sharing is not applicable to this article as the work is based on a literature review. All sources used are listed in the references section.

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.

Data availability

No data was used for the research described in the article.

References

[1] Accenture, 2021. *Technology Vision 2021*. Dublin: Accenture. Available from : https://www.accenture.com/gb-en/insights/technology/technology-trends-2021 [Accessed 27/01/2023].

[2] Gartner, 2022. *Top Strategic Technology Trends for 2022*. Stamford: Gartner. Available from: https://www.gartner.co.uk/en/information-technology/insights/top-technology-trends [Accessed 27/01/2023].

[3] Grieves, 2015. *Digital Twin: manufacturing excellence through virtual factory replication.* White Paper, available from: https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/DELMIA/PDF/Whitepaper/DELMIA-APRISO-Digital-Twin-Whitepaper.pdf [Accessed 27/01/2023].

[4] A. Bolton, M. Enzer, J. Schooling, et al., The Gemini Principles: Guiding values for the national digital twin and information management framework, Centre for Digital Built Britain and Digital Framework Task Group, 2018.

[5] NSF Workshop on Cyber-Physical Systems: research motivation, techniques and roadmap. Austin, Texas, 16-17 October 2006. Available from: https://web.archive.org/web/20080517071555/http://varma.ece.cmu.edu/cps/ [Accessed 20/04/2023].

[6] C.J. Anumba, N. Roofigari-Esfahan (Eds.), Cyber-Physical Systems in the Built Environment, Springer International Publishing, Cham, 2020.

[7] G. Augenbroe, 2019, Time for reflection. Retirement presentation to graduate students, Atlanta, August 2019.

[8] W.L. Oberkampf, C.J. Roy (Eds.), Verification and Validation in Scientific Computing, Cambridge University Press, 2010.

