BUILD SIMUL (2023) 16: 831–852
https://doi.org/10.1007/s12273-023-0984-6

Challenges and opportunities of machine learning control in building operations

Liang Zhang[1] (✉), **Zhelun Chen**[2], **Xiangyu Zhang**[3], **Amanda Pertzborn**[4], **Xin Jin**[3]

1. The University of Arizona, 1209 E 2nd St, Tucson, AZ, USA
2. Drexel University, 3141 Chestnut St, Philadelphia, PA, USA
3. National Renewable Energy Laboratory, 15013 Denver W Pkwy, Golden, CO, USA
4. National Institute of Standards and Technology, 100 Bureau Dr, Gaithersburg, MD, USA

Review Article

Building Systems and Components

Abstract

Machine learning control (MLC) is a highly flexible and adaptable method that enables the design, modeling, tuning, and maintenance of building controllers to be more accurate, automated, flexible, and adaptable. The research topic of MLC in building energy systems is developing rapidly, but to our knowledge, no review has been published that specifically and systematically focuses on MLC for building energy systems. This paper provides a systematic review of MLC in building energy systems. We review technical papers in two major categories of applications of machine learning in building control: (1) building system and component modeling for control, and (2) control process learning. We identify MLC topics that have been well-studied and those that need further research in the field of building operation control. We also identify the gaps between the present and future application of MLC and predict future trends and opportunities.

Keywords

machine learning;
building operation control;
building energy system;
reinforcement learning

Article History

Received: 23 September 2022
Revised: 09 December 2022
Accepted: 26 December 2022

1 Introduction

Globally, buildings use about 40% of primary energy and are responsible for approximately 30% of greenhouse gas emissions (Costa et al. 2013). Building control and operation have a significant impact on building energy efficiency and occupant comfort (Oldewurtel et al. 2012). Moreover, the development of the smart grid has led to a revolution in power infrastructure from centralized one-way communication to decentralized two-way communication (Zhang 2018). The National Energy Technology Laboratory estimates that more than one-fourth of U.S. electricity demand could be dispatchable from buildings through advanced whole-building control, operation strategies, and smart grid infrastructure (Hagerman 2014).

Conventional building control in most building automation systems (BAS) is rule-based feedback control, relying on pre-determined logic and schedules of building equipment operation, and realized by classical control techniques, such as proportional-integral-derivative (PID) control (Hong et al. 2020). The increasing complexity of building control tasks, such as control in building-to-grid-integration, occupancy-based control, and prediction-based control, has introduced challenges to conventional building control strategies, because: (1) building systems are massively nonlinear, which is difficult for conventional control methods to capture; (2) the system dynamics are unknown or hard to capture; (3) the dimensionality of control objectives and measurements is high; (4) predictive information (such as weather, occupancy, and occupant behaviors) is not taken into consideration, leading to sub-optimal performance (Hong et al. 2020); and (5) conventional control strategies are not sufficiently customized to the specific building and climate, and they are unable to adapt to changes (such as retrofits) to the building (Hong et al. 2020).

The increase in data availability provides opportunities

E-mail: liangzhang1@arizona.edu

Nomenclature

A3C	advantage actor-critic	MLC	machine learning control
AHU	air handling unit	MPC	model predictive control
ANFIS	adaptive neuro-fuzzy inference system	NN	neural network
BAS	building automation system	PCA	principal component analysis
BPNN	back-propagation neural network	PID	proportional-integral-derivative
DDPG	deep deterministic policy gradients	PPO	proximal policy optimization
DQN	deep Q-network	RBF	radial basis function
HVAC	heating, ventilation, and air-conditioning	RL	reinforcement learning
IoT	Internet of things	SSSS	sub-keyword synonym subtopics searching
ML	machine learning	SVM	support vector machine

for machine learning (ML) methods that rely on large amounts of data that have been deployed in the fields of engineering, manufacturing, healthcare, education, marketing, financial modeling, and policing (Jordan and Mitchell 2015). There is an explosion of building data due to decreasing hardware costs, increasing data accessibility, fast computers, fast simulations, and advances in Internet of things (IoT) and BAS. Machine learning control (MLC) combines ML, intelligent control, and control theory, to solve control problems. Although not officially defined, MLC can refer to any control where ML techniques are partially or wholly applied. Although MLC are not subject to physics-based models, and therefore lack theoretical rigor, they are highly flexible and adaptable methods. Due to its adaptability, MLC can address the challenges that conventional building controls encounter so that the design, modeling, tuning, and maintenance of building controllers can be more accurate, automated, flexible, and adaptable, making control more accessible in a real building system.

Research of MLC in building energy systems has developed rapidly in recent years, covering research topics including ML-based model predictive control (MPC), reinforcement learning (RL) control, ML-improved traditional control (e.g., PID control), ML-based fuzzy control, and ML-based feature engineering for control (Zhang and Wen 2019), applied at different levels of building control (building-level, system-level and component-level), using various ML algorithms. Although some review papers are relevant to MLC, many of them simply list the control applications of ML for building systems. Other review papers do not review the exact topic of MLC in building energy systems, but either review a broader scope, like MLC in the building life cycle (Hong et al. 2020), or a narrower scope or specific sub-category of MLC, such as neural networks (NN) (Naidu and Rieger 2011). A review that specifically and systematically focuses on MLC in building energy systems is lacking.

This paper fills that gap by providing a systematic review of MLC in building energy systems. The paper organization is presented in Figure 1. First, we introduce the paper search methodology in Section 2. Then, we summarize the review papers on MLC and identify the need for this review paper in Section 3. In Section 4, we review technical papers in two major categories of applications of ML in building control: (1) building system and component modeling for control and (2) control process learning. In Section 5, we conclude by re-emphasizing topics that are well-studied and those that are not but have enormous potential. Finally, we identify the gaps between present and future applications of MLC and predict future trends and research opportunities.

