Energy & Buildings 198 (2019) 170–186

Contents lists available at ScienceDirect

Energy & Buildings

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

Surrogate modelling for sustainable building design – A review

Paul Westermann*, Ralph Evins

Energy and Sustainable Cities Group Department of Civil Engineering University of Victoria 3800 Finnerty Road, Victoria BC, Canada

ARTICLE INFO

Article history:
Received 24 January 2019
Revised 15 April 2019
Accepted 26 May 2019
Available online 29 May 2019

Keywords:
Sustainable building design
Building performance simulation
Surrogate model
Meta-model
Early design
Uncertainty analysis
Sensitivity analysis
Building design optimisation

ABSTRACT

Statistical models can be used as surrogates of detailed simulation models. Their key advantage is that they are evaluated at low computational cost which can remove computational barriers in building performance simulation. This comprehensive review discusses significant publications in sustainable building design research where surrogate modelling was applied.

First, we familiarize the reader with the field and begin by explaining the use of surrogate modelling for building design with regard to applications in the conceptual design stage, for sensitivity and uncertainty analysis, and for building design optimisation. This is complemented with practical instructions on the steps required to derive a surrogate model. Next, publications in the field are discussed and significant methodological findings highlighted. We have aggregated 57 studies in a comprehensive table with details on objective, sampling strategy and surrogate model type. Based on the literature major research trends are extracted and useful practical aspects outlined.

As surrogate modelling may contribute to many sustainable building design problems, this review summarizes and aggregates past successes, and serves as practical guide to make surrogate modelling accessible for future researchers.

1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) recognizes the potential for the current building stock to stabilize or reduce its global energy use by mid-century [1]. The high performance of current building technologies and understanding of how to integrate them, make energy efficient buildings and retrofits also economically viable.

However, the building sector transforms slowly. The International Energy Agency (IEA) observed that it lags behind in the clean-energy transition as defined in the Paris Agreement [2]. One key challenge faced by the sector is that each building and retrofit is unique and has to be customized due to varying purpose, location and cultural context. Taking into account that the existing building stock of 150 billion square meters will grow by an annual rate of 3.7 billion square meters until 2026 [3] and that buildings are currently designed in a largely individual fashion by architects and engineers, facilitating and automating the design processes will be crucial to the spread of sustainable buildings.

Recent advances in machine learning paired with growing data availability are pushing the automation of analytical problems like sustainable building design [4,5]. Three fundamental types of data exist in the building domain:

(a) Building sensor data (e.g. smart meters, internet of things (IoT) sensors, building management systems)
(b) Building stock data (e.g. annual energy demand and floor area for a large set of buildings)
(c) Building simulation data (stored results of building simulation)

The first two types are particularly useful for optimising building operation [6,7], designing building-specific retrofit options [8] *(a)*, or for conducting energy mapping and building performance benchmarking in a certain geographic area covered by the building stock data *(b)* [9].

Both types of data are composed of historical observations on already existing buildings. Statistical prediction models trained on that data clearly may not be accurate for new building technologies or unique design concepts. Hence, building simulation relying on physical laws remains crucial for the design of new buildings. Its validity is not bound to observations, but instead any new design, retrofit option or building technology can be modelled.

Abbreviations: BPS, Building Performance Simulation; GP, Gaussian Process model; ANN, artificial neural network; MARS, multivariate regression splines; SVM, support vector machine; PCE, polynomial chaos expansion; RF, random forest; RBF, radial basis function; LSTM, long-short term memory network; LHS, latin hypercube sampling; DoE, design of experiments; iid, independent and ideally distributed; SA, sensitivity analysis; UA, uncertainty analysis; BDO, building design optimisation.

* Corresponding author.
E-mail addresses: pwestermann@uvic.ca (P. Westermann), revins@uvic.ca (R. Evins).

https://doi.org/10.1016/j.enbuild.2019.05.057
0378-7788/

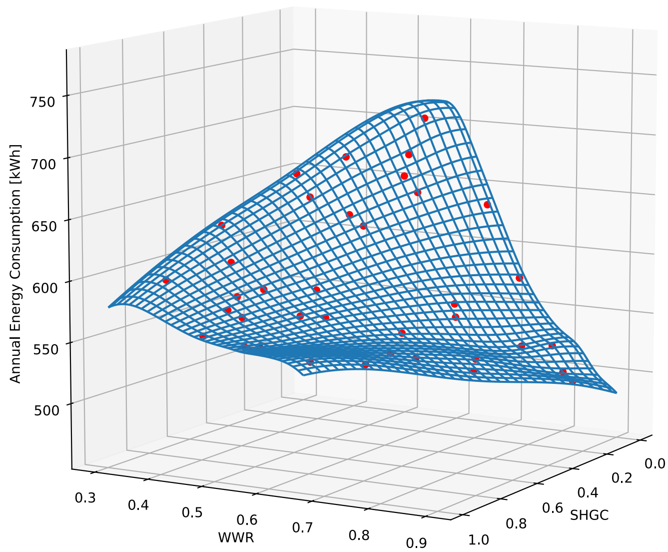

Fig. 1. Example of the application of surrogate modelling for sustainable building design evaluation. This surrogate estimates annual energy consumption based on window-to-wall ratio (WWR) and solar heat gain coefficient (SHGC). It was fitted to previously collected simulation samples (red dots) and was then evaluated at a finer resolution (every intersection of the blue mesh). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

However, current building simulation software has high computational cost and setting up a building model is time intensive [10]. Needed architects and designers do not fully integrate it into their daily work [11]. Surrogate models [12–14], or meta-models, are promising to provide building performance assessment which is physical knowledge based but much faster than simulation-based design analysis [15].

The idea of surrogate modelling is to emulate an expensive high-fidelity model, in this case a building simulation model, using a statistical model. The surrogate is trained on a small set of simulation in- and output data (*c*). Once it is validated to approximate the detailed simulation model well enough, it can be used to almost instantly predict outcomes of the high-fidelity simulation given an appropriate set of building design information.

In this work we are largely concerned with surrogates that predict aggregated design metrics (e.g. annual energy use) rather than detailed time series (e.g. hourly energy use). The process is illustrated in Fig. 1 for a problem with two inputs and one output. Here a Gaussian process model was trained to predict annual energy demand based on window-to-wall ratio and solar-heat-gain coefficient. In general (deep) artificial neural networks, support vector machines, or radial-basis function networks are common choices [16].

It is important to stress that the models studied in this review are trained on synthetic data. They are only accurate within the limitations of the simulation program and the input data used. The error induced by the simulation program as well as the modelling error of the surrogate must be balanced against the significant benefits that surrogate models bring. Both causes of errors must be addressed together, as the more accurate the simulation, the more accurate the surrogate must be to capture its behaviour. We assume that the reader is familiar with the possible errors in building simulation [17] and therefore take synthetic data as sufficient.

The review is structured as follows:

In the first two sections we familiarize the reader with the field. Section 2 covers the background on the use of surrogate modelling for the conceptual design stage (2.1), sensitivity and uncertainty analysis (2.2–2.2) and design optimization (2.4). Section 3 gives details on the steps to derive a surrogate model split into problem definition (3.1), simulation base model implementation (3.2), sampling (3.3) and surrogate model fitting (3.4). This is complemented with a list of existing surrogate modelling tools (3.5).

The reviewed literature is presented in Sections 4 and 5. First, we outline the scope of this review and refer to other reviews in related fields like energy demand forecasting (4.1). After giving an overview of the research topics (4.2) and the applied methods found (4.2.1–4.2.3), the papers are discussed thoroughly grouped by the four use cases as introduced in Section 2. We summarize findings drawn from the literature in a comprehensive list in Section 5 covering research trends and practical aspects of surrogate model fitting.

Finally, we conclude and give suggestions for future research in Section 6.

2. Surrogate models for building design

Based on existing literature (see Table 2), four stages of the building design process are found to significantly benefit from surrogate modelling:

1. Conceptual design stage
2. Sensitivity analysis
3. Uncertainty analysis
4. Optimisation

In the following section, each stage is explained in detail and the associated use of surrogate modelling explained. The section is summarized in Table 1.

2.1. Conceptual design stage

The early design or conceptual design stage happens at the very beginning of the building design process. At this point, the design is most flexible. Many parameters are roughly determined (e.g. building geometry and system types), which have a substantial impact on the final environmental and economic performance of the building [18].

Architects derive design concepts together with other stakeholders in a dynamic process. This can involve quick and drastic design changes [19] where the whole concept of the building is

Table 1
Summary on the use of surrogate models for building performance design analysis.

Analysis type	Use of surrogate
Conceptual design	• Fast feedback for design concepts; design space exploration • Fast analysis of impact of design decisions on design variability
Sensitivity analysis	• Fast variance-based global SA
Uncertainty analysis	• Fast building performance probability distribution derivation • (Model calibration)[a]
Optimisation	• Acceleration of optimisation process • Enabling gradient-based optimisation

[a] Beyond the scope of this review.

Table 2
Considered literature and its properties.

				Subject						Surrogate Model Type													Sampling Methodology				
Index (#)	Reference	First author	Year	Early Design	Sensitivity Analysis	Uncertainty Quantification	Optimisation	Comparison of meta-modelling techniques	Building Stock Analysis	ANN	RBF	Gaussian Processes	Linear Regression incl. (LASSO, stepwise, nth order terms)	MARS	SVM	Random Forest	Ensembles	Others	Favoured model (only if comparison done	# Inputs	# Outputs	Multiple climate zones?	Sampling Method	Sampling Method: adaptive (a) / static (s)	Simulation tool	# Samples (Total)	Checked impact of number of samples?
1	[77]	Gratia	2002	x									x							8	3		Doelhart matrix	s	TRNSYS	1980	
2	[26]	Ritter	2015	x					x				x						Polynomial	7	5		factorial	s	EnergyPlus	2187	
3	[25]	Geyer	2014	x									x							8	4		factorial	s	EnergyPlus	4096	x
4	[78]	Edwards	2017	x				x		x			x							156	80-90		Markov order	s	EnergyPlus	8*10^-6	
5	[79]	Yi	2015	x									x							8	2		L18 Taguchi	s	EnergyPlus	18	
6	[81]	Maltais	2017	x	x		x						x							15	2		manual	s	Daysim	1900	x
7	[84]	Korolija	2012	x									x					x		2	3		full factorial	s	EnergyPlus	23040	
8	[86]	Catalina	2013	x									x							3	1	x	full factorial	s	TRNSYS	8748	
9	[87]	Hygh	2012	x	x								x						Stepwise LR	28	3	x	full factorial	s	EnergyPlus	20000	
10	[88]	Lam	2010	x									x						LR	12	1	x	manual	s	DOE-2.1	1021	
11	[89]	Jaffal	2009	x									x							11	1	x	L12/18 Taguchi, FCC35, BB188, D68, D136, D80, D160	s	TRNSYS	60-230	x
12	[90]	Rackes	2016	x											x					38	1	x	Sobol	s	EnergyPlus	16k - 50k	
13	[91]	Romani	2015	x				x					x						Polynomial	11	2	x	D-optimal	s	TRNSYS	115-210	
14	[92]	Ascione	2017	x					x	x									ANN	29	3	x	LHS	s	EnergyPlus	500	
15	[93]	Jaffal	2017	x									x						Polynomial	10	4	x	factorial, Box-Behnken	s	TRNSYS	120	
16	[94]	Singaravel	2018	x						x									ANN	15	1	x	LHS	s	EnergyPlus	1001	
17	[23]	Ostergard	2017	x	x								x					x	Others	9	3		Sobol	s	Be10	6000	
18	[24]	Basbagill	2014	x		x														27	2		LHS, orthogonal array	s	eQuest	7623	
19	[97]	Hopfe	2012		x	x	x					x								10	2		optimality criterion	a	VA114	n/a	
20	[98]	Kim	2016		x	x		x				x							GPR	47	1		LHS	s	EnergyPlus	15	x
21	[36]	Tian	2011		x								x						SRRC	9	3		manual	s	EnergyPlus	2400	
22	[99]	Das	2014		x												x		ensemble	9	2	x	LHS	s	CONTAM	500	
23	[32]	Rivalin	2018		x	x		x				x						x	PCE	24	1		manual	s	TRNSYS	200	
24	[29]	Eisenhower	2011		x	x												x		1009	10		pseudo Monte Carlo	s	EneryPlus	5000	
25	[101]	Tsanas	2012		x			x					x			x			RF	8	2		full factorial	s	Ecotect	768	
26	[102]	de Wilde	2010		x	x							x	x					MARS	10	3		LHS	s	EnergyPlus	2000	
27	[103]	Tian	2015		x			x				x	x	x	x			x	ensemble	28	2		LHS	s	IDEAS (Modelica)	600	x
28	[105]	Chen	2017		x			x					x	x					MARS	10	3		LHS	s	EnergyPlus	5000	x
29	[21]	Hester	2017	x		x							x						Stepwise LR	20	1		full factorial	s	EnergyPlus	20000	
30	[108]	Papadopoulos	2016			x							x						MLR	6	1		LHS	s	EnergyPlus	1200	
31	[113]	Wong	2010	x			x			x									ANN	9	4		manual	s	EnergyPlus	11315	
32	[114]	Magnier	2010				x			x										20	5		LHS	s	TRNSYS	450	
33	[115]	Asadi	2014				x			x										4	3		LHS	s	TRNSYS	1045	
34	[116]	Aydin	2015				x			x									ANN	8	2		manual	s	EnergyPlus	105	
35	[107]	Eisenhower	2012			x	x								x					7-1009	2		quasi - Monte Carlo	s	EnergyPlus	5000	
36	[106]	Chen	2018		x		x								x				SVM	10	1	x	LHS	s	EnergyPlus	5610	
37	[117]	Chen	2017				x						x	x	x				SVM	9	3	x	LHS	s	EnergyPlus	5610	
38	[119]	Stavrakakis	2012				x				x									2	1		manual	s	Fluent (CFD)	63	
39	[72]	Prada	2018				x	x			x	x	x	x	x				MARS	6	2		pseudo Monte Carlo, Sobol, LHS	a	TRNSYS	50-27k	x
40	[96]	Dhariwal	2017		x		x						x							26	6		factorial	s	EnergyPlus	113-193	
41	[120]	Carreras	2016				x						x							5	2		full factorial	s	EnergyPlus	7776	
42	[121]	Gengembre	2012				x					x								8	4		expected improvement	a	R5C1	150	x
43	[123]	Tresidder	2012				x					x								10	2		optimality criterion	a	EnergyPlus	84-100	x
44	[124]	Zhang	2013				x					x								15	1		expected improvement	a	EnergyPlus	195	
45	[125]	Gilan	2016				x					x								9	2		optimality criterion	a	EnergyPlus	300-1100	x
46	[127]	Zemella	2011				x			x										5	5		optimality criterion	a	EnergyPlus	1500	x
47	[128]	Xu	2016				x						x		x					6	2		optimality criterion	a	EnergyPlus	>90	x
48	[129]	Brownlee	2015				x				x									50	3	x	optimality criterion	a	EnergyPlus	5000	
49	[130]	Wortmann	2018				x				x									13-40	3		optimality criterion	a	EnergyPlus	500	
50	[131]	Van Gelder	2013					x					x	x					MARS	14	2		LHS	s	IDEAS (Modelica)	500	
51	[132]	Cheng	2014					x						x						8	2		full factorial	s	Ecotect	768	
52	[133]	Chou	2014					x		x					x	x	x		SVR, ensemble	8	2		full factorial	s	Ecotect	768	
53	[134]	Symonds	2015					x		x	x								ANN	9	6		LHS	s	EnergyPlus	1000	
54	[15]	Van Gelder	2014					x		x	x	x		x		x			ANN	14	2		LHS	s	IDEAS (Modelica)	500	x
55	[28]	Yang	2016				x	x			x	x	x			x		x	SVM	42	2		initial: LHS, adaptive: space filling	a	EnergyPlus/Daysim	360-430	
56	[30]	Ostergard	2018		x			x		x		x	x	x	x	x			GPR	14	4		Sobol	s	BSim	32-8100	x
57	[83]	Chlela	2009					x					x						Polynomial	13	3		factorial	s	SIMBAD	32-470	x