[9] I. Beausoleil-Morrison, Fundamentals of building performance simulation, Routledge, Abingdon, 2021.
[10] Clarke, Energy simulation in building design, Butterworth-Heinemann, Oxford, 2001.
[11] J. Hensen, R. Lamberts (Eds.), Building performance simulation for design and operation, 2nd edition., Routledge, Abingdon, 2019. ISBN 9781138392199.
[12] de Wilde, P., 2018. *Building Performance Analysis*. Wiley Blackwell. ISBN: 978-1-119-34192-5.
[13] H. Wang, Z. Zhai, Advances in building simulation and computational techniques: a review between 1987 and 2014, Energy and Buildings 128 (2016) 319–335, https://doi.org/10.1016/j.enbuild.2016.06.080.
[14] S. Oh, J. Haberl, Origins of analysis methods used to design high-performance commercial buildings: whole-building energy simulation, Science and Technology for the Built Environment 22 (1) (2016) 118–137, https://doi.org/10.1080/23744731.2015.1063958.
[15] G. Augenbroe, Trends in building simulation, Building and Environment 37 (8/9) (2002) 891–902, https://doi.org/10.1016/S0360-1323(02)00041-0.
[16] J. Spitler, Building performance simulation: the now and not yet, HVAC&R Research 12 (3a) (2006) 549–551, https://doi.org/10.1080/10789669.2006.10391194.
[17] J. Clarke, A vision for building performance simulation: a position paper prepared on behalf of the IBPSA Board, Journal of Building Performance Simulation 8 (2) (2015) 39–43, https://doi.org/10.1080/19401493.2015.1007699.
[18] P. de Wilde, The gap between predicted and measured energy performance of buildings: a framework for investigation, Automation in Construction 41 (2014) 40–49, https://doi.org/10.1016/j.autcon.2014.02.009.
[19] L. Zhang, Y. Pan, X. Wu, M. Skibniewski, Artificial Intelligence in Construction Engineering and Management, Springer, Singapore, 2021.
[20] K. Murphy, Machine Learning: a Probablistic Perspective, MIT Press. ISNB, Cambridge, 2012, pp. 978–0262018029.
[21] J. Viola, Y. Chen, Digital-Twin-Enabled Smart Control Engineering, Springer, Cham, 2023.
[22] R. Alur, Principles of Cyber-Physical Systems, MIT Press, Cambridge, 2015.
[23] A. Gyasi-Agyei, Wireless Internet of Things: Principles and Practice, World Scientific Publishing, Singapore, 2020.
[24] M. Zaki, W. Meira, Data Mining and Machine Learning: Fundamental Concepts and Algorithms, Cambridge University Press, Cambrigde, 2020.
[25] Royal Society, Machine learing: the power and promise of computers that learn by example, The Royal Society, London, 2017.
[26] T. Derek, J. Clements-Croome, What do we mean by intelligent buildings? Automation in Construction 6 (5-6) (1997) 395–400.
[27] Bishop, C., 2007. *Pattern Recognition and Machine Learning.* Singapore: Springer. ISBN: 978-0387310732.
[28] Brunton, S.. B. Noack and P.Koumoutsakos, 2022. Machine Learning for Fluid Mechanics. *Annual Review of Fluid Mechanics*, .52 (1), 477-508. https://www.annualreviews.org/doi/10.1146/annurev-fluid-010719-060214.
[29] D. Chakraborty, H. Elzarka, Advanced machine learning techniques for building performance simulation: a comparative analysis, Journal of Building Performance Simulation 12 (2) (2019) 193–207, https://doi.org/10.1080/19401493.2018.1498538.
[30] M. Wiering, M. van Otterlo (Eds.), Reinforcement Learning: State-of-the-Art, Springer, Heidelberg, 2012.
[31] M. Ayoub, A review on machine learning algorithms to predict daylighting inside buildings, Solar Energy 202 (2020) 249–275, https://doi.org/10.1016/j.solener.2020.03.104.
[32] H. Sun, H. Burton, H. Huang, Machine learning applications for building structural design and performance assessment: state-of-the-art review, Journal of Building Engineering 33 (2021), 101816, https://doi.org/10.1016/j.jobe.2020.101816.
[33] H.D. Deng, Fannon and M. Eckelman, Predictive modelling for US commercial building energy use: a comparison of existing statistical and machine learning algorithms using CBECS microdata, Energy and Buildings 163 (2018) 34–43, https://doi.org/10.1016/j.enbuild.2017.12.031.
[34] Seyedzadeh, S., F. Rahimian, I Glesk and M. Roper, 2018. Machine learning for estimation of building energy consumption and performance: a review. *Visualization in Engineering*, 6, 5. https://doi.org/s40327-018-0064-7.
[35] S. Fathi, R. Srinivasan, A. Fenner, S. Fathi, Machine learning applications in urban building energy performance forecasting: a systematic review, Renewable and Sustainable Energy Reviews 133 (2020), 110287, https://doi.org/10.1016/j.rser.2020.110287.
[36] Centre for Digital Built Britain, 2022. *Digital Twins* [online]. Cambridge, United Kingdom. Available from: https://www.cdbb.cam.ac.uk/research/digital-twins [Accessed 14/04/2022].
[37] Brilakis, I., Y. Pan, A. Borrmann, H. Mayer, F. Rhein, C. Vos, A. Pettinato and S. Wagner, 2019. *Built Environment Digital Twinning.* Munich: Technical University of Munich and Siemens AG. Available from: https://publications.cms.bgu.tum.de/reports/2020_Brilakis_BuiltEnvDT.pdf. [Accessed 27/01/2023].
[38] M. Singh, E. Fuenmayor, E. Hinchy, Y. Qiao, N. Murray, D. Devine, Digital Twin: origin to future, Applied System Innovation 4 (2021) 36, https://doi.org/10.3390/asi4020036.
[39] R. Ganguli, S. Adhikari, The digital twin of discrete dynamic systems: initial approaches and future challenges, Applied Mathematical Modelling 77 (2020) 1110–1128, https://doi.org/10.1016/j.apm.2019.09.036.
[40] P. Korenhof, V. Blok, S. Kloppenburg, Steering representations – towards a critical understanding of Digital Twins, Philosophy & Technology 34 (2021) 1751–1773, https://doi.org/10.1007/s13347-021-00484-1.