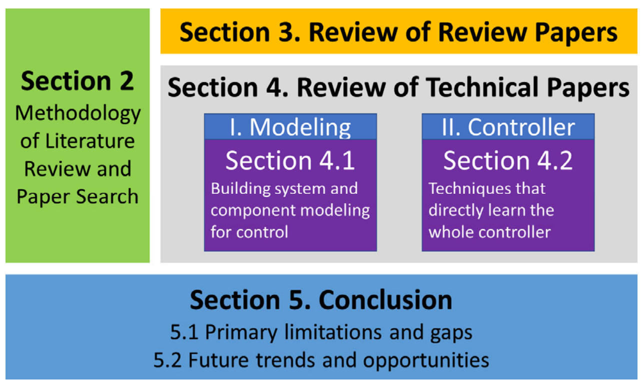

Fig. 1 Review paper structure

2 Literature search methodology: Sub-Keyword Synonym Subtopics Search

To conduct a comprehensive paper search and review, we utilize a paper search methodology called Sub-keyword Synonym Subtopics Search (SSSS) introduced in Zhang et al. (2021b). This methodology exhausts relevant papers by multiple searches with synonym and subtopic sub-keywords (Zhang et al. 2021b). For the research topic of MLC in buildings as an example, a researcher could use "machine learning control in HVAC" (heating, ventilation, and air-conditioning) instead of "machine learning control in buildings", but essentially, the two search terms reflect the same research field; one research field can have many synonyms when searching on literature databases such as Google Scholar. There are also many subtopics for one research field. For example, "reinforcement learning control in buildings" and "ML-based PID control in buildings" are both relevant topics to this review paper. As a result, a mechanism is needed to conduct multiple searches to exhaust the search keywords reflecting synonym and subtopic, and subsequently to exhaust the relevant papers. In SSSS methodology, each search keyword consists of multiple sub-keywords: there are multiple synonyms or subtopics for each sub-keyword, and the set of search keywords consists of the full combination of every possible value of the sub-keywords (Zhang et al. 2021b).

We use keywords composed of three sub-keywords in this paper. The first sub-keyword defines the subtopics of the "machine learning" concept. The list of the first sub-keyword is: "machine learning," "reinforcement learning," "deep learning," "neural network," "intelligent," "genetic algorithm," "self-tuning," "self-learning," and "advanced." The second sub-keyword defines subtopics of control applications. The list of the second sub-keyword is: "control," "daylight control," "windows control," "MPC," "model predictive control," "PID control," and "fuzzy control." The list of the third sub-keyword is: "in buildings," "HVAC," and "building system," which narrows the search within building energy systems.

We use Google Scholar as the search engine in this paper, and the list of search keywords in Google Scholar is the full combination of each element in each sub-keyword list. One example search keyword was: "machine learning, control, in buildings." The total number of search keywords in this paper is $9 \times 7 \times 3 = 189$. As summarized in Table 1, in this paper, the first 10 papers per search are considered; the publication year of the papers ranges from 2010 to 2021; and the citation threshold is five for papers published before 2018. The SSSS methodology is automated using a module coded in Python (https://github.com/lz356/SSSS (Zhang et al. 2021b)). The total number of papers returned by the search is: 189 keywords × 10 papers/keyword = 1,890 papers, but there are duplicate papers found in the search, so the final number of non-duplicate papers is 850.

Table 1 Parameters of SSSS for ML control in buildings

Parameter	Values
Sub-keyword 1	Machine learning, reinforcement learning, deep learning, neural network, intelligent, genetic algorithm, self-tuning, self-learning, advanced
Sub-keyword 2	Control, daylight control, windows control, MPC, model predictive control, PID control, fuzzy control
Sub-keyword 3	In buildings, HVAC, building system
Citation threshold	5 for papers before 2018
Number of papers per search	10
Year from	2010
Year to	2022

Of those 850 papers, we included 101 papers in this review article based on their relevance, novelty, and quality. Specifically, we selected a paper if it met any of the following criteria: (1) a classic paper published in a prestigious journal with at least 50 citations, (2) a new paper applying novel methodologies and algorithms in MLC, and (3) a paper with insights and conclusions that benefit the discussion. Some papers were included in the citations of the paper found in the search instead of directly from the search; these papers were manually included for review. Figure 2 is a relational graph that shows the research topics/keywords and citation relationships among the papers that are reviewed in this paper. As indicated by the largest circles, the most discussed topics in the papers related to HVAC and energy consumption are NNs, model predictive control, and reinforcement learning. The most trending keywords we reviewed about MLC in buildings include demand flexibility, occupant behavior, and deep neural network.

3 Review of review papers

We summarize the existing review papers on MLC in Table 2. Most review papers do not include papers with the exact scope of MLC in building energy systems. This first set of review papers had the narrow scope of specific control applications using ML. Salimi and Hammad (2019) reviewed occupancy monitoring-based control, and ML methods are only one of the classes of applied techniques. Han et al. (2019) reviewed RL methodologies for controlling occupant comfort in buildings. They first introduced the general RL method and then reviewed applications of RL for comfort control in buildings.

The next set of papers had a narrow focus on specific ML algorithms in MLC. Naidu and Rieger (2011) reviewed MLC applications in terms of hard-control techniques, soft-control techniques, and the fusion of hard- and soft-control techniques, using only NN as the ML algorithm. Vázquez-Canteli and Nagy (2019) reviewed energy systems and RL methods and RL applications in demand response modeling and dynamic response.