modified. Currently, building simulation cannot keep up with the speed in the early design phase [11,20]. One reason is that setting up a simulation for one specific concept involves the manual definition of many parameters [21]. Furthermore, the simulation runtime itself is long and may interrupt the train of thought in the creativity process of the architect: ideally the program feedback time would be less than 10 seconds [22].

As a consequence of these drawbacks, researchers have derived requirements for early design tools. [23] point out that a tool for fast global design space exploration is required to quickly evaluate a large bandwidth of different initial design concepts. To reduce complexity in that process, only a few interesting parameters should be considered [20]. This may lead to facilitation of simulation, but should be balanced with simplification [19]. Lastly, Hester et al. [21] and Basbagill et al. [24] suggest early design tools should provide distributions of the performance of the building as an output. This is because at early stage many parameters are uncertain or defined as a range of possible values (*design variability*), and hence simulation results should incorporate that uncertainty.

How a surrogate model helps. Surrogate modelling simplifies the interaction between the building designer and the building simulation process in two ways. First, as surrogates are evaluated instantly (<0.1 s [15]), they are able to provide rapid point estimates [25], or distribution estimates [21] of the building performance. This enables designers to rapidly assess a design concept and explore the design space. Second, in comparison to simulation-based parametric analysis which generates discrete results, surrogate models provide continuous relationships between design variables and building performance metrics. Due to the complexity of the state-of-the art surrogate models, they are capable to capture variable interactions and extract non-linear, multi-modal behaviour [23].

Lastly, the computational layout of surrogate models is lightweight and could be embedded into existing modelling software [26].

2.2. Sensitivity analysis

Sensitivity analysis (SA) is used to rank the importance of parameters on some outcome variable [27,28]. Often it serves as a preliminary step prior to early design, uncertainty analysis (see Section 2.3) or optimisation (see Section 2.4) to reduce problem complexity. There are two different approaches: local and global methods.

In *local methods* inputs of one specific design are perturbed to approximate their partial derivatives. This provides sensitivities of inputs for the considered design. However, in a non-linear building design space sensitivities may change among different building designs [29,30] and local methods may not be suitable for general conclusions on the sensitivity of parameters.

Global methods study the influence of parameters over the whole design space. Apart from fast parameter screening methods, global analysis is computationally more demanding compared to local methods [29]. Two different methods for global analysis exist. First, the structure of the model and its parameters (or: coefficients) may be interpreted as for example in linear regression based SA. Second, in the variance-based approach a large set of simulation samples is statistically analysed. The latter is model-free and studies the impact of one parameter (*first order* sensitivity) or the combinatorial impact of multiple parameters (*total* sensitivity) on the variance of the output.

How a surrogate model helps. Local and global methods are based on simulation samples. Fast surrogate model evaluations speed up the process of sample generation [27]. They could be particularly helpful for variance-based methods which demand large number of samples. For example, the derivation of Sobol indices is sample intensive and usually limited to a small number of parameters due to computational costs [31]. In this case, the speed of a surrogate model enables an increase in the number of parameters to be studied [32].

On the other side, SA also plays a crucial role for surrogate models. Using SA, the most relevant surrogate model inputs can be determined and thus the model complexity reduced. Furthermore, when the surrogate model is very complex (as with a black-box model), SA can be used alongside the surrogate model to obtain a better understanding of the model behaviour.

2.3. Uncertainty analysis

While the purpose of SA is to quantify the effect of a change in one input on the output, uncertainty analysis (UA) studies the likeliness of a change in outputs induced by uncertain inputs [33,34]. A probabilistic view of building performance is very important. It enables quality assurance of building performance under uncertainty as for example required for energy performance contracting [32], to quantify the robustness of the design towards some exogenous variable change (e.g. climate change [35]) or to support the early design stage when many design parameters are uncertain (see Section 4.3.1.2). Sensitivity analysis may be a part of UA to screen the parameter set for the most impactful ones to reduce computational cost [31,32].

Ongoing research was reviewed in [36]. Generally, uncertainties in building design may be grouped into three categories [37]:

- Uncertainty in design parameters during the planning phase,
- uncertainty in physical parameters caused by fluctuations of material properties,
- uncertainty in scenario parameters due to assumptions of internal (e.g. usage of the building) and external (weather and climate data) conditions.

Different ways to quantify that uncertainty exist. Most commonly, uncertainty in parameters is forward propagated to receive a probability distribution of building performance like energy consumption or carbon emissions [36]. This may be done following the external or the internal approach [33].

The former assumes a building simulation model to be a black-box model. The model is used to produce a probability distribution of outcomes given a random set of possible design parameter combinations. The Monte-Carlo method may be the most popular external approach method. In the internal approach the simulation model is modified and uncertainty distributions in parameters is propagated to the model outputs [33].

To conduct the external approach the uncertainty of parameters is required. Usually, it is based on expert knowledge or results from inverse parameter uncertainty estimation if measurement data is available [38]. Bayesian calibration is a common approach for parameter uncertainty estimates and found in [38] or [39] for the building design context.

How a surrogate model helps. Surrogate models are particularly useful to accelerate the derivation of building performance distributions with the external approach which requires a significant number of simulation samples. Depending on the specific approach different numbers of simulation runs are required, varying between 60 and 80 samples for joint uncertainty propagation of all parameters in a Monte Carlo simulation [40] to larger numbers like 2^N or $2N+1$ if the impact of individual parameters and their interactions are broken down as in the factorial or differential method [33].

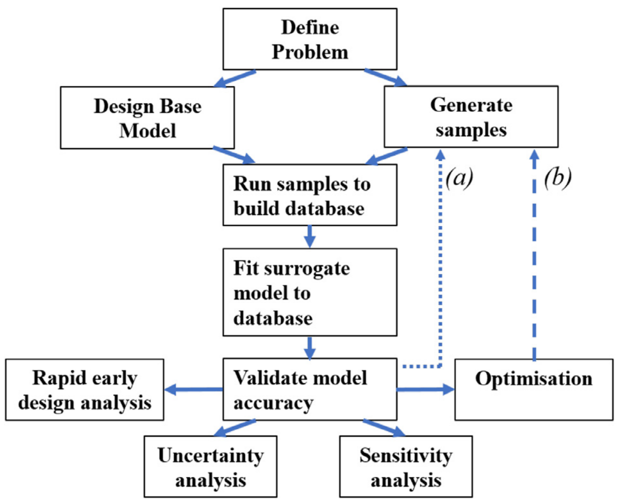

Fig. 2. Overview of the steps to derive a surrogate model. Two approaches exist. In the *sequential approach* sampling and surrogate model fitting happens subsequently. In the *iterative approach*, sampling and surrogate fitting happens iteratively where samples are picked by identifying parts of the design space with unsatisfying model accuracy *(a)* or based on an optimality criterion defined for an optimisation task *(b)*.

2.4. Design optimisation

Building design optimisation (BDO) is one of the fastest growing fields in building simulation research. It is reviewed in [41] and [42]. The goal is to find building designs which optimize a performance objective subject to constraints (e.g. comfort, system size, etc.).

In most common BDO, the fitness function to be optimized is computed using building simulation software. Different optimization algorithms exist that range from direct search, integer programming and gradient-based methods to meta-heuristics like genetic algorithms (GA). Many algorithms are introduced in the reviews above and some of them compared in [43]. The most prevalent approach is GA [41], which is easily implemented and capable of dealing with a wide variety of problems including discrete and continuous variables (e.g. heating system type versus wall thickness), multiple objectives, and discontinuities prevailing in building simulation software [44].

Following [42] an optimisation process may be split into three steps:

1) Preprocessing:	Formulation of the optimization problem; selection of optimizer
2) Optimization:	Running and monitoring of the optimizer; checking of termination criterion
3) Postprocessing:	Visualization of optimization results (e.g. Pareto front); possibly robustness evaluation

The procedure of numerical optimization is iterative, which involves many building simulation runs and may take multiple hours or days until convergence is achieved.

How a surrogate model helps. Surrogate models may speed up convergence rate of BDO. They are applied in two different ways (see Fig. 2 in [13]). In the direct surrogate-based optimisation approach the surrogate model is fitted initially and then used for optimisation.[1]The iterative approach iterates between fitting the surrogate and adding potentially optimal points to the training data.

In other engineering domains where complex simulations are imperative and too expensive without surrogate models (e.g. aerospace engineering [13,14,45]), surrogate models are well established and extensive know-how exists that is yet to be transferred to the building domain. Regarding building performance optimisation, the characteristic of surrogate models to smooth the original fitness function [46] is especially promising as building simulation results were found to have discontinuities [43]. Removing the discontinuities enables the use of optimization algorithms with potentially better performance than meta-heuristics like GA.

3. Surrogate model derivation

The steps to derive a surrogate model are shown in Fig. 2. First, the design problem and the associated design parameters have to be defined. Then the building designer implements an initial building model and picks design samples to be simulated using some sampling strategy. The parameter set defined for each sample is used to modify the base model and run building simulations with it. Results are stored in a database of inputs (design parameter values) and outputs (simulation results, e.g. annual energy consumption). Afterwards, a surrogate model is fitted to the input-output data. Last, the model is validated by computing the model accuracy. It quantifies the deviation of surrogate predictions from simulation outcomes for the same set of inputs.

Most commonly surrogate derivation happens *sequentially*. First sample locations are generated using some Design of Experiments (DoE) strategy and then the surrogate model is fitted. As the samples are defined prior to simulation and not adjusted depending on model outcomes, we refer to this approach as *static sampling*.

The *iterative* approach intertwines sample definition and surrogate model fitting. Samples are iteratively added to the database based on surrogate predictions and simulation results. Therefore, surrogate accuracy and design space complexity *(a)*, or an optimisation criterion *(b)* are evaluated to identify optimal choices for further samples.

In the following we provide details on each step in Fig. 2.

3.1. Problem definition

In the first step design parameters, the inputs to the surrogate model (also known as ‘features’), and design objectives, the outputs of the surrogate model, are defined. The selection of inputs and outputs is important as changing them at later stage may require additional high-fidelity model simulations.

Outputs are chosen based on the design objective. Similar to optimisation methods, a surrogate supports studying a specific aspect of building design, e.g. energy efficiency, which is encoded in the surrogate outputs.

The number of design parameters should be limited to circumvent the curse of dimensionality: the number of simulation samples that are needed to create an accurate surrogate of the design space grows exponentially with the number of parameters [47]. Parameters may be chosen based on the design task, or global SA if the most important parameters should be considered [48,49] (see Section 2.2). Besides deciding which parameters to choose, an associated range of possible values needs to be defined.