[41] F. Tao, H. Zhang, A. Liu, A. Nee, Digital Twin in industry: state-of-the-art, IEEE Transactions on Industrial Informatics 15 (4) (2019) 2405–2415, https://doi.org/10.1109/TII.2018.2873186.
[42] J. Delgado, L. Oyedele, Digital Twins for the built environment: learning from conceptual and process models in manufacturing, Advanced Engineering Informatics 49 (2021), 101332, https://doi.org/10.1016/j.aei.2021.101332.
[43] National Science Foundation, 2022. *Cyber-Physical Systems (CPS)*. Washington, USA. Availabe from: www.nsf.gov [Accessed 15/04/2022].
[44] IBM, 2022. *Internet of Things webpage*. Armonk, USA. Availabe from: www.ibm.com [Accessed 06/05/2022].
[45] M. Jia, A. Komeily, Y. Wang, R. Srinivasan, Adopting the Internet of Things for the development of smart buildings: a review of enabling technologies and applications, Automation in Construction 101 (2019) 111–126, https://doi.org/10.1016/j.autcon.2019.01.023.
[46] S. Tang, D. Shelden, C. Eastman, P. Pishdad-Bozorgi, X. Gao, A review of building information modeling (BIM) and the internet of things (IoT) devices integration: present status and future trends, Automation in Construction 101 (2019) 127–139, https://doi.org/10.1016/j.autcon.2019.01.020.
[47] IBM, 2022. *Data Mining webpage*. Armonk, USA. Availabe from: www.ibm.com [Accessed 15/04/2022].
[48] M. Pena, F. Biscarri, J. Guerro, I. Monedero, C. Leon, Rule-based system to detect energy efficiency anomalies in smart buildings, a data mining approach, Smart Systems With Applications 56 (2016) 242–255, https://doi.org/10.1016/j.eswa.2016.03.002.
[49] C. Fan, F. Xiao, Z. Li, J. Wang, Unsupervised data analytics in mining big building operational data for energy efficiency enhancement: a review, Energy and Buildings 159 (2018) 296–308, https://doi.org/10.1016/j.enbuild.2017.11.008.
[50] Y. Zhao, C. Zhang, Y. Zhang, Z. Wang, J. Li, A review of data mining technologies in building energy systems: load prediction, pattern identification, fault detection and diagnosis, Energy and Built Environment 1 (2020) 149–164, https://doi.org/10.1016/j.enbenv.2019.11.003.
[51] B. Qolomany, A. Al-Fuqaha, A. Gupta, D. Benhaddou, S. Alwajidi, J. Qadir, A. Fong, Leveraging machine learning and big data for smart buildings: a comprehensive survey, IEEE Access 7 (2019) 90316–90356, https://doi.org/10.1109/ACCESS.2019.2926642.
[52] Oracle, 2022. Big Data webpage. Reading, UK. Available form www.oracle.com/uk/big-data/what-is-big-data/ [Accessed 07/06/2022].
[53] B. Eisenhower, Z. O'Neill, S. Narayanan, V. Fonoberov, I. Mezić, A methodology for meta-model based optimization in building energy models, Energy and Buildings 47 (2012) 292–301, https://doi.org/10.1016/j.enbuild.2011.12.001.
[54] P. de Wilde, C. Martinez-Ortiz, D. Pearson, I. Beynon, M. Beck, N. Barlow, Building simulation approaches for the training of automated data analysis tools in building energy management, Advanced Engineering Informatics 27 (4) (2013) 457–465, https://doi.org/10.1016/j.aei.2013.05.001.
[55] H. Sha, M. Moujahed, D. Qi, Machine learning-based cooling load prediction and optimal control for mechanical ventilative cooling in high-rise buildings, Energy and Buildings 242 (2021), 110980, https://doi.org/10.1016/j.enbuild.2021.110980.
[56] M. Zorn, N. Catunda, L. Claus, N. Kobylinska, M. Frey, T. Wortmann, Replacing energy simulations with surrogate models for design space exploration, Bauphysik 44 (6) (2022) 311–316, https://doi.org/10.1002/bapi.202200034.
[57] A. Chakrabarty, E. Maddalena, H. Qiao, C. Laughman, Scalable Bayesian optimization for model calibration: case study on coupled building and HVAC dynamics, Energy and Buildings 253 (2021), 111460, https://doi.org/10.1016/j.enbuild.2021.111460.
[58] P. Geyer, S. Singaravel, Component-based machine learning for performance prediction in building design, Applied Energy 228 (2018) 1439–1453, https://doi.org/10.1016/j.apenergy.2018.07.011.
[59] P. Westermann, M. Welzel, R. Evins, Using a deep temporal convolutional network as a building energy surrogate model that spans multiple climate zones, Applied Energy 278 (2020), 115563, https://doi.org/10.1016/j.apenergy.2020.115563.
[60] C. Deb, A. Schlueter, Review of data-driven energy modelling techniques for building retrofit, Renewable and Sustainable Energy Reviews 144 (2021), 110990, https://doi.org/10.1016/j.rser.2021.110990.
[61] A. Foucquier, S. Robert, F. Suard, L. Stéphan, A. Jay, State of the art in building modelling and energy performances prediction: a review, Renewable and Sustainable Energy Reviews 23 (2013) 272–288, https://doi.org/10.1016/j.rser.2013.03.004.
[62] S. Papadopoulos, E. Azar, W. Woon, C. Kontokosta, Evaluation of tree-based ensemble learning algorithms for building energy performance simulation, Journal of Building Performance Simulation 11 (3) (2018) 322–332, https://doi.org/10.1080/19401493.2017.1354919.
[63] M. Singh, S. Singaravel, R. Klein, P. Geyer, Quick energy prediction and comparison of options at the early design stage, Advanced Engineering Informatics 46 (2020), 101185, https://doi.org/10.1016/j.aei.2020.101185.
[64] R. Ward, C. Wong, A. Chong, R. Choudhary, S. Ramasamy, A study on the transferability of computational models of building electricity load patterns across climatic zones, Energy and Buildings 237 (2021), 110826, https://doi.org/10.1016/j.enbuild.2021.110826.
[65] X. Li, R. Yao, Modelling heating and cooling energy demand for building stock using a hybrid approach, Energy and Buildings 235 (2021), 110740, https://doi.org/10.1016/j.enbuild.2021.110740.
[66] J. Sanyal, J. New, R. Edwards, L. Parker, Calibrating building energy models using supercomputer trained machine learning agents, Concurrency and