The third set of papers had a larger scope that included MLC in one section of the review. Hong et al. (2020) reviewed ML in the building life cycle. They reviewed two

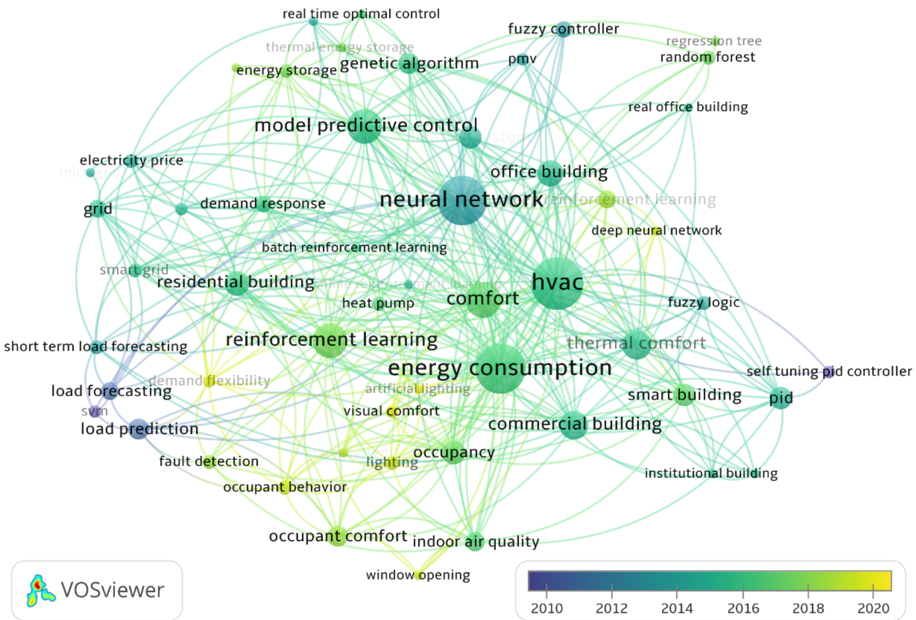

Fig. 2 Relational graph reflecting covered topics, keywords, and citation relationships

categories of MLC: ML-based MPC and RL-based control. Ahmad et al. (2016) reviewed computational intelligence techniques for HVAC systems, and MLC is only one of the applications; NN is the only ML algorithm included in the review.

To sum up Table 2, in terms of algorithms, RL control (Han et al. 2019; Mason and Grijalva 2019; Vázquez-Canteli and Nagy 2019; Wang and Hong 2020) and NN control (Naidu and Rieger 2011; Mirinejad et al. 2012; Kumar et al. 2013; Ahmad et al. 2016; Hidalgo-León et al. 2019; Wagiman et al. 2020) are the most widely reviewed algorithm topics for MLC in buildings; in terms of applications, there are papers with larger (Ahmad et al. 2016; Hidalgo-León et al. 2019; Hong et al. 2020) and smaller scopes (Mirinejad et al. 2012; Han et al. 2019; Salimi and Hammad 2019; Vázquez-Canteli and Nagy 2019), but no review paper is found to focus on the exact scope of MLC in building energy systems. This paper can fill this gap. The target readers of this paper are not only researchers from the building industry who would be exposed to cutting-edge machine learning control research but also those from the machine learning industry who could gain an understanding of the potential and challenges in applying MLC in buildings, which provides a unique perspective that other existing review papers cannot provide.

4 Review of technical papers

The ML application that gains the most attention is developing ML models for the building system and components. As mentioned in Section 1, in real buildings, equipment, systems, and envelopes have dynamics that are not fully known, so the modeling of them can be learned and formulated just using data with ML techniques. Building energy modeling, equipment performance modeling (e.g., coefficient of performance modeling), and indoor environment (e.g., room temperature) modeling, are typical applications in this category. Instead of spending a lot of effort building an ML model from data and then integrating it into the control and optimization process, a second way of applying ML is to directly learn the whole controller with ML techniques. RL is a typical application that directly learns a controller.

The rest of this section will review the technical papers. Section 4.1 reviews papers related to building system and component modeling for control, while Section 4.2 is a review of papers that learn the control process.

4.1 Building system and component modeling for control

Building system and component modeling is a major application of ML in building control systems. Real building equipment, systems, envelopes, indoor and outdoor environments, have dynamics that are not fully known, and the modeling of them can be overly complex, nonlinear, time-consuming, and uncertain. With adequate data and suitable model formulation, ML techniques can capture the complex dynamics in real buildings using data with limited expert knowledge. After the extended search, the existing literature was categorized into two groups according to the research focus. The first group focuses on ML modeling for the building system and components that has the potential