3.2. Base model implementation

In this step, an initial building design is implemented in physics-based building simulation software like EnergyPlus [50]. Contextual parameters, i.e. those not part of the list of design parameters, are carefully set depending on the problem (e.g. building location, climate, etc.).

[1] Some existing literature refers to model-based optimisation instead of surrogate-based optimisation. This should not be confused with simulation models used for optimization. For clarity we specifically refer to surrogate models.

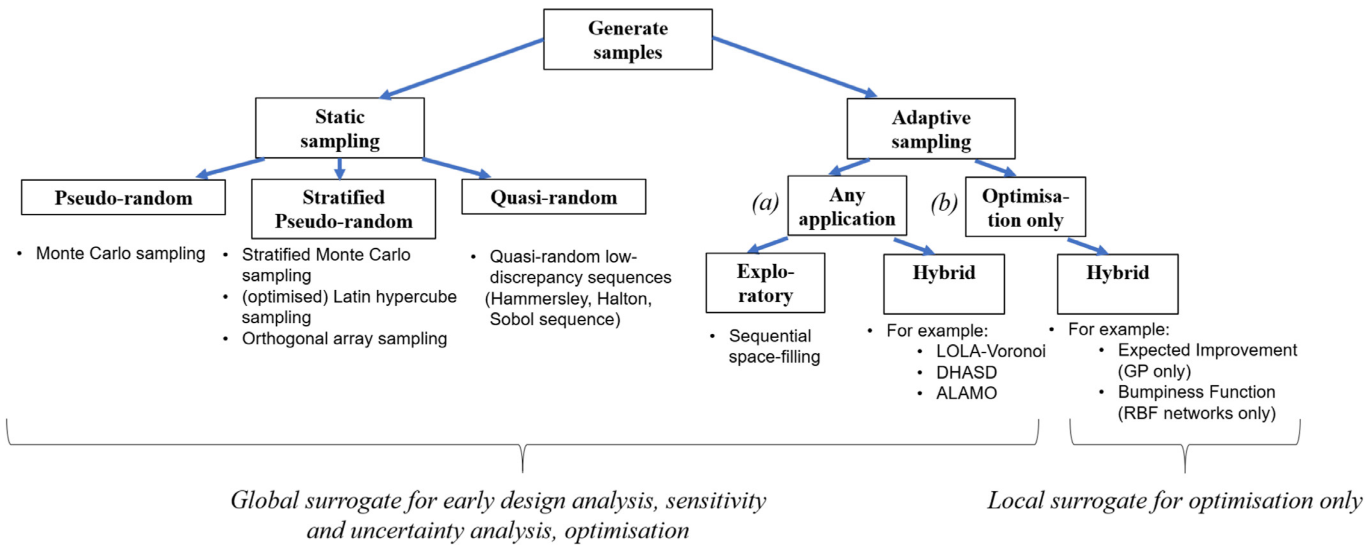

Fig. 3. Overview of different sampling methods [52].

3.3. Database generation

After the selection of parameter inputs and their range, a sampling strategy is chosen (see Fig. 3). The goal of all sampling strategies (also known as design of experiments, DoE) is to select points in the design space to maximise information gain per simulation run while minimizing sampling time. Recent reviews on DoE strategies are given by Yondo et al. [51] and Garud et al. [52].

As outlined above, two types of sampling methods exist. In static sampling all sample locations are defined in one shot prior to model fitting. This provides a global surrogate model being accurate on the whole design space. Common methods include *pseudo-random* sampling like Monte Carlo sampling, *quasi-random* sampling like Hammersly, Halton or Sobol's sequences, and *stratified pseudo-random* sampling like stratified Monte Carlo sampling, latin-hypercube sampling (LHS), or orthogonal array sampling. It is not obvious which of the provided algorithms performs best and depends on the number of variables and samples. A comparison of the methods is given in [52]. Looking at building related literature, we found that LHS is the most applied sampling scheme.

A caveat of static sampling is that it may require a lot of samples to reach an acceptable level of accuracy and therefore, adaptive sampling algorithms are sometimes favourable [51]. The goal of adaptive sampling is to balance *exploration* of under-sampled areas of the design space and *exploitation* of information gained from surrogate or simulation outcomes. Different exploration and exploitation metrics exist, called space infill criteria. They enable to identify under-sampled and complex *(a)*, or potentially optimal *(b)* areas. Before adaptive sampling is applied the surrogate is initiated on a seed of samples (found using a static sampling algorithm). While the adaptive sampling strategy *(a)* produces a *global* surrogate, *(b)* generates a surrogate model which is accurate *locally* where the design space is interesting with regard to a certain design objective. Adaptive sampling methods for global surrogate derivation *(a)* are addressed in [52] and for optimisation *(b)* in [53].

If a global surrogate is wanted, a straight-forward way of adaptive sampling is to iteratively reapply space-filling sampling (see static sampling algorithms) which is purely *explorative*. However, this may lead to inefficient sampling as it does not differentiate between complex and rather uniform areas. Therefore, taking both exploration and exploitation into account may be favourable (*hybrid*). For optimisation purposes, we only consider *hybrid* adaptive sampling methods. Pure exploitation would cause the algorithm to get stuck in local optima. An often applied sample infill criterion for optimisation is the expected improvement (EI) metric which balances model uncertainty with potential optimal performance [54].

To visualise the difference between static and adaptive sampling we derive a surrogate model (Gaussian Process) for optimisation of the Branin test function as shown in Fig. 4. We selected 20 samples using static sampling as well as adaptive sampling (path *(b)* in Fig. 3). The white dots in both plots show the locations of samples using the static approach. In case of adaptive sampling the white dots represent the initial seed to train a first model.

While static sampling leads to a uniform placement of the samples, adaptive sampling quickly identifies the areas where the test function may be optimal (here minimal). This is done by picking locations where the expected improvement criterion is the highest [54].

This small experiment showcases how sampling can follow a specific objective and possibly, increase sampling efficiency to achieve a certain accuracy in the area of interest.

3.4. Surrogate model fitting

Model construction happens in three steps.

1. Data preprocessing and model type selection
2. Model training and hyper-parameter optimisation
3. Model validation

For brevity and because of an abundance of existing literature, we only provide a small introduction to the field and the existing types of surrogate models. The interested reader is referred to [55] for an introduction on machine learning, to [14] for a book on surrogate modelling, and to [30] where different surrogate modelling techniques for building design are compared.

3.4.1. Data preprocessing and model type selection

The input and output data format must be suitable for the surrogate modelling approach of choice. For example, most approaches require the inputs to be numerical instead of categorical. In that case, categorical variables can be transformed to dummy variables [55]. Once formatted correctly, the data is split into training and test samples. A random separation of 20% of the data for testing is suitable. Finally, some model types require the inputs to

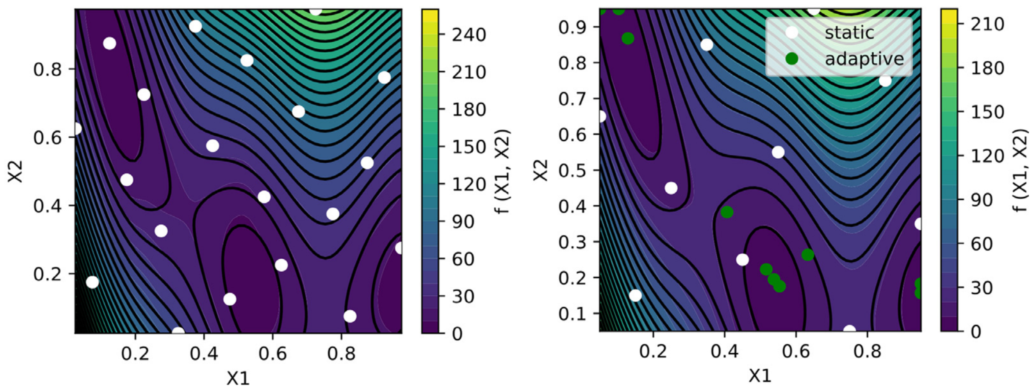

Fig. 4. Showcasing the difference between static (left) and adaptive (right) sampling. On the left 20 samples are chosen based on LHS. On the right, first an initial set of 10 samples was picked using static sampling (LHS) followed by 10 adaptively selected samples using the expected improvement criterion [54].

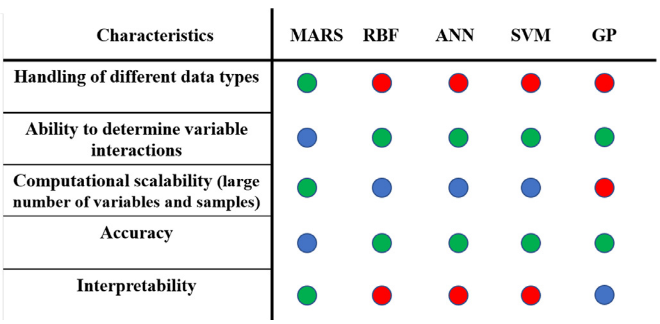

Fig. 5. Comparison of different non-parametric surrogate models based on [55, p. 351]. Green, blue and red dots indicate good, medium and poor performance with regard to the characteristics listed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

be normalized to the same range which ensures equal weighting of variables during model training.

The selection of the surrogate model type is primarily driven by reaching the highest surrogate accuracy possible. Sometimes a trade-off between optimum accuracy and an interpretable model structure is favoured [48,56]. Although each model type has advantages and disadvantages with regard to certain modelling requirements as shown in Fig. 5, many authors suggest the initial use of multiple models to find the most suitable one [13,15].

Model types may be grouped into parametric models and non-parametric models [56,57]. The former uses assumptions on the functional relationship of inputs and outputs. Based on that assumption, a data model is derived whose parameters are calibrated using the collected data. In non-parametric modelling the goal is not to find the correct parameter values of a predefined data model but to find the underlying functional relationship between inputs X and outputs y [57]. In building design, performance metrics like energy consumption may behave non-linearly, featuring discontinuities and multiple modes [30,43,44]. Understanding that behaviour and manually encoding it in a parametric model may be difficult and time consuming. Non-parametric, algorithmic modelling automates this process and thus, may be more suitable for to quickly modelling the relationship of design parameters and performance metrics. In the following, examples for the two model types are given.

3.4.1.1. Parametric models. Multiple linear regression is the most popular parametric model. Its structure and variables are specified manually preliminary to model training. The structure can include variable interaction terms or variables transformed by taking its nth order as done in polynomial regression. Even if variables are combined or transformed, linear regression remains *linear in parameter* meaning no model parameter appears as an exponent or is multiplied or divided by another parameter.

Other parametric models can be developed but they all share a common disadvantage. Unless knowledge allows to derive a valid assumption for the structure of the data model, they are prone to provide questionable analytical findings and lower prediction performance in comparison to algorithmic models [57].

3.4.1.2. Non-parametric models. Different types of non-parametric methods exist. They include artificial neural networks (ANN), radial basis functions networks (RBF), support vector machines (SVM), multivariate adaptive regression splines (MARS), Gaussian Process models (GP) and others. The model types differ in their generic structure.

MARS models may be considered as an extension to linear regression models which automatically identify variable interactions and suitable variable transformations. This is done by a linear combination of multiple basis functions applied to the input vector. Here, the basis function is commonly a hinge function or a multiplication of multiple hinge functions [58]. The hinge function enables piecewise behaviour of the resulting model which is characteristic for MARS models. The multiplication of multiple hinge functions enables to model arbitrary high order relationships and variable interactions.

RBF networks also use linear combinations of basis functions [59]. They use Gaussians as basis functions and apply them to the distance of the input vector to a center vector associated to each Gaussian. Functions that only depend on the distance to a center vector are radially symmetric which explains the name of this model.

Another model type pivoting non-linear basis functions to model versatile mathematical relationships is the ANN. An ANNs consists of multiple cells, called neurons, which receive inputs from and send their outputs to other neurons. Inside a cell the inputs are weighted, summed up and used in a basis function. Typically, sigmoid basis functions are used which imitate the spiking of a neuron in a human brain. Chaining up multiple layers consisting of multiple neurons gives the ANN a high degree of flexibility and in theory, it is capable to model any mathematical function [55].

In GP, observations are considered as realisations of a multivariate Gaussian distribution. The multivariate Gaussian is used as a prior distribution and this distribution is conditioned by existing data. This leads to a posterior distribution of possible functions which generated the data [60].

Support vector machines were originally designed for classification problems. In support vector classification a hyperplane is determined with maximal margin towards the closest observation of each class. The same method is used in support vector regression to find the hyperplane which centres all observations optimally [61].

3.4.2. Model training and hyper-parameter optimisation

After the data is prepared and the model is selected, its parameters and weights are found using a specific training algorithm. For example ANNs are trained via the well-known *backpropagation* algorithm. Apart from training the model weights, non-parametric models require *hyper-parameters* to be specified. Hyper-parameters allow to tune the variance of the predictions of the surrogate. They should be optimised to balance variance with bias to avoid *overfitting* the model on the training data. An overfitted model does not generalize well on unseen data. To select the hyper-parameters usually multiple different settings are compared in a grid-search or Bayesian optimisation is used [62].

3.4.3. Model validation

Model validation is done using dedicated test data. The accuracy of the model is quantified using different performance metrics. Typical choices are mean absolute error (MAE) or the coefficient of determination (R^2) which quantifies how much of the variance in the data is explained by the model.

3.5. Tools

Existing tools may be sorted into two groups: dedicated surrogate modelling toolboxes and those covering portions of the surrogate derivation process.