Computation: Practice and Experience 26 (2014) 2122–2133, https://doi.org/10.1002/cpe.3267.

[67] A. Yezioro, B. Dong, F. Leite, An applied artificial intelligence approach towards assessing building performance simulation tools, Energy and Buildings 40 (2008) 612–620, https://doi.org/10.1016/j.enbuild.2007.04.014.

[68] A. Brownlee, J. Wright, Constrained, mixed-integer and multi-objective optimization of building designs by NSGA-II with fitness approximation, Applied Soft Computing 33 (2015) 114–126, https://doi.org/10.1016/j.asoc.2015.04.010.

[69] R. Edwards, J. New, L. Parker, B. Cui, J. Dong, Constructing large scale surrogate models from big data and artificial intelligence, Applied Energy 202 (2017) 685–699, https://doi.org/10.1016/j.apenergy.2017.05.155.

[70] B. Ekici, O. Turkcan, M. Turrin, I. Sariyildiz, M. Tasgetiren, Optimising high-rise buildings for self-sufficiency in energy consumption and food production using artificial intelligence: case of Europoint Complex in Rotterdam, Energies 15 (2022) 660, https://doi.org/10.3390/en15020660.

[71] Y. Kim, K. Ahn, C. Park, Issues of application of machine learning models for virtual and real-life buildings, Sustainability 8 (2016) 543, https://doi.org/10.3390/su8060543.

[72] W. Mazuroski, J. Berger, R. Oliveira, N. Mendes, An artitifical intelligence-based method to efficiently bring CFD to building simulation, Journal of Building Performance Simulation 11 (5) (2018) 588–603, https://doi.org/10.1080/19401493.2017.1414880.

[73] H. Nourkojouri, N.S. Shafavi, M. Tahsildoost, Z.S. Zomorodian, Development of a machine-learning framework for overall daylight and visual comfort assessment in early design stages, Journal of Daylighting 8 (2) (2021) 270–283.

[74] H. Sha, P. Xu, Z. Yang, Y. Chen, J. Tang, Overview of computational intelligence for building energy system design, Renewable and Sustainable Energy Reviews 108 (2019) 76–90, https://doi.org/10.1016/j.rser.2019.03.018.

[75] S. Sharif, A. Hammad, Developing surrogate ANN for selecting near-optimal building energy renovation methods considering energy consumption, LCC and LCA, Journal of Building Engineering 25 (2019), 100790, https://doi.org/10.1016/j.jobe.2019.100790.