Table 2 Summary of review papers

No.	Reference	Topics	Scope
1	Hong et al. 2020	- Problem formulation - MPC - RL-based control	Reviewed ML in the building life cycle; MLC is one section
2	Han et al. 2019	- Introduction of general RL methods - Applications of RL for comfort control in buildings	Reviewed RL methodologies for controlling occupant comfort in buildings
3	Salimi and Hammad 2019	- Review of control systems in occupancy monitoring-based control	Only reviewed occupancy monitoring-based control, and ML is one of the classes of applied techniques
4	Naidu and Rieger 2011	- Hard-control techniques, such as PID, optimal, robust, and adaptive control - Soft-control techniques, such as NNs, fuzzy logic, and genetic algorithm - Fusion of hard- and soft-control techniques	Not focus on ML but reviewed some NN applications
5	Vázquez-Canteli and Nagy 2019	- Energy systems and RL methods - Demand response modeling and dynamic response	Focused on RL for demand response
6	Ahmad et al. 2016	- Review of computational intelligence techniques: genetic algorithm, NN, PCA (principal component analysis), etc.	Focused on computational intelligence techniques, including NN
7	Wang and Hong 2020	- What RL algorithms are used in what building control problems - How these control problems are modeled as RL problems - How states/rewards are set in existing literature - Discussion on remaining issues	Reviewed RL for building control problems, including controls for whole building, HVAC, and water heater
8	Mason and Grijalva 2019	- Introduction of general RL methods - Reviewed by control application: HVAC, water heating, home management system and grid-interactive applications	Reviewed RL for building control problems
9	Kumar et al. 2013	- Reviewed different NN models used for building energy analysis - Reviewed different NN applications in the following areas: load estimation, indoor air temperature prediction, energy consumption prediction, and HVAC system	Very limited content about building system control. The paper focused on NN in terms of ML algorithms
10	Mirinejad et al. 2012	- Reviewed applications of intelligent controls including fuzzy controller and auto-tuned PID	Focused on fuzzy control, where NN may or may not be employed
11	Wagiman et al. 2020	- Reviewed by control technique: controller-based, optimization-based, and hybrid of both	Reviewed limited ML applications. Some controllers might be NN-based, but not discussed in detail
12	Hidalgo-León et al. 2019	- Reviewed building energy consumption reduction techniques: occupancy detection, HVAC system control, lighting control, and energy prediction and estimation	Discussed building system control, but not based on ML in general. It contains one lighting control application that includes NN
13	Merabti et al. 2016	- Reviewed applications of intelligent controls include PID, fuzzy, fuzzy PID, adaptive fuzzy PID, NN, neuro-fuzzy, and genetic algorithm	Compared pros and cons of different intelligent control methods. Some methods are NN-based
14	Afroz et al. 2018	- Physics-based model for control - Data-driven model for control - Gray-box model for control	Reviewed modeling techniques used in building HVAC control systems, data-driven is only one part and only the algorithm of NN is reviewed

for future use in control. These papers focus on the algorithms and development of ML models. The second group of literature focuses on control and ML modeling is only a part of the research. These papers focus on the control side with the integration of ML techniques.

The model-focused papers focus on ML modeling in building equipment, systems, envelopes, occupancy, indoor and outdoor environment that has the potential for future control purposes. These papers discuss the formulation of models in terms of data collection and processing, ML algorithm, algorithm tuning and testing, supporting techniques, and so on. We reviewed ML modeling for building energy and cooling/heating load (Yildiz et al. 2017; Seyedzadeh et al. 2018; Zhang et al. 2021b), building equipment (chiller, cooling tower, etc.) energy and operation (Mosavi et al. 2019), indoor air temperature and humidity (Cifuentes et al. 2020), occupancy (Dai et al. 2020), daylighting (Ayoub 2020), indoor air quality (Wei et al. 2019), thermal comfort (Ma et al. 2021), and outdoor weather forecasting (Lazos et al. 2014). The papers cited in the previous sentence are review papers, and representative technical papers are summarized in Table 3, which highlights the specific ML algorithms and the modeling output. ML algorithms include NN, support vector machines (SVM), tree-based algorithms, deep learning, hybrid algorithms, ensemble learning, autoregressive algorithms, Bayesian networks, extreme learning machines, case-based reasoning, meta learning, k-nearest neighbors, Gaussian process and mixture models, and fuzzy timeseries algorithms. Table 3 shows that building energy consumption and cooling/heating load modeling applies more diverse ML algorithms than other applications and there is a trend that other applications are using more diverse ML algorithms. The modeling-focused studies are well-explored by many researchers, covering sub-topics such as parameter tuning/optimization, model improvement by supporting techniques such as clustering, and interpretation of ML models. These papers focus on the formulation of modeling while integration between the models and the controllers is not thoroughly discussed.

The control-focused papers focus on further integrating ML models to control. Specifically, papers reviewed in this section focus on control studies in which ML techniques are applied to model the building system and components in control. The modeling includes building energy consumption, HVAC energy and performance, chiller energy and performance, occupancy, cooling load, indoor (zone

Table 3 Representative studies on building system and component modeling for control

No.	Reference	ML algorithm	Modeling output	Model or control focused
1. Building energy modeling				
1.1	Fan et al. 2017	Deep neural network, multiple linear regression, elastic net, random forests, gradient boosting machines, support vector machine, extreme gradient boosting trees	Short-term day-ahead building cooling load	Model
1.2	Monfet et al. 2014	Case-based reasoning	Commercial building energy load	Model
1.3	Cui et al. 2016	Meta learning	Short-term building energy load	Model
1.4	Ahmad et al. 2017b	Random forest and NN	Building energy load	Model
1.5	Lusis et al. 2017	NN, regression trees, and SVM	Day-ahead residential building energy load	Model
1.6	Yu et al. 2010	Decision tree and NN	Residential building energy load modeling and building energy performance indexes	Model
1.7	Kwok and Lee 2011	Probabilistic entropy-based NN	Building cooling load	Model
1.8	Guo et al. 2018	Extreme learning machine, multiple linear regression, SVM and BPNN	Building heating load	Model
1.9	Idowu et al. 2016	SVM, regression tree, feed forward NN, and multiple linear regression	District heating load	Model
1.10	Wahid and Kim 2016	k nearest neighbor	Building energy load	Model
1.11	Cheng and Cao 2014	Multivariate adaptive regression splines and artificial bee colony	Building energy performance	Model
1.12	Fan et al. 2019a	Deep recurrent NN	Building energy load	Model
1.13	Kamel et al. 2020	NN, fuzzy inductive reasoning, Lasso regression, SVM	Building energy load	Control
1.14	Jain et al. 2017	Regression trees and ensemble learning	Building energy load	Control
1.15	Smarra et al. 2018	Random forests	Building energy load	Control
1.16	Lee et al. 2015	NN	Building energy load	Control
1.17	Wang et al. 2019	Long short-term memory networks	Building internal load	Control
1.18	Cole et al. 2014	NN models	Building energy load	Control
1.19	Chen et al. 2015	NNs and SVM	Home energy load	Control
1.20	Huang et al. 2015b	NN	HVAC system load	Control
1.21	Manjarres et al. 2017	Random forest	HVAC system load	Control
1.22	Cox et al. 2019	Non-linear autoregressive with exogenous inputs NN	Thermal load of district cooling load	Control
1.23	Verrilli et al. 2017	Generalized regression NN	Building energy load	Control