The first group of tools covers all steps from 3 to 5 from above. They offer different DoE strategies and surrogate types. Due to the excellent performance of surrogate models on optimisation problems, toolboxes are often designed specifically for optimisation purposes. Matlab users are referred to Matsumoto [63] or to SUMO [64]. A Python option is the SMT toolbox, which focusses on gradient-based optimisation,[2] although the choice of methods is rather limited.

Other tools only provide software for specific steps of surrogate model derivation. The Python toolbox PyDOE offers a set of different static sampling methods.[3] The EPPY toolbox allows to access EnergyPlus input files in Python, which enables to quickly transfer generate simulation models given a set of samples.[4] The well known machine learning toolboxes ScikitLearn [65], Tensorflow [66] and PyTorch [67] all feature different surrogate model types and model validation schemes.

Opossum, a plug-in to Grasshopper, is the only surrogate toolbox dedicated to building design [68]. It can only be used for BDO problems and not for deriving global surrogate models. Opossum is based on the Python toolbox RBFOpt [69].

4. Review of surrogate modelling for building design

A significant amount of literature exists to explore and realise the potential of surrogate models, also termed meta-model or response surface model, for building design. The literature review was started with a search through publications listed in Google Scholar and Web of Science using the terms “surrogate model”, “building design” and “building performance design”.[5] This provided a list of 30 publications. The list was extended by analysing their bibliography. Finally, 57 sources were found, shown in Table 2.

Apart from that, previous reviews in the wider context of surrogate modelling applications were collected. They are introduced in the following Section (4.1).

Although a lot of effort was invested to compile a representative set of ongoing research, the intention of this review is not to be exhaustive. In particular, applications of statistical models trained on non-simulation data (e.g. [70]) or simplified physical models are disregarded (e.g. [71]). This also involves applications of surrogate models for model calibration as in [38].

4.1. Previous reviews

This is the first review on the use of surrogate models for Building Performance Simulation (BPS). However, multiple papers include review sections addressing applications of surrogates for building design.

A review and comparison of model types are found in [16] and [15]. Prada et al. [72] looked at the suitability of different types of surrogate models for evolutionary building design optimisation. A comprehensive review of the use of data for building design may be found in [5]. In [73] and [36] surrogates are mentioned in an overview of literature in the field of uncertainty quantification and in [27] they are part of a review on sensitivity analysis. In [42] and [41] sections cover surrogate models applied to building design optimisation.

Other fields in computational building science use similar techniques as in surrogate modelling, for example [74] and [6] reviewed data driven energy demand forecasting.

In other engineering domains where computational experiments are costly surrogate modelling has been applied extensively. An overview of the application of surrogate modelling in aerospace engineering is given by Wang and Shan [13], Forrester et al. [14], Simpson et al. [45] and Queipo et al. [75].

4.2. Overview of publications

Table 2 and Fig. 6 give a summary of each publication including subject, surrogate model type and sampling strategy. Most of the papers address building design optimisation (22 publications), and leverage surrogate models at the early design stage. Also a wide distribution in the fields of sensitivity (16) and uncertainty analysis (9) was found.

Aside from the applications of surrogates for building design problems, 16 papers compare the suitability of different types of surrogate models for building design.

4.2.1. Surrogate model types

Fig. 6(b) shows that in half of the studies, parametric models are used (compare Section 3). Apart from multiple linear regression, this group encompasses polynomial, stepwise, and LASSO regression. The second most models found in literature are Gaussian Process models (GP) and third most common are artificial neural networks (ANN). Other model types include multivariate regression splines (MARS), support vector machine (SVM), random forest (RF), radial basis function (RBF) and model ensembles. An introduction to the models is found in [76].

4.2.2. Sampling strategies

Eleven studies used adaptive sampling (Fig. 6(d)). All but five use them in combination with a GP model. The majority of papers used static sampling strategies with a strong preference towards latin hypercube sampling (LHS)(15). Other sampling strategies include random, orthogonal array, full-factorial, Box-Behnken design

[2] https://github.com/SMTorg/smt.

[3] https://pythonhosted.org/pyDOE/.

[4] https://pythonhosted.org/eppy/.

[5] As Google Scholar does not support the use of parentheses multiple searches equivalent to [(“surrogate model” or “meta-model” or “metamodel”) AND (“building design” or “building performance design”)] were conducted.

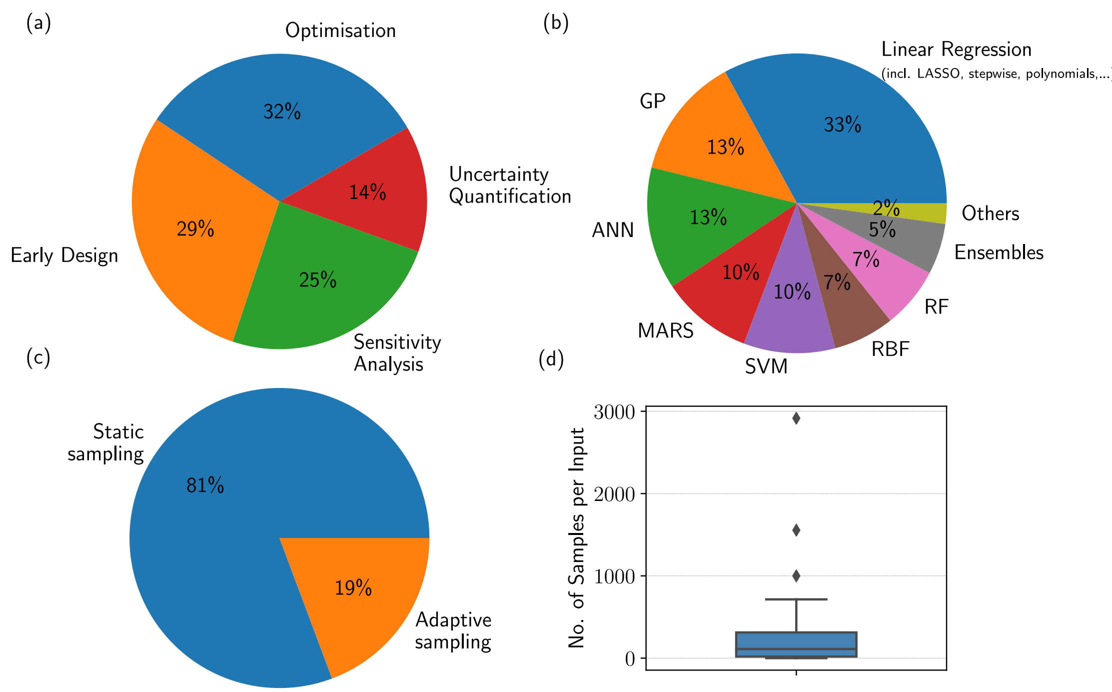

Fig. 6. Overview of publications. Figure (a) highlights the applications of surrogate modelling found in the building design research domain. Figure (b) shows the different surrogate model types used. Figures (c) and (d) focus on the sampling methods to derive an accurate surrogate. Figure (c) shows the share of papers that used static sampling instead of adaptive sampling. Figure (c) indicates how many samples per input were collected in each paper.

and L12-Taguchi tables based sampling. A limited number of studies used manual sampling or evaluated all possible combinations of design parameters (full-factorial).

The number of simulation samples per input is shown in Fig. 6(c) and Table 2. The range is large spanning from single digits to thousands of samples per input. Quantifying the sampling efficiency by number of samples per input is questionable as the design space does not increase linearly but exponentially with each input parameter added. Another option would be to quantify the share of the design space covered by samples [25]. However, as input variables may be continuous, discrete or categorical, and their ranges change drastically among the different studies, the design space size of each study would have to be calculated individually. This is beyond the scope of this review.

4.2.3. Model objectives (outputs) and parameters (inputs)

Fig. 7 shows which inputs were used for different model objectives (outputs). Annual energy demand (and energy use intensity) is the most common output. Another big fraction of papers approximated heating and cooling demand.

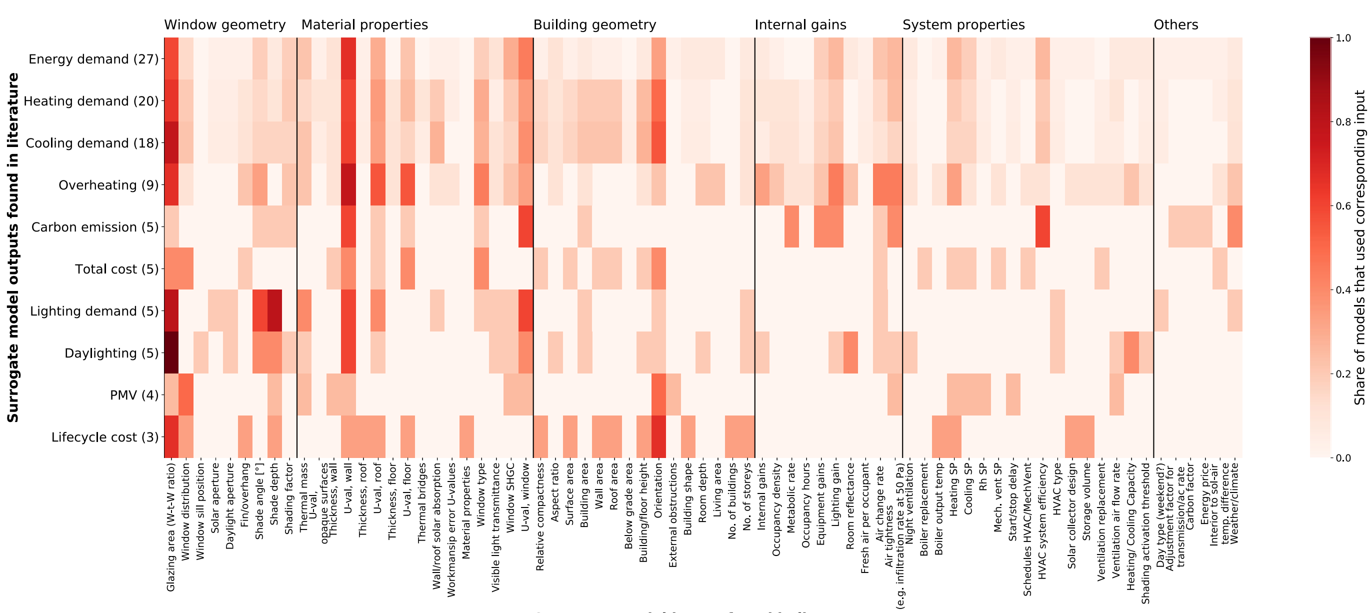

Fig. 7. Usage of in- and output variables in the literature. The figure shows the share of models which used a specific input for a specific output. Next to the outputs (y-axis) the number of associated studies is shown in brackets.

To model the energy demand mainly inputs on building geometry, window geometry and material properties are used. Many authors rely on window-to-wall ratio and wall thermal transmittance as model inputs.

4.3. Discussion of papers grouped by purpose

4.3.1. Early design

As presented in Section 2.1, surrogate models may play a key role in overcoming current limitations of building simulation tools for early design applications. The following literature used surrogates

- To provide rapid feedback during early design,
- and to run large amounts of simulations quickly to provide estimates of *design variability*.

As the amount of existing literature is large, we subdivide the literature corresponding to the geographical scope on which the surrogate model was validated. The scope ranges from one specific building location to multiple climate zones.

4.3.1.1. Rapid feedback. The first group of authors used surrogates for the design of **one building**:

In an early work from 2002, [77, #1] embedded a polynomial regression model into an early design tool which computes energy demand based on 10 parameter inputs. The paper shows that the concept of surrogates has existed for a long time. However, the provided method is not as robust in comparison to more recent publications. For example, surrogate modelling errors are not computed on separate test data.

A similar tool, the design space exploration assistance method (DSEAM), embeds a surrogate model into a CAD tool to provide instant feedback during early design [26, #2]. Performance predictions are visualised as a three-dimensional surface and in a parallel coordinate plot. The method is showcased for the design process of an urban office building. The visualisation techniques enabled to intuitively identify the most influential design parameters for reducing energy consumption.

Geyer and Schlueter [25, #3] considers the use of surrogate models covering a large geographical scope and multiple building types rather suitable for educational purposes but foresees the need for highly individualised surrogates to ensure high accuracy for a specific building. Therefore, they developed an automated surrogate derivation method customized for building design practitioners. They fitted a polynomial model with an algorithm which automatically selects exponents and interaction terms. Consequently, their modelling scheme is non-parametric, which is unusual for polynomials, and allows to combine high surrogate interpretability and accuracy.

Instead of developing holistic tools the following papers looked at more specific elements of surrogate modelling. The authors of [78, #4] tried to overcome the curse of dimensionality, i.e. the computational cost incurred by high numbers of surrogate inputs and outputs. They used ANNs and LASSO regression, which scale more efficiently than MARS, GP and RBF, to build models with 156 inputs and 80 to 90 outputs that are time resolved at 15 min. Their model was trained on an extensive EnergyPlus dataset (267 TB) generated on large cluster computer. This clearly bound their method to users with access to large hardware, unless their surrogate would generalize well, i.e. it could be reused for a variety of building design problems. Unfortunately, this was not addressed by the authors. An advantage of the 15 min-resolution is that it enables users to not only speed up building design, but also automate tasks like system sizing and demand profiling.

Yi et al. [79, #5] computed emergy[6] using EnergyPlus output data and trained a surrogate model on that postprocessed data. They achieved a prediction performance of $R^2 \approx 0.62$[7] which is low in comparison to studies where surrogates estimated energy demand.