[76] Q.i. Tang, J.i. Dang, Y. Cui, X. Wang, J. Jia, Machine learning-based fast seismic risk assessment of building structures, Journal of Earthquake Engineering 26 (15) (2022) 8041–8062.

[77] E. Thrampoulidis, G. Mavromatidis, A. Lucchi, K. Orehounig, A machine learning-based surrogate model to approximate optimal building retrofit solutions, Applied Energy 281 (2021), 116024, https://doi.org/10.1016/j.apenergy.2020.116024.

[78] A. Nutkiewicz, Z. Yang, R. Jain, Data-driven Urban Energy Simulation (DUE-S)): a framework for integrating engineering simulation and machine learning methods in a multi-scale urban energy modeling workflow, Applied Energy 225 (2018) 1176–1189, https://doi.org/10.1016/j.apenergy.2018.05.023.

[79] H. Carstens, X. Xia, S. Yadavalli, Low-cost energy meter calibration method for measurement and verification, Applied Energy 188 (2017) 563–575, https://doi.org/10.1016/j.apenergy.2016.12.028.

[80] F. Causone, S. Carlucci, M. Ferrando, A. Marchenko, S. Erba, A data-driven procedure to model occupancy and occupant-related electric load profiles in residential buildings for energy simulation, Energy and Buildings 2020 (2019), 109342, https://doi.org/10.1016/j.enbuild.2019.109342.

[81] Z. Deng, Q. Chen, Reinforcement learning of occupant behavior model for cross-building transfer learning to various HVAC control systems, Energy and Buildings 238 (2021), 110860, https://doi.org/10.1016/j.enbuild.2021.110860.

[82] H. Mo, H. Sun, J. Liu, S. Wei, Developing window behavior models for residential building using XGBoost algorithm, Energy and Buildings 205 (2019), 109564, https://doi.org/10.1016/j.enbuild.2019.109564.

[83] M. Vollmer, M. Langer, F. Banihashemi, H. Harter, D. Kierdorf, W. Lang, Prediction of window handle state using machine learning, Bauphysik 46 (6) (2020) 352–359, https://doi.org/10.1002/bapi.202000032.

[84] V. Sánchez, A. Marijuan, Integrated model concept for district energy management optimisation platforms, Applied Thermal Engineering 196 (2021), 117233, https://doi.org/10.1016/j.applthermaleng.2021.117233.

[85] M. Martinez-Comesaña, L. Febrero-Garrido, E. Granada-Álvarez, J. Martínez-Torres, S. Martínez-Mariño, Heat loss coefficient estimation applied to existing buildings through machine learning models, Applied Sciences 10 (2020) 8968, https://doi.org/10.3390/app10248968.

[86] K. Le, R. Bourdais, H. Guéguen, From hybrid model predictive control to logical control for shading system: a support vector machine approach, Energy and Buildings 84 (2014) 352–359, https://doi.org/10.1016/j.enbuild.2014.07.084.

[87] M. Dey, S. Rana, S. Dudley, Smart building creation in large scale HVAC environments through automated fault detection and diagnosis, Future Generation Computer Systems 108 (2020) 950–966, https://doi.org/10.1016/j.future.2018.02.019.

[88] C. Fu, C. Miller, Using Google Trends as a proxy for occupant behavior to predict building energy consumption, Applied Energy 310 (2022), 118343, https://doi.org/10.1016/j.apenergy.2021.118343.

[89] J. Roth, A. Martin, C. Miller, R. Jain, SynCity: using open data to create a synthetic city of hourly building energy estimates by integrating data-driven and physics-based methods, Applied Energy 280 (2020), 115981, https://doi.org/10.1016/j.apenergy.2020.115981.

[90] A. Liguori, R. Markovic, T. Dam, J. Frisch, C. van Treeck, F. Causone, Indoor environment data time-series reconstruction using autoencoder neural networks, Building and Environment 191 (2021), 107623, https://doi.org/10.1016/j.buildenv.2021.107623.

[91] R. Veiga, A. Veloso, A. Melo, R. Lamberts, Application of machine learning to estimate building energy use intensities, Energy and Buildings 249 (2021), 111219, https://doi.org/10.1016/j.enbuild.2021.111219.

[92] M. Singh, C. Deb, P. Geyer, Early-stage desgin support combining maching learning and building information modelling, Automation in Construction 136 (2022) 2022, https://doi.org/10.1016/j.autcon.2022.104147.

[93] C. De Gaetani, M. Mert, F. Migliaccio, Interoperability analyses of BIM platforms for construction management, Applied Sciences 10 (2020) 4437, https://doi.org/10.3390/app10134437.

[94] B. Bass, J. New, W. Copeland, Potential energy, demand, emissions, and cost savings distributions for buildings in a utility's service area, Energies 14 (2021) 132, https://doi.org/10.3390/en14010132.

[95] M. Brennenstuhl, D. Lust, D. Pietruschka, D. Schneider, Demand-side management based power-to-heat and power-to-gas optimimization strategies for PV and wind self-consumption in a residential building cluster, Energies 14 (2021) 6712, https://doi.org/10.3390/en14206712.

[96] N. Buckley, G. Mills, S. Letellier-Duchesne, K. Benis, Designing an energy-resilient neighbourhood using an urban energy model, Energies 14 (2021) 4445, https://doi.org/10.3390/en14154445.

[97] P. Marchione, F. Ruperto, Prototyping a Digital Twin – a case study of a 'U-shaped' military building, International Journal of Energy Production and Management 7 (1) (2022) 83–94, https://doi.org/10.2495/EQ-V7-N1-83-94.