Table 3 Representative studies on building system and component modeling for control (Continued)

No.	Reference	ML algorithm	Modeling output	Model or control focused
2. Equipment energy and performance modeling				
2.1	Swider et al. 2001	NN	Chiller performance	Model
2.2	Hosoz et al. 2007	NN	Thermal performance of cooling tower	Model
2.3	Karunamurthy et al. 2020	Linear regression	Thermal performance of cooling tower	Model
2.4	Kim et al. 2019	NN	Energy consumption model of chiller	Model
2.5	O'Neill and O'Neill 2016	Bayesian networks	HVAC hot water energy consumption	Model
2.6	Sala-Cardoso et al. 2018	Adaptive neuro-fuzzy inference system (ANFIS)	HVAC power demand	Control
2.7	Kim and Park 2014	Gaussian process model	Chiller energy consumption	Control
2.8	Chow et al. 2002	NN	Performance of chiller	Control
2.9	Park et al. 2019c	NN	Chiller operation	Control
2.10	Yang et al. 2021	NN	Optimal cooling power set point	Control
2.11	Park et al. 2019c	NN and hybrid NN	Chiller power consumption	Control
2.12	Park et al. 2019b	NN	Daily operation schedule of air handling unit (AHU) and chiller	Control
2.13	Kumar et al. 2021	Linear and logistic regression	AHU fan speed operation	Control
3. Indoor air temperature and humidity				
3.1	Mba et al. 2016	NN	Indoor temperature and relative humidity	Model
3.2	Potočnik et al. 2019	NN, ARX, and extreme learning machine	Indoor temperature	Model
3.3	Alawadi et al. 2022	36 machine learning algorithms	Indoor temperature	Model
3.4	Mustafaraj et al. 2011	NN	Indoor temperature and relative humidity	Model
3.5	Xu et al. 2019	Long short-term memory	Indoor air temperature	Model
3.6	Liang et al. 2015	Auto-regressive moving average exogenous	Return air temperature	Control
3.7	Ma et al. 2012	Autoregressive exogenous	Zone air temperature and power measurement	Control
3.8	Li et al. 2013	BPNN	Room temperature	Control
4. Indoor air quality				
4.1	Tang et al. 2014	Multilayer perceptron ensemble	Building energy consumption and air quality index	Model
4.1	Liu et al. 2013	Output error model	Indoor air quality	Control
4.2	Huang et al. 2015a	NN	Indoor air temperature and building energy consumption	Control
4.3	Kim et al. 2016	NN	Indoor air temperature, building energy consumption, and daylight illuminance	Control
4.4	Yang et al. 2020	NN	Indoor PMV data	Control
5. Thermal Comfort				
5.1	Chaudhuri et al. 2017	SVM, NN, logistic regression, linear discriminant analysis, k-nearest neighbors, and classification trees	Thermal comfort	Model
5.2	Wu et al. 2018	Ensemble machine learning	Thermal comfort (thermal sensation, effective temperature, standard effective temperature and PMV	Model
5.3	Garnier et al. 2014	Feedforward NN	Non-linear behavior of the PMV index (thermal comfort)	Control
5.4	Zhou et al. 2015	Convex piecewise linear classifier	Thermal comfort	Control
5.5	Ruano et al. 2016	Radial basis function (RBF) NNs	PMV and Energy estimation	Control
6. Occupancy				
6.1	Yang et al. 2014	SVM	Occupancy	Model
6.2	Chen and Jiang 2018	Generative adversarial network	Occupancy	Model
6.3	Ryu and Moon 2016	Decision tree	Occupancy	Model

Table 3 Representative studies on building system and component modeling for control (Continued)

No.	Reference	ML algorithm	Modeling output	Model or control focused
6.4	Jin et al. 2021	1-week seasonal period with an NN structure	Occupancy	Model
6.5	Li and Dong 2018	NN and SVM	Occupancy	Control
6.6	Mosaico et al. 2019	Deep transfer learning	Occupancy	Control
7. Daylighting				
7.1	Ahmad et al. 2017a	Random forest and NN	Daylighting	Model
7.2	Beccali et al. 2018	NN	Daylighting	Control
7.3	Hu and Olbina 2011	NN	Daylighting (for split blinds control)	Control
8. Weather forecast				
8.1	Florita and Henze 2009	Moving average models and NN	Weather forecasting model building applications	Model
8.2	Dong and Lam 2014	recurrent NN	Weather forecasting model for building applications	Control
8.3	Javed et al. 2014	Novel random NN compared with similar NN	Energy consumption in buildings, and solar radiation prediction	Control
8.4	Finck et al. 2019	NN	Short-term solar radiation, heating system energy load, energy load of all thermal zones	Control
8.5	Wei et al. 2015	Chi-squared automatic interaction detector, boosting tree, random forest, multi-layer perceptron, multivariate adaptive regression splines, and SVM	Energy consumption, indoor air temperature, indoor air humidity, and CO2 concentration	Control
9. Others				
9.1	Atabay et al. 2013	NN	Building thermal behavior	Control
9.2	Parisio et al. 2014	Least-square SVM	Renewable power generation and the demand for the day ahead	Control
9.3	Oldewurtel et al. 2010	Least-squares SVM	Electricity tariff price forecasting	Control
9.4	Yousaf et al. 2021	ANFIS and SVM	Electricity price forecasting	Control