Maltais and Gosselin [81, #6] studied daylighting with regard to comfort (glare index) and lighting demand. They fitted a polynomial model for both outputs and found a large accuracy difference although the same simulation data was used for model fitting ($R^2_{GI} \approx 0.95$, $R^2_{LD} \approx 0.78$). The finding that accuracy varies for different model outputs was observed in other publications (e.g. [30,82,83]).

Lastly, Korolija et al. [84, #7] provides an interesting extension to surrogates of whole building simulation programs. The authors state that including detailed HVAC systems into simulations increase simulation run time by 1.3 to 3.7 and that it requires advanced know-how which architects may not have. Therefore, they derived a polynomial model to map building energy demand to energy requirements of secondary HVAC systems which distribute thermal energy inside the building (e.g. ducts). They found that, although a simplistic polynomial was used, accuracy is satisfactory with a relative error of less than 10% in more than 80% of the cases for cooling and heating requirement. In future work one could study if accuracy can be improved with a non-parametric instead of a polynomial model. Nonetheless, the given performance seems good enough to save time by replacing detailed HVAC modelling with a surrogate during early design.

In the second group of studies, surrogates were derived considering **multiple climate regions**:

Catalina et al. conducted two studies to apply linear regression to receive rapid information on heating demand during early design [85][86, #8]. In their latter study the regression model included quadratic terms and was fitted using iteratively re-weighted terms. The model uses only three inputs (heat loss coefficient, south equivalent surface and difference of indoor to outdoor temperature). Although the model was trained and validated on simulation data, the model predictions were also compared to actual measurements of buildings. A significant difference of predicted to measured annual heating demand is found and could only be taken into account by introducing a building-specific correction term to their polynomial.

Hygh et al. [87, #9] applied multivariate linear regression to approximate energy use in four different climate zones. In comparison to the previous study, 27 input parameters were used leading to only slightly higher accuracy ($R^2 > 0.98$) than the previous paper ($R^2 \approx 0.97$) but allowing a wider variety of designs to be analysed using one surrogate. Similar to [87], [88, #10] and [89, #11] used one multivariate linear regression model per climate zone. Both authors point out that the accuracy varied for different climate zones. All three studies used multiple models for each climate zone.

The question arises if variables capturing the characteristics of different climate zones can be found. When used as inputs, they could enable the use of one surrogate covering all climate zones. An approach is given in [90, #12], where a SVM model is derived that estimates energy use of naturally ventilated commercial buildings in Brazil. They generated a dataset with 418 different weather files. The weather data was then reduced to a few statistical variables and used as surrogate model inputs. Romani et al. [91, #13] also used only one model (full quadratic polynomial model) to predict the heating and cooling demand in Morocco (four cli-

[6] Energy refers to the amount of solar energy embodied in the energy used up during a service or production [80].

[7] Calculated from the provided F-score.

mate zones). They found variable interaction helped to compensate for the use of one model only.

Thirdly, one study was found covering **multiple building types** within one climate region [92, #14]. They derived multiple single-output ANNs to find optimal retrofit strategies for Southern Italy. Their six separate networks estimate heating demand, cooling demand and occupant comfort of the existing non-retrofitted and retrofitted building stock.

Recently, some researchers focussed on finding **general** surrogate models without a climatic or geographical scope.

In [93, #15] the same authors as in [89] tried to find more generalizable surrogates by integrating know-how on building physics to derive more meaningful features (inputs) from common design parameters. Their features include energy gains due to transmission, air change rate and solar heat gain. Thy were calculated assuming steady-state behaviour but fed to a surrogate which approximates dynamic, i.e. non-steady-state, simulation. They achieved high accuracy scores for a single-building case study ($R^2 \approx 0.99$) but further details and benchmarking are required.

Singaravel and Geyer aimed to decompose a "monolithic surrogate model" into multiple components [82,94, #16]. In the first approach they suggest fitting multiple ANNs, each approximating heat gain through an individual building element like a wall or a window. Adding up the outputs of the individual models they computed the whole building performance. In their second approach they compartmentalised one surrogate into multiple approximating heating and cooling demand. The use of the recurrent long-short term memory network (LSTM) allowed to model dynamic effects. They studied the generalizability of both approaches on three test cases. Based on the results of the most complex building design case, the first approach ($R^2_{cooling} \approx 0.98$, $R^2_{heating} \approx 0.85$) seems to outperform the zonal LSTM model, but no final conclusion is drawn by the authors.

4.3.1.2. Design Variability. After studying methods for early building design [11, #17], Ostergard et al. developed a new design methodology to guide sustainable building design with multiple stakeholders involved [23]. They propose to first evaluate the performance of large number of designs, using a surrogate model, and sequentially filter them using specifications on the final building performance (outputs). This provides distributions of possible design choices (inputs). To visualise the impact of performance specifications, the parallel coordinates plot was favoured.

Instead of specifying outputs, Hester et al. [21] determine the change of the output distribution if one of the design parameters is decided. They used a linear regression model to run Monte Carlo simulations after each design choice. Part of the authors conclusion is that not only the speed of the surrogate is helpful, but also the reduced number of parameters required to provide a performance estimate.

A similar study was done by Basbagill et al. [24, #18]. The authors constructed probability distributions for life cycle cost and performance treating decision parameters as random variables. The distributions are derived from a database generated with eQuest [95] using orthogonal array sampling. Although no surrogate is used (but a fast physics based simulation model instead), the methodological steps are similar to those of surrogate modelling, and show that fast simulation software is an alternative to surrogate modelling.

4.3.2. Sensitivity analysis

There are three cases found in literature where surrogate modelling and SA are combined (compare Section 2.2). *(i)* use SA prior to surrogate modelling for variable selection, *(ii)* use surrogates to accelerate variance-based SA, and *(iii)* use SA complementary to surrogate modelling to increase analytical insight into the data. While in general non-parametric methods are preferred for building surrogate models, parametric methods are a regular choice for SA as one can easily access variable importance estimates by looking at the regression coefficients (e.g. linear regression).

For brevity, this section does not cover the full literature on SA in which linear regression was applied. Further literature can be found in [27].

4.3.2.1. Variable selection for surrogate models. Many studies use SA for variable selection. Here we summarize eight contributions which share a similar approach.

Dhariwal and Banerjee [96, #41] conducted fractional factorial design-based sampling and determined the most impactful parameters using Morris' method. The parameters found are used as model inputs to a second order polynomial surrogate (response surface model).

Hopfe et al. [97] computed standardized rank regression coefficients (SRRC) to quantify the impact of uncertain parameters. They chose the five most influential to complement the design parameters as inputs of the GP. Similarly, SRRCs were used in [98, #20].

Multiple SA methods (Pearsson, Spearman, Kolm and Krusk coefficient) were computed in [99, #22] to take linear, monotonic and non-monotonic, and asymmetric variable correlations into account. The most important variables were used as inputs of an ANN to emulate internal air quality simulations of a building stock. They found differing results, which may be caused by the way the four SA methods handle non-linearities. They suggest the use of simple scatter plots to discover non-linearities and to pick the SA accordingly.

Maltais and Gosselin [81, #6] used linear regression and variance-based SA prior to fitting a polynomial to estimate natural daylighting performance. As part of their study, they looked at the numbers of samples required to achieve stable sensitivity coefficients. While standard regression coefficients stabilized after 600 runs, the Sobol indices (variance-based SA) converged after 1900 runs. The data generated from those 1900 runs was subsequently used to train a surrogate model.

Ostergard et al. [30, #56] compared different surrogate modelling techniques. To facilitate model fitting of many different model types, they applied a global SA on the *hyper-parameters* using Smirnov two-sample statistics.

The same authors in [23] and [32] chose surrogate inputs by ranking parameters with the Morris screening method. The method was favoured as it is fast, requires fewer simulation samples and its qualitative ranking of variable sensitivities is close to more complex SA methods.

4.3.2.2. Surrogate model based sensitivity analysis. In [32, #23] two different kinds of surrogate models and two types of sensitivity analysis are applied. The authors fitted a surrogate (polynomial chaos expansion model) to conduct a variance-based SA with 24,000 samples. For the derivation of the surrogate, Morris screening was conducted beforehand to find input parameters as introduced in the section above.

Eisenhower et al. [29, #24] emulates the design space with high dimensional model representations [100] to compute variance-based global sensitivities for thousands of parameters. This would take multiple days without a surrogate.

Tsanas and Xifara [101, #25] used a random forest model (RF) to estimate the energy performance of a building. RFs provide parameter importance ranking through the impurity metric which the authors compared to SRRC-based ranking. They found slight differences and warned of the limitations of linear regression in dealing with collinearity.

In [102, #26] both variance-based SA metrics using a MARS model and SRRCs were computed to find the most important parameters for building-related carbon emissions taking climate change into account. Although the latter ignores variable interactions, both approaches provided similar results. 6000 simulations were required to conduct the variance-based SA in their case study on a UK office building. Without the use of a MARS model this analysis would take multiple days.

4.3.2.3. Interpretation of surrogate models. The last application of SA is to provide insight into the functional behaviour of a surrogate model, if its mathematical structure is too complex to be comprehensible intuitively.

[103, #27] use the Correlation-Adjusted corRelation (CAR) score [104] to understand variable importance for a set of campus buildings. The same data is used to derive multiple surrogates. The combination of both statistical approaches provides interpretive and predictive tools to the building designer.

Similarly, Hygh et al. [87, #9] and Chen et al. [105, #28] computed standardized regression coefficients alongside training a stepwise linear regression and a MARS surrogate model. The latter also used bootstrapping methods to validate the robustness of the sensitivity coefficients. The same authors conducted a SA and a heuristic optimisation in [106].

4.3.3. Uncertainty analysis

Output uncertainty quantification is similar to the assessment of design variability during the conceptual design stage (see Section 4.3.1), and often conducted alongside a sensitivity analysis (see Section 4.3.2). Like design variability assessment and sensitivity analysis, current uncertainty quantification methods mostly rely on sampling based methods, i.e. input parameter distributions are converted to output distributions using Monte Carlo simulations [73]. The idea is to use surrogates to accelerate Monte Carlo simulations [36], however the existing literature is rather limited.

Hester et al. [21, #29] sequentially generate probability distributions of the output after each design parameter is specified. This visualises the converging distribution of the output with each design decision taken. Here, they used a linear regression surrogate model to avoid long computation times.

Eisenhower et al. also derived probability distribution on comfort and annual energy demand based on 1009 input parameters uniformly distributed within 20% of their baseline [107]. They compared the distributions of both 5000 simulation runs as well as SVM evaluations and found high agreement in the mean and variance.

Rivalin et al. [32, #23] and Kim [98, #20] studied the use of Gaussian process emulators and polynomial chaos expansion (PCE). The former paper first applies LHS to derive the PCE model. Once they have an accurate model they re-apply LHS to derive the model output dispersion and distribution faster than with random Monte Carlo simulation.

Papadopoulos and Azar [108, #30] use a surrogate model to study the influence of varying levels of control of occupants and facility management under uncertain occupant behaviour. After training the surrogate, a linear regression model, they generate 11^3 cases each with a different level of control of occupants on lighting, equipment and thermostat setpoints. They underpin the cases with uncertainty of human behaviour and generate 25 samples for each case, such that the surrogate model is evaluated for 33275 samples. They visualise the results in an appealing three-dimensional map.

One of the reasons for the scarcity of existing literature may be the findings of Macdonald [109] and Lomas and Eppel [40], who stated that, disregarding the number of uncertain parameters, between 60 and 100 samples are required to receive an accurate probability distribution of the outputs. Based on existing literature this number of samples is probably not sufficient to derive an accurate surrogate model (see column "number of samples" in Table 2). Thus the majority of papers used standard building performance simulation for sampling, sometimes running them on high performance computing facilities [29,35,110,111]. A workbench for propagating input uncertainties to performance uncertainties using EnergyPlus may be found in [112].

4.3.4. Design optimisation

In this section we review the papers which replace simulation models by surrogate models to accelerate the search for optimal building design parameters. The literature can be sorted into two different groups: *(i)* Direct surrogate-based optimisation (*(a)* in Figs. 2 and 3) and *(ii)* iterative surrogate-based optimisation (*(b)* in Figs. 2 and 3).

4.3.4.1. Direct surrogate-based optimisation. An early application of surrogate models for BDO is found in [113, #31]. Wong et al. used an ANN based grid search to determine optimal selections of solar aperture, daylight aperture, overhangs and side fins to minimize annual energy consumption. The authors limited the grid search to only 41 surrogate model runs although their surrogate model should be cheap to evaluate much more samples.

Magnier and Haghigat [114, #32] used an ANN and the NSGA-II optimizer to minimize energy consumption and comfort. They reported that the surrogate-based optimisation achieved an accuracy of within 1% of simulation-based optimisation and only required seven minutes, but stress that generating the database underlying the surrogate took three weeks. Nonetheless, if the same number of model evaluations during optimisation would have been conducted with a simulation, the process would have taken 10 years. The relatively long simulation time might be caused by choosing 2 min time-steps for their simulation, while having a workstation with a 1.66 GHz processor.