[98] H. Hosamo, P. Svennevig, K. Svidt, D. Han, H. Nielsen, A Ditigal Twin predictive maintenance framework of air handling units based on automatic fault detection and diagnosis, Energy and Buildings 261 (2022), 111988, https://doi.org/10.1016/j.enbuild.2022.111988.

[99] S. Kalantari, S. Pourjabar, T. Xu, J. Kan, Developing and user-testing a 'Digital Twins' prototyping tool for architectural design, Automation in Construction 135 (2022), 104140, https://doi.org/10.1016/j.autcon.2022.104140.

[100] H. Cai, F. Khayatian, P. Heer, Experiment strategy for evaluating advanced building energy management system, Journal of Physics: Conference Series 2042 (1) (2021) 012030.

[101] W. Wu, B. Dong, Q. Wang, M. Kong, D. Yan, J. An, Y. Liu, A novel mobility-based approach to derive urban-scale building occupant profiles and analyze impacts on building energy consumption, Applied Energy 278 (2020), 115656, https://doi.org/10.1016/j.apenergy.2020.115656.

[102] N. Aste, M. Manfren, G. Marenzi, Building automation and control systems and performance optimization: a framework for analysis, Renewable and Sustainable Energy Reviews 75 (2017) 313–330, https://doi.org/10.1016/j.rser.2016.10.072.

[103] L. Tagliabue, F. Cecconi, S. Rinaldi, A. Ciribini, Data driven indoor air quality prediction in educational facilities based on IoT network, Energy and Buildings 236 (2021), 110782, https://doi.org/10.1016/j.enbuild.2021.110782.

[104] S. Gilani, W. O'Brien, Review of current methods, opportunities, and challenges for in-situ monitoring to support occupant modelling in office spaces, Journal of Building Performance Simulation 10 (5) (2017) 444–470, https://doi.org/10.1080/19401493.2016.1255258.

[105] B. Delinchant, H. Dang, H. Vu, D. Nguyen, Massive arrival of low-cost and low-consuming sensors in buildings: toward new building energy services, IOP Conference Series: Earth and Environmental Sciences 307 (2019), 012006, https://doi.org/10.1088/1755-1315/307/1/012006.

[106] S. Noye, R. Martinez, L. Canieletto, M. De Carli, A. Aguirre, A review of advanced ground source heat pump control: artificial intelligence for autonomous and adaptive control, Renewable and Sustainable Energy Reviews 153 (2022), 111685, https://doi.org/10.1016/j.rser.2021.111685.

[107] K. Wang, Y. Zhao, R. Gangadhari, Z. Li, Analyzing the adoption challenges of Internet of Things (IoT) and Artificial Intelligence (AI) for smart cities in China, Sustainability 13 (2021) 10983, https://doi.org/10.3390/su131910983.

[108] R.P. Kramer, A.W.M. van Schijndel, H.L. Schellen, The Importance of integrally simulating the building, HVAC and control systems, and occupant's impact for energy predictions of buildings including temperature and humidity control: validated case study museum Hermitage, Amsterdam. *Journal of Building Performance Simulation* 10 (3) (2017) 272–293.

[109] S. Attia, J. Hensen, L. Beltrán, A. De Herde, Selection criteria for building performance simulation tools: contrasting architects' and engineers needs, Journal of Building Performance Simulation 5 (3) (2012) 155–169, https://doi.org/10.1080/19401493.2010.549573.

[110] S. Alsaadani, C. Bleil De Souza, Performer, consumer or expert? A critical review of building performance simulation training paradigms for building design decision-making, Journal of Building Performance Simulation 12 (3) (2019) 289–307.

[111] A. Kathirgamanathan, M. De Rosa, E. Mangina, D. Finn, Data-driven predictive control for unlocking building energy flexibility, Renewable and Sustainable Energy Reviews 135 (2021), 110120, https://doi.org/10.1016/j.rser.2020.110120.

[112] J. Arroyo, C. Manna, F. Spiessens, L. Helsen, Reinforcement model predictive control (RL-MPC) for building energy management, Applied Energy 309 (2022), 118346, https://doi.org/10.1016/j.apenergy.2021.118346.

[113] M. Frank, D. Drikakis, V. Charissis, Machine-Learning methods for computational science and engineering, Computation 8 (2020) 15, https://doi.org/10.3390/computation8010015.

[114] M. Giacomini, K. Veroy, P. Díez, Special issue on credible high-fidelity and low-cost simulations in computational engineering, International Journal for Numerical Methods in Engineering 121 (2020) 5151–5152, https://doi.org/10.1002/nme.6551.

[115] O. Gómez-Carmona, D. Casado-Manislla, F. Kraemer, D. López-de-Ipiña, J. García-Zubia, Exploring the computational cost of machine learning at the edge for