temperature and thermal comfort) and outdoor environment. Table 3 also lists the literature in terms of the ML algorithms and the quantity of interest that the algorithms try to predict. The most commonly predicted variables for control-focused ML studies are whole building (Chen et al. 2015), chiller (Park et al. 2019c), HVAC system (Manjarres et al. 2017), district cooling (Cox et al. 2019), indoor air temperature (Xu et al. 2019), equipment temperature (Liang et al. 2015), thermal comfort (Zhou et al. 2015), occupancy (Li and Dong 2018), weather (Finck et al. 2019), building thermal behavior (Atabay et al. 2013), electricity tariff (Oldewurtel et al. 2010), and renewable energy generation (Parisio et al. 2014). In terms of algorithms selected, NNs, SVM, and autoregressive methods are found in the literature. The ML algorithms selected in the control-focused literature are much less diversified than those in the modeling-focused literature.

HVAC plant structure is very complex where time varying system dynamics, slow-moving processes with time delays, and non-ideal behavior of actuators prevail, and substantial disturbances, constraints, and uncertainties are imposed by running the total HVAC system dynamics (Afroz et al. 2018). ML modeling plays a key role in precise and automated modeling of the system in a data-driven way by handling disturbances, constraints, and uncertainties existing within the HVAC system dynamics, thus benefiting model-based control.

As observed from the literature review, ML has been widely applied to almost every component and modeling task in the building. However, ML is not always the best way to model a system or component, and the application of ML should be more selective in the modeling task. In other words, there is still a large gap between “we can use ML modeling” and “we should use ML modeling.” Comparison between ML modeling methods and other non-ML modeling methods should be conducted in a standard and comprehensive way. These comparisons are essential to demonstrate the advantages of ML modeling for control and persuade industry to apply ML in complex system modeling and control.

Although many papers claim that their ML models are built for control, the integration of the ML model with control is not always well covered. When applied to and integrated with control applications, more problems will appear, such as inference speed, optimization compatibility, abnormal inference, and interpretability. These topics should also be covered by conducting comprehensive experiments in real systems. Most papers examine the feasibility of ML in certain modeling tasks and prove its increased accuracy. Although accuracy is the most important metric for modeling, it is not always the most important metric in the context of control. The reliability and extendibility of the model are also important, especially when the model inferences are based on inputs that are vastly different from the training data.

In summary, the key challenges for building system and component modeling for control are threefold. First, more automated, streamlined, plug-n-play, and scalable ML modeling as well as validation is lacking. Second, using biased data or data with limited range of inputs in ML model training and development is not reliable for predicting unseen situations (Zhang 2021; Zhang and Wen 2021). Third, more studies focusing on improving the computation efficiency of dealing with big data and optimization in the modeling and application process are needed, since ML modeling is dealing with big data and sophisticated optimizations, and it is challenging in terms of computational capability and efficiency to apply it for control in real-time.

4.2 Control process learning

Instead of using ML to model control components, the papers reviewed in this section cover techniques that directly learn the whole controller using ML techniques. From the reviewed literature, learning the control can be categorized into three groups: RL (Section 4.2.1), ML-based PID control (Section 4.2.2), and ML-based fuzzy control (Section 4.2.3). We summarize the challenges and opportunities in Section 4.2.4.

4.2.1 Reinforcement learning (RL)

In recent years, especially since 2017, there has been an increase in the number of studies investigating the use of deep RL for building optimal control. Compared with traditional building optimal control approaches (e.g., MPC), RL controllers are expected to achieve a similar if not better control performance while reducing implementation costs. Specifically, as discussed in Zhang et al. (2021d), although online MPC is considered mainstream in building optimal control, it requires the solution of an optimization problem on-the-fly within each control interval, which might only be attainable on advanced computing platforms that are not cost-effective. Explicit MPC, though able to remove the dependency on on-demand computation, is only suitable for small scale problems with a shorter prediction horizon and smaller state dimension (Mayne 2014). In addition to the high cost of computing hardware and optimization software, MPC-based approaches often require an accurate yet simple building model that can be formulated into the optimization problem (i.e., differentiable). This inevitably increases modeling costs. In contrast, a controller based on RL does not require intensive real-time computation and the control actions are obtained by evaluating an RL control policy, which can be easily deployed on an edge device (i.e., cost effective devices with limited computation capability) due to the light online computation requirement. Admittedly, the offline computational requirements for RL are significant, but this can be addressed by using cloud computing system with affordable costs. In addition, RL does not require a mathematically expressible building model; instead, the building model can take the form of a first-principle simulator (e.g., EnergyPlus™), a data-driven model (e.g., non-differentiable ML model) or even a real building in some cases (e.g., when the RL controller is already properly pretrained). As a result, RL controllers stand out as good candidates for building optimal control, with lower implementation costs when compared with state-of-the-art controllers. We review existing RL literature from several different perspectives.

Table 4 contains the list of RL papers reviewed, and it includes the following information for each paper: the system to be controlled (e.g., HVAC), the level of control (e.g., setpoint calculation, device on/off determination), the RL algorithm(s), the data source, and the relevant time intervals used in the study. In addition, the table includes information about the reward structure, which is a key element of RL. These topics are discussed in the next few paragraphs.