Shortly after that Asadi et al. [115, #33] published a similar study focussing on retrofit optimisation (between one and three objectives) using a validated EnergyPlus base model for the surrogate derivation. Sample simulation took three days and their model achieved a decent accuracy (MRE < 2.5%) on a validation set. Like other authors, they did not report the accuracy of the optimality candidates. They point out that the speed of surrogate-based optimisation (< 9 min) enables designers to explore different design strategies at early stage. In comparison, simulation-based exhaustive search would have taken 75 days. They suggest increasing the number of design parameters and incorporating surrogate uncertainty prediction to further expand the insight for architects. A study similar to [114] and [115] can be found in [116, #34], which focussed on L-shaped multi-story office buildings.

While the previous authors used an ANN model in combination with a genetic algorithm, Eisenhower et al. [107, #35] and Chen et al. [106, #36][117, #37] used SVM models. While the latter used NSGA-II like previous papers, Eisenhower et al. leveraged that a surrogate model enables the use of gradient-based instead of derivative-free optimizers. Comparing both on a multi-object optimization of comfort and energy demand, the results were equally stable but gradient-based optimizers converged significantly faster (a few seconds instead of some minutes). In all three studies sensitivity analysis was conducted to reduce the number of design parameters prior to optimization. Eisenhower et al. showed that similar optima were found with seven input parameters as with 1009 parameters. Hence, increasing the number of inputs barely increased the optimality score.

Constrained gradient-based optimisation (sequential-quadratic programming [118]) was used in [119, #38] together with an RBF model to maximise comfort of naturally ventilated buildings using window geometry. Their surrogate was trained on simulation data generated by a sequence of computational fluid dynamics (CFD) and building energy simulation. To validate the optima, they also ran simulation at the optima and found a MAE of $< 10\%$ in the RBF-optimization outcomes. The validation did not include a comparison of surrogate-based and simulation-based optimization. As the building considered was simplistic, further research is required for a general conclusion on their method.

A comparison of different surrogate models and different sampling approaches for evolutionary design optimisation is given in [72, #39]. The models were compared with regard to their efficiency, efficacy and solution quality. The authors recommend MARS models over GP, RBF and SVM models due to higher accuracy. Part of their study was an analysis of potential time savings using surrogate models (including sampling). They found savings of more than 80% to be feasible, particularly for complex design spaces. In comparison to the other papers they measured optimum accuracy not only at specific points, but computed the generational distance between the surrogate-based and the simulation-based Pareto fronts. One of many findings is that increasing the number of training samples leads to a lower generational distance.

Above, only non-parametric models were presented which are complex and difficult to interpret. Some authors prefer simpler approaches like polynomials. [96, #40] used a second-order approximation and benchmarked it against simulation-based optimization with and without parameter importance analysis. Carreras et al. [120, #40] optimised a cubic house to minimize cost and environmental life-cycle performance. They used cubic spline interpolation as a surrogate and reduced gradient optimization to determine the best insulation thickness. Including the time for database generation they found a time reduction of 8 times, down to 21.3 hours in comparison to simulation-based optimisation.

4.3.4.2. Iterative surrogate-based optimisation. In comparison to direct optimisation, iterative surrogate-based optimisation relies on a space infill criterion which balances exploration and exploitation. This difference leads to a changing preference on surrogate model type. While in the previous section a lot of studies used ANN or SVM models, here often GP models are applied which can quantify model uncertainty. This can serve as an *exploration* criterion for adaptive sampling (see Section 2).

An early work on the use of GP in the BDO domain by Gengembre et al. [121, #42] minimized life-cycle cost and energy consumption using the constrained efficient global optimizer [122]. The GP model was updated with samples chosen to maximise the expected improvement criterion.

Similarly, in [123, #43] and [124, #44] the expected improvement criterion is used. In the former study, the optimisation process is benchmarked against simulation-based optimisation using NSGA-II. It was found to have a steeper convergence curve and to require fewer high-fidelity model simulations. However, in the case of multi-objective optimisation this could not be confirmed.

Gilan et al. [125, #45] used a combination of GP and NSGA-II. In comparison to other studies, they computed the space infill criterion based on a whole area of the design space instead of an individual point. They calculated the mean posterior variance of the offspring from each iteration of the optimizer (50 samples). Comparing their method to direct surrogate-based optimisation, they found good agreement of the Pareto Front while cutting runtime by two thirds.

Hopfe et al. [97, #19] performed optimisation using the SMS-EMOA algorithm [126]. As the objective function they used the mean value of 201 perturbations around a point proposed by the optimizer. The goal is to find more robust solutions. To the best of our knowledge, they could replace their GP model with any other surrogate model type. They claimed that their approach helps to reduce the number of samples needed to find an optimal solution by 5–20% compared to simulation-based optimization.

Besides the literature on GP, the following publications used methods which are independent of the surrogate type. They rely only on model predictions instead of posterior variance estimates provided by GP models. One early study [127, #46] updated ANNs at each iteration with samples selected by NSGA-II leading to a locally accurate surrogate. The model was initialised on a set of 50 samples and the optimizer provided 50 samples at each iteration. In a similar fashion, [128, #47] used an SVM and [129, #48] a RBF model to minimize building cost. The former reported their method reduced optimization time by up to 60% on a case study.

Lastly, Wortmann developed Opossum as plug-in for Grasshopper. His tool, which is based on RBFOpt [69], adaptively trains and optimizes an RBF model [130, #49]. In a comparison with eight other optimization schemes, RBFopt was found to be the fastest converging and second most stable (after direct search) to optimize the energy demand of a building with 13 design parameters. Furthermore, it was the best performing algorithm in maximizing useful daylight illuminance (UDI) while minimizing glaring effects. This paper clearly shows the great potential of iterative surrogate-based optimization for building design problems.

5. Trends and practical aspects

This review confirms that surrogate models are a strong element in current building performance simulation and optimisation research, and results have shown that they are a suitable alternative to common building simulation models in certain cases. Performance analysis during the conceptual building design stage, sensitivity and uncertainty analysis, as well as building design optimisation are more accessible, primarily due to the large reduction of computational cost.

In the following section, we list application trends and practical aspects extracted from the reviewed literature.

5.1. Trends in the application of surrogate models

- As surrogate models lower the computational burden of early design, sensitivity and uncertainty analysis, it becomes possible to get insight into building performance over the whole space of potential design options.
  A good way to visualise this is the parallel coordinates plot [23,26]. In comparison to simulation-based design exploration, this allows users to intuitively explore multi-dimensional spaces and find promising designs in a limited time.
- The value of surrogate models hinges on the decrease in time to conduct a certain analysis while maintaining high accuracy. Although many examples on the use of surrogates for early design, SA and UA exist, there is a lack of understanding how large the time savings can be. Only in papers on optimisation analysis, thorough analyses of time savings were found.
- Surrogate-based optimisation showed promising first results to speed-up building design optimisation.
  The listed publications achieved a time reduction of up to 80% [72] and the identified optima have proven to have high quality in comparison to full simulation-based black-box optimisation [72,107,130]. Examples for time savings are given in Table 3. An open question is whether direct or iterative surrogate-based optimisation better fits the requirements of building designers. Iterative surrogate-based optimisation may be fast and efficient [130], but as stated in [107], direct optimisation using a global surrogate allows to

Table 3
Examples from the literature for potential of optimization time reduction while maintaining accuracy of the optimum.

Optimization strategy		Multi objective	Optimizer	Time reduction	Comparison to simulation-based optimum
Direct surrogate-based					
	[114]	x	NSGA-II	-97% (10y to 3w)	Energy demand: $<2.5\%$, Overheating: $<25\%$, max. error on global validation set
	[96]	x	NSGA-II	-55% (27h to 12h)	$<10\%$ max. error on global validation set
	[107]		IPOPT/NOMAD	+4% (49h to 52h)	0.7% deviation from true optimum
	[117]	x	NSGA-II	n/a (7.4h to n/a)	1.59% deviation from true optimum
Iterative surrogate-based					
	[125]	x	NSGA-II	-52.2% (14h to 7h)	3.17% difference in hypervolume of Pareto Front
	[72]	x	NSGA-II	-82% (71h to 13h)	75%-80% samples of original Pareto Front found
	[128]	x	NSGA-II	-60% (23h to 9h)	optima are close to true Pareto Front but have low diversity (spread metric Δ increases from 0.41 to 1.01)

easily change the optimisation objective or optimizer settings without rerunning simulations.

- Recently, researchers have been trying to find more general surrogates applicable to many different problems. One may envision that if a surrogate is highly generalizable, it could fully replace building simulation tools for the most common types of building projects.
  The maximum scope of a single surrogate model has been broached by multiple publications. Most authors used surrogate models only for a specific building and therefore, Geyer and Schlueter [25] focussed on automating the surrogate derivation process. Others have fitted a single surrogate to estimate the performance of multiple buildings of a specific type [92] or in one climate region [86,87]. In future, one could capture weather data in a few descriptive variables and use them as inputs to a general surrogate as introduced in [90].
  Another option is given by Singaravel et al. [94]. In their grey-box approach they used domain knowledge to split one surrogate into multiple physical entities representing energy fluxes through walls, floors, etc. Further research is required to support their promising initial findings.
- Lastly, most surrogate model types lack interpretability of their mathematical structure and are not suitable to answer analytical questions. One way around is to increase the number of surrogate model outputs. For example in the aforementioned grey-box approach multiple physical meaningful metrics are estimated. Another example was given in [78] where time resolved energy use instead of annual performance metrics were reported.

5.2. Practical aspects

- In the reviewed publications, it seems feasible to explain more than 95% of the variance in simulation results ($R^2 > 0.95$) for energy, heating and cooling demand estimates with one surrogate model for one or more buildings in one or more climate zones. The accuracy was found to be lower for other kinds of output (e.g. *Max* CO_2 [30], or *Overheating* [131]).
- Both model selection and hyper-parameter optimisation are important to achieve high accuracy as shown in different comparisons of surrogate modelling techniques for building design [15,30,72,132–134]. For example, [30] advocates the use of ANN for extensive analysis, GP for non-experts to get high accuracy, and MLR for quick, automated surrogate modelling.
  In general, accuracy may be improved by using non-linear, parameter-free models instead of parametric ones. However, we observed that even with models that are *linear in parameter*, especially polynomials, an accuracy of $R^2 > 0.95$ is achievable in some cases [25,26,91].
  As important as model selection is hyperparameter optimization as standard model settings usually yield insufficient accuracy [15]. A simple grid search may already yield a large increase in accuracy [106]. It is promising that recent publications relied on validated, sophisticated hyperparameter optimisation methods using state-of-the-art toolboxes (see Section 3).
- A frequently reported problem is the limited number of inputs a surrogate models can handle without exploding computational cost (*curse of dimensionality*).
  It is popular to integrate sensitivity analysis into the surrogate derivation process to determine the most important parameters. A surrogate model using only those inputs has proven to be accurate and to provide sufficiently optimised design options [106,107].
- While good model selection does not necessarily increase accuracy, it may increase sampling efficiency. [15] and [30] both found GP models to be sample efficient while RBF and MARS models require a lot of samples to reach high accuracy.
- The best choice of sampling algorithm is uncertain. Most studies within that review used latin hypercube sampling (see Table 2).
  No study exists which compares all static sampling schemes at once. In [135] Sobol's sampling was used for Monte Carlo simulation and provided more precise and robust output distributions than latin-hypercube and random sampling.
  Furthermore, a comparison of static and adaptive sampling in the field of building surrogate models is yet to be done. In other research domains adaptive sampling strategies have successfully shown to require less simulation runs until the surrogate reaches a certain accuracy [136].

6. Conclusion

This review provides a thorough discussion of publications that use surrogate models for sustainable building design.

The publications are sorted according to application area into conceptual design, sensitivity and uncertainty analysis, and building design optimisation. In particular, the use of surrogate models as a tool to give insight and understanding into high dimensional building design spaces was found to be popular in current research. Furthermore, multiple publications have shown that embedding surrogate models into optimisation procedures accelerates the process significantly.

Apart from the analysis of research trends, this review serves as a practical guide. A detailed introduction to the process of deriving a surrogate model is given. The publications reviewed are categorized in both a large table and multiple figures providing a convenient technical overview of the field. Finally, practical aspects of all publications are summarized in a separate section with regard to model accuracy, model type, input selection and sampling strategy.

We expect future research to focus on lowering the computational cost for deriving a surrogate model and to increase the interpretability of models. The former could be achieved by implementing advanced sampling strategies, or by extending the scope of a single surrogate model from one to multiple buildings such that the derivation process does not have to be repeated for every analysis. Low interpretability can be avoided by compartmentalising surrogate models into multiple physically meaningful submodels.

Surrogate modelling has already been shown to lower the burden for architects and engineers to assess sustainable building designs using advanced performance analysis. We envisage, that it will play a key role in achieving sustainability in the future building stock.

Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgement

This work has been supported by grant funding from CANARIE via the BESOS project (CANARIE RS-327).