human-centric internet of things, Future Generation Computer Systems 112 (2020) 670–683, https://doi.org/10.1016/j.future.2020.06.013.
[116] Y. Lin, L. Zhao, X. Liu, W. Yang, X. Hao, L. Tian, Design optimization of a passive building with green roof through machine learning and group intelligent algorithm, Buildings 11 (2021) 192, https://doi.org/10.3390/buildings11050192.
[117] A. Verma, S. Prakash, A. Kumar, ANN-based energy consumption prediction model up to 2050 for a residential building: towards sustainable decision making, Environmental Progress and Sustainable Engineering 40 (2020) e13544.
[118] Y. Ang, Z. Berzolla, S. Letellier-Duchesne, V. Jusiega, C. Reinhart, UBEM.io: a web-based framework to rapidly generate urban building energy models for carbon reduction technology pathways, Sustainable Cities and Society 77 (2022), 103534, https://doi.org/10.1016/j.scs.2021.103534.
[119] J. Vázquez-Canteli, S. Ulyanin, J. Kämpf, Z. Nagy, Fusing TensorFlow with building energy simulation for intelligent energy management in smart cities, Sustainable Cities and Society 45 (2019) 243–257, https://doi.org/10.1016/j.scs.2018.11.021.
[120] M. Amayri, A. Arora, S. Ploix, S. Bandhyopadyay, Q. Ngo, V. Badarla, Estimating occupancy in heterogeneous sensor environment, Energy and Buildings 129 (2016) 46–58, https://doi.org/10.1016/j.enbuild.2021.111758.
[121] M. Arslan, C. Cruz, D. Ginhac, Understanding occupant behaviors in dynamic environments using ObiDE framework, Building and Environment 166 (2019), 106412, https://doi.org/10.1016/j.buildenv.2019.106412.
[122] X. Dai, J. Liu, Z. Zhang, A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings, Energy and Buildings 223 (2020), 110159, https://doi.org/10.1016/j.enbuild.2020.110159.
[123] H. Elkhoukhi, Y. NaitMalek, M. Bakhouya, A. Berouine, A. Kharbouch, F. Lachhab, M. Hanifi, D, ElOuadghiri and M. Essaaidi, A platform architecture for occupancy detection using stream processing and machine learning approaches, Concurrency and Computation: Practice and Experience 32 (2019) 35651, https://doi.org/10.1002/cpe.5651.
[124] B. Huchuk, S. Sanner, W. O'Brien, Comparison of machine learning models for occupancy prediction in residential buildings using connected thermostat data, Building and Environment 160 (2019), 106177, https://doi.org/10.1016/j.buildenv.2019.106177.
[125] S. Ryu, H. Moon, Development of an occupancy prediction model using indoor environmental data based on machine learning techniques, Building and Environment 107 (2016) 1–9, https://doi.org/10.1016/j.buildenv.2016.06.039].
[126] J. Silva, M. Amayri, S. Ploix, P. Reignier, C. Silva, Cooperative and interactive learning to estimate human behaviours for energy applications, Energy and Buildings 258 (2022), 111727, https://doi.org/10.1016/j.enbuild.2021.111727.
[127] H. Yi, Visualized co-simulation of adaptive human behavior and dynamic building performance: an agent-based model (ABM) and artificial intelligence (AI) approach for smart architectural design, Sustainability 12 (2020) 6672, https://doi.org/10.3390/su12166672.
[128] M. Aftab, C. Chen, C. Chau, T. Rahwan, Automated HVAC control with real-time occupancy recognition and simulation-guided model predictive control in low-cost embedded system, Energy and Buildings 154 (2017) 141–156, https://doi.org/10.1016/j.enbuild.2017.07.077.
[129] J. Drgoňa, D. Picard, M. Kvasnica, L. Helsen, Approximate model predictive building control via machine learning, Applied Energy 218 (2018) 199–216, https://doi.org/10.1016/j.apenergy.2018.02.156.
[130] S. Taheri, A. Razban, Learing-based CO2 concentration prediction: application to indoor air quality control using demand-controlled ventilation, Building and Environment 205 (2021), 108164, https://doi.org/10.1016/j.buildenv.2021.108164.
[131] S. Grieu, O. Faugeroux, A. Traoré, B. Claudet, J. Bodnar, Artificial intelligence tools and inverse methods for estimating the thermal diffusivity of building materials, Energy and Buildings 43 (2011) 543–554, https://doi.org/10.1016/j.enbuild.2010.10.020.
[132] Z. Wu, N. Li, J. Peng, H. Cui, P. Liu, H. Li, X. Li, Using an ensemble maching learning methodology-bagging to predict occupant's thermal comfort in buildings, Energy and Buildings 173 (2018) 117–127, https://doi.org/10.1016/j.enbuild.2018.05.031.
[133] W. Zhang, F. Liu, R. Fan, Improved thermal comfort modelling for smart buildings: a data analytics study, Electrical Power and Energy Systems 103 (2018) 634–643, https://doi.org/10.1016/j.ijepes.2018.06.026.
[134] M. Hosseini, A. Bigtashi, B. Lee, A systematic approach in constructing typical meteorological year weather files using machine learning, Energy and Buildings 226 (2020), 110375, https://doi.org/10.1016/j.enbuild.2020.110375.
[135] R. Kalyanam, S. Hoffmann, A novel approach to enhance the generalization capability of the hourly solar diffuse horizonal irradiance models on diverse climates, Energies 13 (2020) 4868, https://doi.org/10.3390/en13184868.
[136] M. Ayoub, Self-organizing profiles to characterize representative temporal settings for daylight simulations, Solar Energy 214 (2021) 248–267, https://doi.org/10.1016/j.solener.2020.11.051.
[137] J. Lee, M. Boubekri, F. Liang, Impact of building design parameters on daylighting metrics using an analysis, prediction, and optimization approach based on statistical learning technique, Sustainability 11 (2019) 1474, https://doi.org/10.3390/su11051474.
[138] K. Charturvedi, T. Kolbe, A requirement analysis on extending semantic 3D city models for supporting time-dependent properties, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 4 (4) (2019) W9, https://doi.org/10.5194/isprs-annals-IV-4-W9-19-2019.
[139] E. Garrison, J. New, Quality control methods for advanced metering infrastructure data, Smart Cities 4 (2021) 195–203, https://doi.org/10.3390/smartcities4010012.
[140] M. Sibilla, F. Abanda, Multi-criteria decision making optimization framework for positive energy blocks for cities, Sustainability 14 (2022) 446, https://doi.org/10.3390/su14010446.
[141] J. Simonsson, K. Atta, G. Schweiger, W. Birk, Experiences from city-scale simulation of thermal grids, Resources 10 (2021) 10, https://doi.org/10.3390/resources10020010.
[142] G. Porsani, K. Del Valle, A. de Lersundi, S.-O. Gutiérrez, C. Bandera, Interoperability between Building Information Modelling (BIM) and Building Energy Model (BEM), Applied Sciences 11 (2021) 2167, https://doi.org/10.3390/app11052167.
[143] L. Zhao, H. Zhang, Q. Wang, H. Wang, T. Dede, Digital-Twin based evaluation of Nearly Zero-Energy Building for existing buildings based on scan-to-BIM, Advances in Civil Engineering 2021 (2021) 1–11.
[144] G. Kontes, G. Giannakis, V. Sánchez, P. de Augustin-Camacho, A. Romero-Amorrotu, N. Panagiotidou, D. Rovas, S. Steiger, C. Mutschler, G. Gruen, Simulation-based evaluation and optimization of control strategies in buildings, Energies 11 (2018) 3376, https://doi.org/10.3390/en11123376.
[145] M. Narayanan, A. Ferreira de Lima, A. Dantas, W. Commerell, Development of a couples TRNSYS-MATLAB simulation framework for model predictive control of integrated electrical and thermal residential renewable energy systems, Energies 13 (2020) 5761, https://doi.org/10.3390/en13215761.
[146] A. Chong, Y. Gu, H. Jia, Calibrating building energy simulation models: a review of the basics to guide future work, Energy and Buildings 253 (2021), 111533, https://doi.org/10.1016/j.enbuild.2021.111533.
[147] M.A. Jafari, E. Zaidan, A. Ghofrani, K. Mahani, F. Farzan, Improving building energy footprint and asset performance using Digital Twin technology, IFAC PapersOnLine 53 (3) (2020) 386–391.
[148] M.A. Jafari, A. Ghofrani, E. Zaidan, A. Abulibdeh, Improving building energy footprint and asset performance using digital twin technology, Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction 174 (2) (2021) 57–65.
[149] Duch-Zebrowska, P. and K. Zielonko-Jung, 2021. Integrating digital twin technology into large panel system estates retrofit projects. *Urban Planning*, 6 (4), 164-171. https://doi.org/10.17645/up.v6i4.4464.
[150] A. Massafra, G. Predari, R. Gulli, Towards Digital Twin driven cultural heritage management: a HBIM-based workflow for energy improvement of modern buildings, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 46 (5) (2022) W1, https://doi.org/10.5194/isprs-archives-XLVI-5-W1-2022-149-2022.
[151] R. Pavon, M. Alberti, A. Alvarez, I. Carrasco, Use of BIM-FM to transform large conventional public buildings into efficient and smart sustainable buildings, Energies 14 (2021) 3127, https://doi.org/10.3390/en14113127.
[152] G. Lydon, S. Caranovic, I. Hischier, A. Schlueter, Coupled simulation of thermally active building systems to support a digital twin, Energy and Buildings 202 (2019), 109298, https://doi.org/10.1016/j.enbuild.2019.07.015.
[153] E. Togashi, M. Miyata, Y. Yamamoto, The first world championship in cybernetic building optimization, Journal of Building Performance Simulation 13 (3) (2020) 391–408, https://doi.org/10.1080/19401493.2020.1741685.
[154] H. Johra, E.A. Petrova, L. Rohde, M.Z. Pomianowski, Digital Twins of building physics experimental laboratory setups for effective e-learning, Journal of Physics: Conference Series 2069 (1) (2021) 012190.
[155] J. Böke, U. Knaack, M. Hemmerling, State-of-the-art of intelligent building envelopes in the context of intelligent technical systems, Intelligent Buildings International 11 (2) (2019) 27–45, https://doi.org/10.1080/17508975.2018.1447437.
[156] P. Li, Y. Lu, D. Yan, J. Xiao, H. Wu, Scientometric mapping of smart building research: towards a framework of human-cyber-physical system (HCPS), Automation in Construction 129 (2021), 103776, https://doi.org/10.1016/j.autcon.2021.103776.
[157] A. Karbasi, A. Farhadi, A cyber-physical system for building automation and control based on a distributed MPC with an efficient method for communication, European Journal of Control 61 (2021) 151–170, https://doi.org/10.1016/j.ejcon.2021.04.008.
[158] M. Schmidt, A. Schülke, A. Venturi, R. Kurpatov, E. Henriquez, Cyber-Physical System for energy efficient stadium operation: methodology and experimental validation, ACM Transactions on Cyber-Physical Systems 1 (1) (2017) 1, https://doi.org/10.48550/arXiv.1807.05059.
[159] Z. Ye, Z. O'Neill, F. Hu, Hardware-based emulator with deep learning model for building energy control and prediction based on occupancy sensor's data, Information 12 (2021) 499, https://doi.org/10.3390/info12120499.
[160] A. Bonci, A. Carbonari, A. Cucchiarelli, L. Messi, M. Pirani, M. Vaccarini, A cyber-physical system approach for building efficiency monitoring, Automation in Construction 102 (2019) 68–85, https://doi.org/10.1016/j.autcon.2019.02.010.
[161] J. Meuer, M. Toetzke, C. Nakhle, S. Windeck, A typology of digital building technologies: implications for policy and industry, IOP Conference Series: Earth and Environmental Sciences 323 (1) (2019) 012053.
[162] A. Brambilla, C. Candido, I. Hettiarachchi, L. Thomas, O. Gocer, K. Gocer, M. Mackey, N. Biloria, T. Alizadeh, S. Sarkar, The potential of harnessing real-time occupancy data for improving energy performance of activity-based workplaces, Energies 15 (2021) 230, https://doi.org/10.3390/en15010230.
[163] B. Mataloto, D. Calé, K. Carimo, J. Ferreira, R. Resende, 3D IoT system for environmental and energy consumption system, Sustainability 13 (2021) 1495, https://doi.org/10.3390/su13031495.