Some papers summarize the characteristics, limitations, pros, and cons of RL modeling in buildings. Chen et al. (2018) mentioned the limitation of RL control: it requires a sufficiently long learning period before it can make optimal decisions under various conditions. In real-world practice, this prerequisite may cause difficulty, but can be alleviated by the assistance of building simulations. Yang et al. (2015) concluded that the most attractive features of RL controls are: (1) direct application into a real-world scenario, (2) limited prior knowledge requirements, (3) self-adaptation to the local environment, and (4) self-adjustment to input variations.

In terms of building type, most papers investigate applications in commercial buildings, a few studies focus on the residential sector, and some do not mention the specific building type. The controllable inputs for RL can be categorized into component level control and system level control. Only three papers include component level control:

Table 4 Summary of RL research

No.	Reference	System type	Level (details) of control	RL algorithms	Data source	Sample time/ control interval/ control horizon	Action space and reward structure	Highlighted features
1	Wei et al. 2017	HVAC	Airflow control for every zone	Deep Q-network (DQN)	Simulation: EnergyPlus co-simulated with BCVTB	1min/15min/ 1 month	Discrete action space Reward: energy consumption and thermal discomfort (temperature only)	NA*
2	Chen et al. 2018	HVAC and window	Heating, cooling, on/ off, window opened and closed for natural ventilation	Tabular Q-learning	Simulation: EnergyPlus	20 min/ 20 min/1 yr.	Discrete action space (5) Reward: energy consumption and occupant discomfort (temperature and humidity).	Mentioned the limitation of RL control
3	Yang et al. 2015	HVAC and renewable energy system	Photovoltaic thermal (PV/T) collectors, geothermal heat pump	Tabular Q-learning and Batch Q-learning with Memory Replay	Real building: residential building in Zurich, Switzerland, compared with simulation results	30 min (PV/T) & event-driven (heat pump)/ 30 min (PV/T) & event-driven (heat pump)/1 month	Discrete action space Reward: Case 1: net power output (thermal energy collected minus electricity consumption); Case 2: heat compensation and heat supply	Summarized advantages and disadvantages of RL
4	Ruelens et al. 2017	HVAC	Heat-pump thermostat	Batch RL: fitted Q-iteration	Simulation: equation-based (model-free)	15 min/ 15 min/24 hr.	Discrete action space Reward: two cost functions related to demand response	NA
5	Barrett and Linder 2015	HVAC	AC on/off (heating only)	Bayesian Inference & Q-learning	Simulation: equation-based (model-free)	1 min/1 min/NA	Discrete action space (4) Reward: heuristic rule learning.	NA
6	Fazenda et al. 2014	HVAC	Electric heater on/off	Q-learning	Simulation: MATLAB	dynamic/ dynamic/ 24 hr.	Discrete action space (2) Reward: tenant preference violation and heating cost.	Component level control
7	Zhang and Lam 2018	HVAC	Radiant heating system	A3C	Real building	5 min/5 min / 3 months	Discrete action space (11) Reward: thermal comfort (PPD) and energy consumption	Real building experiments
8	Li and Xia 2015	HVAC	Cooling setpoint	Multi-grid method of Q-learning	Simulation: MATLAB and EnergyPlus	15 min/ 15 min/NA	Discrete action space Reward: thermal comfort (PPD) and energy consumption	NA
9	Overgaard et al. 2019	HVAC (hydronic based heating)	Mixing loop control valve	TD-learning with eligibility trace	Real building calibrated simulation: office building in Bjerringbro, Denmark.	5 min/5 min /NA	Discrete action space Reward: temperature deviation and energy consumption	Real building calibrated building model
10	Zhang et al. 2019b	HVAC and water	Mullion system supply water temperature setpoint	A3C	Real building calibrated EnergyPlus simulation	5 min/15 min/ 3 months	Discrete action space (11) Reward: thermal comfort (PPD) and energy consumption	NA
11	Ding et al. 2019	HVAC, lighting, blind and window	Building's subsystems, including HVAC, lighting, blind and window systems	Branching Dueling Q-Network (BDQ)	Real building calibrated EnergyPlus simulation	15 min/15 min/ 1 month	Discrete action space (2e6) Reward: thermal comfort, visual comfort, indoor air quality and energy consumption	NA
12	Park et al. 2019a	Lighting	Occupant centered controller for lighting	Dynamic programming, solved by value iteration	Real building: Ernest Cockrell Jr. Hall	1 min/≥1 min/NA	Discrete action space (3) Reward: visual comfort and energy consumption	One of the very few RL papers related to lighting
13	Han et al. 2020	Window	Window opening and closing	Q-learning and SARSA	Real building: office building	10 min/10 min/ 24 hr.	Discrete action space (2) Reward: thermal and indoor air quality discomfort	One of the very few RL papers use real building
14	Costanzo et al. 2016	HVAC	Heating system on/off control	RL, fitted Q-iteration and NN	Real building calibrated simulation	5 min/5 min/ 8 hr.	Discrete action space (2) Reward: energy consumption only.	A backup controller enforces indoor temperature comfort band.
15	Gao et al. 2019	HVAC	Indoor air temperature and humidity setpoint	Deep RL: deep deterministic policy gradients (DDPG)	Simulation: TRNSYS	30 min/30 min/ 24 hr.	Continuous action space (2) Reward: thermal discomfort and energy consumption	NA

Table 4 Summary of RL research (Continued)