References

[1] R.K. Pachauri, M.R. Allen, V.R. Barros, J. Broome, W. Cramer, R. Christ, J.A. Church, L. Clarke, Q. Dahe, P. Dasgupta, et al., Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change, IPCC, 2014.
[2] I.E. Agency, Tracking Clean Energy Progress, 2017.
[3] Navigant Research, Global building stock database, commercial and residential building floor space by country and building type: 2017–2026„ 2017, https://www.navigantresearch.com/reports/global-building-stock-database.
[4] R.S. Michalski, J.G. Carbonell, T.M. Mitchell, Machine Learning: An Artificial Intelligence Approach, Springer Science & Business Media, 2013.
[5] M. Loyola, Big data in building design: a review, J. Inf. Technol. Constr.(ITcon) 23 (13) (2018) 259–284.
[6] Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, X. Zhao, A review of data-driven approaches for prediction and classification of building energy consumption, Renew. Sustain. Energy Rev. 82 (2018) 1027–1047.
[7] A. Ahmad, M. Hassan, M. Abdullah, H. Rahman, F. Hussin, H. Abdullah, R. Saidur, A review on applications of ann and SVM for building electrical energy consumption forecasting, Renew. Sustain. Energy Rev. 33 (2014) 102–109.
[8] H. Lim, Z.J. Zhai, Comprehensive evaluation of the influence of meta-models on Bayesian calibration, Energy Build. 155 (2017) 66–75.
[9] C. Hjortling, F. Björk, M. Berg, T. af Klintberg, Energy mapping of existing building stock in sweden–analysis of data from energy performance certificates, Energy Build. 153 (2017) 341–355.
[10] S. Attia, E. Gratia, A. De Herde, J.L. Hensen, Simulation-based decision support tool for early stages of zero-energy building design, Energy Build. 49 (2012) 2–15.
[11] T. Østergård, R.L. Jensen, S.E. Maagaard, Building simulations supporting decision making in early design–a review, Renew. Sustain. Energy Rev. 61 (2016) 187–201. https://www.sciencedirect.com/science/article/pii/S136403211600280X.
[12] T. Simpson, V. Toropov, V. Balabanov, F. Viana, Design and analysis of computer experiments in multidisciplinary design optimization: a review of how far we have come-or not, in: 12th AIAA/ISSMO multidisciplinary analysis and optimization conference, 2008, p. 5802.
[13] G.G. Wang, S. Shan, Review of metamodeling techniques in support of engineering design optimization, J. Mech. Des. 129 (4) (2007) 370–380.
[14] A. Forrester, A. Keane, et al., Engineering Design via Surrogate Modelling: A Practical Guide, John Wiley & Sons, 2008.
[15] L. Van Gelder, P. Das, H. Janssen, S. Roels, Comparative study of metamodelling techniques in building energy simulation: guidelines for practitioners, Simul. Modell. Pract. Theory 49 (2014) 245–257, doi:10.1016/j.simpat.2014.10.004.
[16] T. Østergård, R.L. Jensen, S.E. Maagaard, A comparison of six metamodeling techniques applied to building performance simulations, Appl. Energy 211 (2018) 89–103.
[17] P. De Wilde, The gap between predicted and measured energy performance of buildings: a framework for investigation, Autom. Constr. 41 (2014) 40–49.
[18] G. Löhnert, A. Dalkowski, W. Sutter, Integrated design process: a guideline for sustainable and solar-optimised building design., International Energy Agency (IEA) Task 23 Optimization of Solar Energy Use in Large Buildings, subtask B. Austria., 2003.
[19] S. Petersen, Simulation-based support for integrated design of new low-energy office buildings, DTU Civil Engineering, Technical University of Denmark, 2011.
[20] S. Attia, J.L. Hensen, L. Beltrán, A. De Herde, Selection criteria for building performance simulation tools: contrasting architects' and engineers' needs, J. Build. Perform. Simul. 5 (3) (2012) 155–169, doi:10.1080/19401493.2010.549573.
[21] J. Hester, J. Gregory, R. Kirchain, Sequential early-design guidance for residential single-family buildings using a probabilistic metamodel of energy consumption, Energy Build. 134 (2017) 202–211, doi:10.1016/j.enbuild.2016.10.047.
[22] R.B. Miller, Response time in man-computer conversational transactions, in: Proceedings of the December 9–11, 1968, fall joint computer conference, part I, ACM, 1968, pp. 267–277.
[23] T. Ostergard, R.L. Jensen, S.E. Maagaard, Early building design: informed decision-making by exploring multidimensional design space using sensitivity analysis, Energy Build. 142 (2017) 8–22, doi:10.1016/j.enbuild.2017.02.059.
[24] J.P. Basbagill, F.L. Flager, M. Lepech, A multi-objective feedback approach for evaluating sequential conceptual building design decisions, Autom. Constr. 45 (2014) 136–150, doi:10.1016/j.autcon.2014.04.015.
[25] P. Geyer, A. Schlueter, Automated metamodel generation for design space exploration and decision-making - a novel method supporting performance-oriented building design and retrofitting, Appl. Energy 119 (2014) 537–556, doi:10.1016/j.apenergy.2013.12.064.
[26] F. Ritter, P. Geyer, A. Borrmann, Simulation-based decision-making in early design stages, in: 32nd CIB W78 conference, Eindhoven, The Netherlands, 2015, pp. 27–29.
[27] W. Tian, A review of sensitivity analysis methods in building energy analysis, Renew. Sustain. Energy Rev. 20 (2013) 411–419.
[28] S. Yang, W. Tian, E. Cubi, Q. Meng, Y. Liu, L. Wei, Comparison of sensitivity analysis methods in building energy assessment, Procedia Eng. 146 (2016) 174–181.
[29] B. Eisenhower, Z. O'Neill, V.A. Fonoberov, I. Mezić, Uncertainty and sensitivity decomposition of building energy models, J. Build. Perform. Simul. 5 (3) (2012) 171–184.
[30] T. Ostergard, R.L. Jensen, S.E. Maagaard, A comparison of six metamodeling techniques applied to building performance simulations, Appl. Energy 211 (2018) 89–103, doi:10.1016/j.apenergy.2017.10.102.
[31] C. Spitz, L. Mora, E. Wurtz, A. Jay, Practical application of uncertainty analysis and sensitivity analysis on an experimental house, Energy Build. 55 (2012) 459–470.
[32] L. Rivalin, P. Stabat, D. Marchio, M. Caciolo, F. Hopquin, A comparison of methods for uncertainty and sensitivity analysis applied to the energy performance of new commercial buildings, Energy Build. 166 (2018) 489–504.
[33] I.A. Macdonald, Quantifying the effects of uncertainty in building simulation, University of Strathclyde Glasgow, 2002 Ph.D. thesis. http://www.esru.strath.ac.uk/Documents/PhD/macdonald_thesis.pdf.
[34] S. de Wit, Uncertainty in building simulation, in: A. Malkawi, G. Augenbroe (Eds.), Advanced building simulation, Routledge, 2004, pp. 39–73.
[35] W. Tian, P. De Wilde, Uncertainty and sensitivity analysis of building performance using probabilistic climate projections: a UK case study, Autom. Constr. 20 (8) (2011) 1096–1109.
[36] W. Tian, Y. Heo, P. De Wilde, Z. Li, D. Yan, C.S. Park, X. Feng, G. Augenbroe, A review of uncertainty analysis in building energy assessment, Renew. Sustain. Energy Rev. 93 (2018) 285–301.
[37] C.J. Hopfe, J.L. Hensen, Uncertainty analysis in building performance simulation for design support, Energy Build. 43 (10) (2011) 2798–2805.
[38] Y. Heo, R. Choudhary, G. Augenbroe, Calibration of building energy models for retrofit analysis under uncertainty, Energy Build. 47 (2012) 550–560.
[39] M. Manfren, N. Aste, R. Moshksar, Calibration and uncertainty analysis for computer models - a meta-model based approach for integrated building energy simulation, Appl. Energy 103 (2013) 627–641, doi:10.1016/j.apenergy.2012.10.031.
[40] K.J. Lomas, H. Eppel, Sensitivity analysis techniques for building thermal simulation programs, Energy Build. 19 (1) (1992) 21–44.
[41] R. Evins, A review of computational optimisation methods applied to sustainable building design, Renew. Sustain. Energy Rev. 22 (2013) 230–245.
[42] A.-T. Nguyen, S. Reiter, P. Rigo, A review on simulation-based optimization methods applied to building performance analysis, Appl. Energy 113 (2014) 1043–1058.