[164] J. Lee, M. Yeo, Condensation control with occupancy activity and effectively mitigate condensation in unheated spaces by real-time sensor control strategy, Sustainability 12 (2020) 4033, https://doi.org/10.3390/su12104033.

[165] N. Chagnon-Lessard, L. Gosselin, S. Barnabé, T. Bello-Ochende, S. Fendt, S. Goers, L. Pereira da Silva, B. Schweiger, R. Simmons, A. Vandersickel, P. Zhang, Smart campuses: extensive review of the last decade of research and current challenges, IEEE Access 9 (2021) 124200–124234, https://doi.org/10.1109/ACCESS.2021.3109516.

[166] A. Capozzoli, D. Grassi, M. Piscitelli, G. Serale, Discovering knowledge from a residential building stock through data mining analysis for engineering sustainability, Energy Procedia 83 (2015) 370–379, https://doi.org/10.1016/j.egypro.2015.12.212.

[167] K. Amasyali, N. El-Gohary, Machine learning for occupant-behavior-sensitive cooling energy consumption prediction in office buildings, Renewable and Sustainable Energy Reviews 142 (2021), 110714, https://doi.org/10.1016/j.rser.2021.110714.

[168] K. Amasyali, N. El-Gohary, Hybrid approach for energy consumption prediction: coupling data-driven and physical approaches, Energy and Buildings 259 (2022), 111758, https://doi.org/10.1016/j.enbuild.2021.111758.

[169] Z. Chen, Y. Soh, Comparing occupancy models and data mining approaches for regular occupancy prediction in commercial buildings, Journal of Building Performance Simulation 10 (5/6) (2017) 545–553, https://doi.org/10.1080/19401493.2016.1199735.

[170] Q. Darakdjian, S. Billé, C. Inard, Data mining of building performance simulations comprising occupant behaviour modelling, Advances in Building Energy Research 13 (2) (2019) 157–173, https://doi.org/10.1080/17512549.2017.1421099.

[171] D. Krishnan, S. Kelly, Y. Kim, A meta-analysis review of occupant behaviour models for assessing demand-side energy consumption, Energies 15 (2022) 1219, https://doi.org/10.3390/en15031219.

[172] M. Ouf, H. Gunay, W. O'Brien, A method to generate design-sensitive occupant-related schedules for building performance simulations, Science and Technology for the Built Environment 25 (2) (2019) 221–232, https://doi.org/10.1080/23744731.2018.1514855.

[173] C. Sun, R. Zhang, S. Sharples, Y. Han, H. Zhang, Thermal comfort, occupant control behaviour and performance gap – a study of office buildings in north-east China using data mining, Building and Environment 149 (2019) 305–321, https://doi.org/10.1016/j.buildenv.2018.12.036.

[174] P. May-Ostendorp, G. Henze, B. Rajagopalan, C. Corbin, Extraction of supervisory building control rules from model predictive control of windows in a mixed mode building, Journal of Building Performance Simulation 6 (3) (2013) 199–219, https://doi.org/10.1080/19401493.2012.665481.

[175] K. Hiyama, Assigning robust default values in building performance simulation software for improved decision-making in the initial stages of building design, The Scientific World Journal 2015 (2015) 1–11.

[176] S. Rouchier, Solving inverse problems in building physics: an overview of guidelines for careful and optimal use of data, Energy and Buildings 166 (2018) 178–195, https://doi.org/10.1016/j.enbuild.2018.02.009.

[177] K. Simpson, J. Whyte, P. Childs, Data-centric innovation in retrofit: a bibliometric review of dwelling retrofit across North Western Europe, Energy and Buildings 229 (2020), 110474, https://doi.org/10.1016/j.enbuild.2020.110474.

[178] M. Abdelrahman, S. Zhan, C. Miller, A. Chong, Data science for building energy efficiency: a comprehensive text-mining driven review of scientific literature, Energy and Buildings 242 (2021), 110885, https://doi.org/10.1016/j.enbuild.2021.110885.

[179] A. Bhatia, V. Garg, P. Haves, V. Pudi, Explainable clustering using hyper-rectangles for building energy simulation data, IOP Conference Series: Earth and Environmental Science 238 (2019), 012068, https://doi.org/10.1088/1755-1315/238/1/012068.

[180] H. Gunay, Z. Shi, I. Wilton, J. Bursill, Disaggregation of commercial building end-uses with automation system data, Energy and Buildings 223 (2020), 110222, https://doi.org/10.1016/j.enbuild.2020.110222.

[181] C. Morbitzer, P. Strachan, C. Simpson, Data mining analysis of building simulation performance data, Building Services Engineering Research and Technology 25 (3) (2004) 253–267, https://doi.org/10.1191/0143624404bt098oa.

[182] Y. Yang, R. Li, T. Huang, Smart meter data analysis of a building cluster for heating load profile quantification and peak load shifting, Energies 13 (2020) 4343, https://doi.org/10.3390/en13174343.

[183] S. Zhan, Z. Liu, A. Chong, D. Yan, Building categorization revisited: a clustering-based approach to using smart meter data for building energy benchmarking, Applied Energy 269 (2020), 114920, https://doi.org/10.1016/j.apenergy.2020.114920.