[43] M. Wetter, J. Wright, A comparison of deterministic and probabilistic optimization algorithms for nonsmooth simulation-based optimization, Build. Environ. 39 (8) (2004) 989–999.
[44] M. Wetter, E. Polak, A convergent optimization method using pattern search algorithms with adaptive precision simulation, Build. Serv. Eng. Res.Technol. 25 (4) (2004) 327–338.
[45] T.W. Simpson, J. Poplinski, P.N. Koch, J.K. Allen, Metamodels for computer-based engineering design: survey and recommendations, Eng. Comput. 17 (2) (2001) 129–150.
[46] J. Free, A. Parkinson, G. Bryce, R. Balling, Approximation of computationally expensive and noisy functions for constrained nonlinear optimization, J. Mech. Transm.Autom. Des. 109 (4) (1987) 528–532.
[47] Y.S. Ong, P. Nair, A. Keane, K. Wong, Surrogate-assisted evolutionary optimization frameworks for high-fidelity engineering design problems, in: Knowledge Incorporation in Evolutionary Computation, Springer, 2005, pp. 307–331.
[48] S. Shan, G.G. Wang, Metamodeling for high dimensional simulation-based design problems, J. Mech. Des. 132 (5) (2010) 051009.
[49] F.A. Viana, T.W. Simpson, V. Balabanov, V. Toropov, Special section on multidisciplinary design optimization: metamodeling in multidisciplinary design optimization: how far have we really come? AIAA J. 52 (4) (2014) 670–690.
[50] D.B. Crawley, L.K. Lawrie, F.C. Winkelmann, W.F. Buhl, Y.J. Huang, C.O. Pedersen, R.K. Strand, R.J. Liesen, D.E. Fisher, M.J. Witte, et al., Energyplus: creating a new-generation building energy simulation program, Energy Build. 33 (4) (2001) 319–331.
[51] R. Yondo, E. Andrés, E. Valero, A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses, Prog. Aerosp. Sci. 96 (2018) 23–61, doi:10.1016/j.paerosci.2017.11.003.
[52] S.S. Garud, I.A. Karimi, M. Kraft, Design of computer experiments: a review, Comput. Chem. Eng. 106 (2017) 71–95.
[53] A. Bhosekar, M. Ierapetritou, Advances in surrogate based modeling, feasibility analysis, and optimization: a review, Comput. Chem. Eng. 108 (2018) 250–267.
[54] J. Močkus, On Bayesian methods for seeking the extremum, in: Optimization Techniques IFIP Technical Conference, Springer, 1975, pp. 400–404.
[55] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed., 2013.
[56] G. James, D. Witten, T. Hastie, R. Tibshirani, An Introduction to Statistical Learning, 112, Springer, 2013.
[57] L. Breiman, Statistical modeling: the two cultures, Stat. Sci. 16 (3) (2001) 199–231.
[58] J.H. Friedman, Multivariate adaptive regression splines, Ann. Stat. (1991) 1–67.
[59] S. Chen, C.F. Cowan, P.M. Grant, Orthogonal least squares learning algorithm for radial basis function networks, IEEE Trans. Neural Netw. 2 (2) (1991) 302–309.
[60] C.E. Rasmussen, Gaussian processes in machine learning, in: Advanced lectures on machine learning, Springer, 2004, pp. 63–71.
[61] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, Stat. Comput. 14 (3) (2004) 199–222.
[62] M. Claesen, B. De Moor, Hyperparameter search in machine learning, arXiv preprint arXiv:1502.02127.
[63] J. Müller, Matsumoto: the matlab surrogate model toolbox for computationally expensive black-box global optimization problems, arXiv preprint arXiv:1404.4261.
[64] D. Gorissen, I. Couckuyt, P. Demeester, T. Dhaene, K. Crombecq, A surrogate modeling and adaptive sampling toolbox for computer based design, J. Mach. Learn. Res. 11 (Jul) (2010) 2051–2055.
[65] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (Oct) (2011) 2825–2830.
[66] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., Tensorflow: a system for large-scale machine learning., in: OSDI, 16, 2016, pp. 265–283.
[67] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, A. Lerer, Automatic differentiation in pytorch, 2017. https://github.com/pytorch.
[68] T. Wortmann, Opossuma model-based optimization tool in grasshopper, in: Proceedings of the 22nd CAADRIA Conference, Hong Kong, CN, 2017, pp. 283–292.
[69] A. Costa, G. Nannicini, Rbfopt: an open-source library for black-box optimization with costly function evaluations, Optim. Online 4538 (2014).
[70] T. Olofsson, S. Andersson, Long-term energy demand predictions based on short-term measured data, Energy Build. 33 (2) (2001) 85–91.
[71] B.B. Ekici, U.T. Aksoy, Prediction of building energy consumption by using artificial neural networks, Adv. Eng. Softw. 40 (5) (2009) 356–362.
[72] A. Prada, A. Gasparella, P. Baggio, On the performance of meta-models in building design optimization, Appl. Energy 225 (2018) 814–826.
[73] S. Bucking, R. Zmeureanu, A. Athienitis, A methodology for identifying the influence of design variations on building energy performance, J. Build. Perform. Simul. 7 (6) (2014) 411–426.
[74] T. Ahmad, H. Chen, Y. Guo, J. Wang, A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: a review, Energy Build. 165 (2018) 301–320.
[75] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, P.K. Tucker, Surrogate-based analysis and optimization, Prog. Aerosp. Sci. 41 (1) (2005) 1–28.
[76] E. Alpaydin, Introduction to Machine Learning, MIT press, 2009.
[77] E. Gratia, A. De Herde, A simple design tool for the thermal study of an office building, Energy Build. 34 (3) (2002) 279–289.
[78] R.E. Edwards, J. New, L.E. Parker, B. Cui, J. Dong, Constructing large scale surrogate models from big data and artificial intelligence, Appl. Energy 202 (2017) 685–699, doi:10.1016/j.apenergy.2017.05.155.
[79] H. Yi, R.S. Srinivasan, W.W. Braham, An integrated energy-emergy approach to building form optimization: use of energy plus, emergy analysis and taguchi-regression method, Build. Environ. 84 (2015) 89–104, doi:10.1016/j.buildenv.2014.10.013.
[80] H.T. Odum, Environmental Accounting: Energy and Environmental Decision Making, Wiley, 1996.
[81] L.-G. Maltais, L. Gosselin, Daylighting 'energy and comfort' performance in office buildings: sensitivity analysis, metamodel and pareto front, J. Build. Eng. 14 (2017) 61–72, doi:10.1016/j.jobe.2017.09.012.
[82] P. Geyer, S. Singaravel, Component-based building performance prediction using systems engineering and machine learning, Appl. Energy 228 (2017) 1439–1453.
[83] F. Chlela, A. Husaunndee, C. Inard, P. RiedeFer, A new methodology for the design of low energy buildings, Energy Build. 41 (9) (2009) 982–990, doi:10.1016/j.enbuild.2009.05.001.
[84] I. Korolija, Y. Zhang, L. Marjanovic-Halburd, V.I. Hanby, Regression models for predicting UK office building energy consumption from heating and cooling demands, Energy Build. 59 (2013) 214–227.
[85] T. Catalina, J. Virgone, E. Blanco, Development and validation of regression models to predict monthly heating demand for residential buildings, Energy Build. 40 (10) (2008) 1825–1832.
[86] T. Catalina, V. Iordache, B. Caracaleanu, Multiple regression model for fast prediction of the heating energy demand, Energy Build. 57 (2013) 302–312.
[87] J.S. Hygh, J.F. DeCarolis, D.B. Hill, S.R. Ranjithan, Multivariate regression as an energy assessment tool in early building design, Build. Environ. 57 (2012) 165–175, doi:10.1016/j.buildenv.2012.04.021.
[88] J.C. Lam, K.K. Wan, D. Liu, C. Tsang, Multiple regression models for energy use in air-conditioned office buildings in different climates, Energy Convers. Manag. 51 (12) (2010) 2692–2697.
[89] I. Jaffal, C. Inard, C. Ghiaus, Fast method to predict building heating demand based on the design of experiments, Energy Build. 41 (6) (2009) 669–677.
[90] A. Rackes, A.P. Melo, R. Lamberts, Naturally comfortable and sustainable: informed design guidance and performance labeling for passive commercial buildings in hot climates, Appl. Energy 174 (2016) 256–274, doi:10.1016/j.apenergy.2016.04.081.
[91] Z. Romani, A. Draoui, F. Allard, Metamodeling the heating and cooling energy needs and simultaneous building envelope optimization for low energy building design in Morocco, Energy Build. 102 (2015) 139–148, doi:10.1016/j.enbuild.2015.04.014.
[92] F. Ascione, N. Bianco, C. De Stasio, G.M. Mauro, G.P. Vanoli, Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: a novel approach, Energy 118 (2017) 999–1017, doi:10.1016/j.energy.2016.10.126.
[93] I. Jaffal, C. Inard, A metamodel for building energy performance, Energy Build. 151 (2017) 501–510, doi:10.1016/j.enbuild.2017.06.072.
[94] S. Singaravel, J. Suykens, P. Geyer, Deep-learning neural-network architectures and methods: using component-based models in building-design energy prediction, Adv. Eng. Inform. 38 (2018) 81–90.
[95] J.J. Hirsch, Associates, equest - the quick energy simulation tool, 2016, http://www.doe2.com/eQUEST/.
[96] J. Dhariwal, R. Banerjee, An approach for building design optimization using design of experiments, Build. Simul. 10 (3) (2017) 323–336, doi:10.1007/s12273-016-0334-z.
[97] C.J. Hopfe, M.T. Emmerich, R. Marijt, J. Hensen, Robust multi-criteria design optimisation in building design, in: Proceedings of building simulation and optimization, Loughborough, UK, 2012, pp. 118–125.
[98] Y.-J. Kim, Comparative study of surrogate models for uncertainty quantification of building energy model: gaussian process emulator vs. polynomial chaos expansion, Energy Build. 133 (2016) 46–58, doi:10.1016/j.enbuild.2016.09.032.
[99] P. Das, C. Shrubsole, B. Jones, I. Hamilton, Z. Chalabi, M. Davies, A. Mavrogianni, J. Taylor, Using probabilistic sampling-based sensitivity analyses for indoor air quality modelling, Build. Environ. 78 (2014) 171–182, doi:10.1016/j.buildenv.2014.04.017.
[100] G. Li, S.-W. Wang, H. Rabitz, S. Wang, P. Jaffé, Global uncertainty assessments by high dimensional model representations (HDMR), Chem. Eng. Sci. 57 (21) (2002) 4445–4460.
[101] A. Tsanas, A. Xifara, Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools, Energy Build. 49 (2012) 560–567, doi:10.1016/j.enbuild.2012.03.003.
[102] P. De Wilde, W. Tian, Predicting the performance of an office under climate change: a study of metrics, sensitivity and zonal resolution, Energy Build. 42 (10) (2010) 1674–1684.
[103] W. Tian, R. Choudhary, G. Augenbroe, S.H. Lee, Importance analysis and metamodel construction with correlated variables in evaluation of thermal performance of campus buildings, Build. Environ. 92 (2015) 61–74, doi:10.1016/j.buildenv.2015.04.021.
[104] V. Zuber, K. Strimmer, High-dimensional regression and variable selection using car scores, Stat. Appl. Genet. Mol.Biol. 10 (1) (2011).

[105] X. Chen, H. Yang, K. Sun, Developing a meta-model for sensitivity analyses and prediction of building performance for passively designed high-rise residential buildings, Appl. Energy 194 (2017) 422–439, doi:10.1016/j.apenergy.2016.08.180.
[106] X. Chen, H. Yang, Integrated energy performance optimization of a passively designed high-rise residential building in different climatic zones of china, Appl. Energy 215 (2018) 145–158, doi:10.1016/j.apenergy.2018.01.099.
[107] B. Eisenhower, Z. O'Neill, S. Narayanan, V.A. Fonoberov, I. Mezic, A methodology for meta-model based optimization in building energy models, Energy Build. 47 (2012) 292–301, doi:10.1016/j.enbuild.2011.12.001.
[108] S. Papadopoulos, E. Azar, Integrating building performance simulation in agent-based modeling using regression surrogate models: a novel human-in-the-loop energy modeling approach, Energy Build. 128 (2016) 214–223, doi:10.1016/j.enbuild.2016.06.079.
[109] I.A. Macdonald, Comparison of sampling techniques on the performance of Monte Carlo based sensitivity analysis, in: Eleventh International IBPSA Conference, 2009, pp. 992–999.
[110] S. Burhenne, O. Tsvetkova, D. Jacob, G.P. Henze, A. Wagner, Uncertainty quantification for combined building performance and cost-benefit analyses, Build. Environ. 62 (2013) 143–154.
[111] K. Soratana, J. Marriott, Increasing innovation in home energy efficiency: Monte Carlo simulation of potential improvements, Energy Build. 42 (6) (2010) 828–833.
[112] B.D. Lee, Y. Sun, G. Augenbroe, C.J. Paredis, Towards better prediction of building performance: a workbench to analyze uncertainty in building simulation, in: 13th International Building Performance Simulation Association Conference, Chambéry, France, 2013.
[113] S.L. Wong, K.K.W. Wan, T.N.T. Lam, Artificial neural networks for energy analysis of office buildings with daylighting, Appl. Energy 87 (2) (2010) 551–557, doi:10.1016/j.apenergy.2009.06.028.
[114] L. Magnier, F. Haghighat, Multiobjective optimization of building design using trnsys simulations, genetic algorithm, and artificial neural network, Build. Environ. 45 (3) (2010) 739–746, doi:10.1016/j.buildenv.2009.08.016.
[115] E. Asadi, M.G. da Silva, C.H. Antunes, L. Dias, L. Glicksman, Multi-objective optimization for building retrofit: a model using genetic algorithm and artificial neural network and an application, Energy Build. 81 (2014) 444–456, doi:10.1016/j.enbuild.2014.06.009.
[116] E.E. Aydin, O. Dursun, I. Chatzikonstantinou, B. Ekici, Optimisation of energy consumption and daylighting using building performance surrogate model, in: Living and Learning: Research for a Better Built Environment: 49th International Conference of the Architectural Science Association, 2015, pp. 536–546. ⟨GotoISI⟩://WOS:000381380100052.
[117] X. Chen, H. Yang, A multi-stage optimization of passively designed high-rise residential buildings in multiple building operation scenarios, Appl. Energy 206 (2017) 541–557, doi:10.1016/j.apenergy.2017.08.204.
[118] P.T. Boggs, J.W. Tolle, Sequential quadratic programming, Acta Numerica 4 (1995) 1–51.
[119] G.M. Stavrakakis, P.L. Zervas, H. Sarimveis, N.C. Markatos, Optimization of window-openings design for thermal comfort in naturally ventilated buildings, Appl. Math. Modell. 36 (1) (2012) 193–211, doi:10.1016/j.apm.2011.05.052.
[120] J. Carreras, C. Pozo, D. Boer, G. Guillen-Gosalbez, J.A. Caballero, R. Ruiz-Femenia, L. Jimenez, Systematic approach for the life cycle multi-objective optimization of buildings combining objective reduction and surrogate modeling, Energy Build. 130 (2016) 506–518, doi:10.1016/j.enbuild.2016.07.062.
[121] E. Gengembre, B. Ladevie, O. Fudym, A. Thuillier, A Kriging constrained efficient global optimization approach applied to low-energy building design problems, Inverse Probl. Sci. Eng. 20 (7) (2012) 1101–1114, doi:10.1080/17415977.2012.727084.
[122] D.R. Jones, M. Schonlau, W.J. Welch, Efficient global optimization of expensive black-box functions, J. Global Optim. 13 (4) (1998) 455–492.
[123] E. Tresidder, Y. Zhang, A.I. Forrester, Acceleration of building design optimisation through the use of Kriging surrogate models, Proc.Build.Simul.Optim. (2012) 1–8.
[124] R. Zhang, F. Liu, A. Schoergendorfer, Y. Hwang, Y.M. Lee, J.L. Snowdon, Optimal selection of building components using sequential design via statistical surrogate models, in: Proceedings of Building Simulation, 2013, pp. 2584–2592.
[125] S.S. Gilan, N. Goyal, B. Dilkina, Acm, Active learning in multi-objective evolutionary algorithms for sustainable building design, in: Gecco'16: Proceedings of the 2016 Genetic and Evolutionary Computation Conference, 2016, pp. 589–596, doi:10.1145/2908812.2908947.
[126] N. Beume, B. Naujoks, M. Emmerich, Sms-emoa: multiobjective selection based on dominated hypervolume, Eur. J. Oper. Res. 181 (3) (2007) 1653–1669.
[127] G. Zemella, D. De March, M. Borrotti, I. Poli, Optimised design of energy efficient building facades via evolutionary neural networks, Energy Build. 43 (12) (2011) 3297–3302, doi:10.1016/j.enbuild.2011.10.006.
[128] W. Xu, A. Chong, O.T. Karaguzel, K.P. Lam, Improving evolutionary algorithm performance for integer type multi-objective building system design optimization, Energy Build. 127 (2016) 714–729, doi:10.1016/j.enbuild.2016.06.043.
[129] A.E.I. Brownlee, J.A. Wright, Constrained, mixed-integer and multi-objective optimisation of building designs by NSGA-II with fitness approximation, Appl. Soft Comput. 33 (2015) 114–126, doi:10.1016/j.asoc.2015.04.010.
[130] T. Wortmann, Genetic evolution vs. function approximation: benchmarking algorithms for architectural design optimization, J. Comput. Des. Eng. (2018), doi:10.1016/j.jcde.2018.09.001.
[131] L. Van Gelder, H. Janssen, S. Roels, Metamodelling in robust low-energy dwelling design, in: 2nd Central European Symposium on Building Physics, Mahdavi, A, OKK-EDITIONS, Vienna, 2013, pp. 93–99.
[132] M.-Y. Cheng, M.-T. Cao, Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines, Appl. Soft Comput. 22 (2014) 178–188, doi:10.1016/j.asoc.2014.05.015.
[133] J.-S. Chou, D.-K. Bui, Modeling heating and cooling loads by artificial intelligence for energy-efficient building design, Energy Build. 82 (2014) 437–446, doi:10.1016/j.enbuild.2014.07.036.
[134] P. Symonds, J. Taylor, Z. Chalabi, M. Davies, Performance of neural networks vs. radial basis functions when forming a metamodel for residential buildings, Int. J. Civil Environ. Struct. Constr. Archit. Eng. 9 (12) (2015) 1446–1450. World Academy of Science, Engineering and Technology.
[135] S. Burhenne, D. Jacob, G.P. Henze, Sampling based on Sobol' sequences for monte carlo techniques applied to building simulations, in: Proc. Int. Conf. Build. Simulat, 2011, pp. 1816–1823.
[136] S.S. Garud, I.A. Karimi, M. Kraft, Smart sampling algorithm for surrogate model development, Comput. Chem. Eng. 96 (2017) 103–114.