No.	Reference	System type	Level (details) of control	RL algorithms	Data source	Sample time/ control interval/ control horizon	Action space and reward structure	Highlighted features
16	Nagy et al. 2018	HVAC	Space heating power	Double deep neural fitted Q-iteration	Simulation	1 hr./1 hr./24 hr.	Discrete action space (6) Reward: thermal discomfort and energy consumption	NA
17	Zhang et al. 2019a	HVAC	HVAC scheduling of zone temperature	Model-based RL	Simulation: EnergyPlus	Dynamic (1-15 min)/ 15 min/75 min	Discrete action space (4) Reward: energy consumption and thermal discomfort (temperature deviation)	NA
18	Peng and Morrison 2016	HVAC	Heat pump thermostat	Model predictive prior reinforcement learning with the adaptive set-point temperature algorithm	Simulation: equation-based gray box	NA	Discrete action space (60) Reward: thermal discomfort (temperature deviation) and energy consumption	NA
19	Kazmi et al. 2018	Water	Hot water system operations (control on hot water production)	Model-based DRL: Partially Observable Markov Decision Process (POMDP)	Real building: thirty-two houses in the Netherlands	(5–15 min)/ (5–15 min)/NA	Discrete action space (2) Reward: thermal discomfort, energy consumption and exploration bonus	NA
20	Chen et al. 2019	HVAC and water	Supply water temperature	Gnu-RL, A Precocial RL	Simulation (EnergyPlus) and real building	5 min/15 min/ 24 hr.	Continuous action space (1) Reward: thermal discomfort (state setpoint deviation) and control effort (magnitude of control signal)	An MPC based non-NN RL policy.
21	Zhang et al. 2021d	HVAC	HVAC ON/OFF signal for multiple zones split AC units.	A3C and Ape-X DQN	Simulation: equation-based	5 min/5 min/ 4 hr.	Discrete action space Reward: thermal discomfort (comfort margin and temperature violation) and power consumption.	NA
22	Raman et al. 2020	HVAC	Control for cooling and dehumidifying coil.	Zap Q-learning	Simulation: equation-based (RC model)	NA	Continuous action space Reward: thermal discomfort (temperature and humidity violation) and power consumption	NA
23	Zhang et al. 2020b	HVAC	Airflow control for every zone and chiller discharge air temperature	A two-stage policy search framework combining Evolution strategy-based RL and PPO	Simulation: equation-based (reduced order model)	5 min/5 min/ 24 hr.	Continuous action space (6) Reward: thermal discomfort (T deviation), power consumption and grid service violation	NA

* NA: not available

electric heater on/off in a residential building (Fazenda et al. 2014), window opening and closing (Han et al. 2020), and mixing loop control valve (Overgaard et al. 2019). In other papers, RL is used for system level control for zone airflow (Wei et al. 2017), HVAC system heating/cooling control (Barrett and Linder 2015; Li and Xia 2015; Costanzo et al. 2016; Chen et al. 2018; Nagy et al. 2018), window system (Chen et al. 2018), distributed energy resources (Yang et al. 2015; Touzani et al. 2021), heat pump operation (Yang et al. 2015; Peng and Morrison 2016; Ruelens et al. 2017), radiant heating system (Zhang and Lam 2018; Zhang et al. 2019b), lighting system (Park et al. 2019a), indoor temperature control (Gao et al. 2019; Zhang et al. 2019a), hot water system (Kazmi et al. 2018), and all building sub-systems (Ding et al. 2019).

Designing a proper reward structure for an RL controller and aligning it with the desired control objective is extremely important. In the domain of building control, two major objectives are: (1) improving the occupant's comfort and (2) minimizing energy consumption. In addition to these two commonly used objectives, under the grid-interactive efficient building (GEB) framework, some studies include rewards related to providing grid services. Although it is common to see some papers use arbitrary numbers (e.g., 0 or −1) to define the reward/penalty under certain circumstances (e.g., Barrett and Linder 2015), in general it is better if the reward system is interpretable, for instance, a weighted-sum objective, and setting the reward to negative control costs (e.g., Chen et al. 2018; Ding et al. 2019).

The action space for RL controllers is usually defined in a straight-forward manner, namely, it is determined by the number and type of control variables. In general, the action space can be either discrete (Wei et al. 2017; Zhang et al. 2019b) or continuous (Raman et al. 2020; Zhang et al. 2020b). Among existing papers, those that consider discrete action space outnumber the ones with a continuous action space for two potential reasons: (1) in some problems discretization of control variables is straight-forward and does not cause significant performance deterioration; (2) problems with continuous action space are, generally speaking, more difficult to learn due to a much larger policy search space. However, due to algorithmic and computational developments, learning an RL control policy for continuous control problems has become tractable.

Another important factor is the availability of the learning environment; although theoretically real buildings can be used for learning, it is not desirable for two major reasons. First, RL is notorious for requiring a large amount of experience to train an optimal controller and directly learning from a building limits the sample efficiency to real-time, so it takes a long time to gain sufficient experience during training (Liu and Henze 2006; Chen et al. 2018). Second, the learning process requires exploration, which will inevitably lead to actions yielding high costs or undesirable consequences in real life, especially at the beginning of the learning phase. As a result, most papers use simulations for the learning environment. Among these papers, some use real-building calibrated simulations (Costanzo et al. 2016; Ding et al. 2019; Overgaard et al. 2019; Zhang et al. 2019b). A few studies use real buildings to validate the controller trained by simulation (Kazmi et al. 2018; Zhang and Lam 2018; Park et al. 2019a; Han et al. 2020), which has many advantages.

Based on the available learning environment and the nature of the action space, an appropriate RL algorithm is chosen. If learning is simulation based and running the simulation is computationally inexpensive, on-policy learning algorithms such as asynchronous advantage actor-critic (A3C) (Mnih et al. 2016) and proximal policy optimization (PPO) (Schulman et al. 2017) can be utilized, for instance (Zhang and Lam 2018; Zhang et al. 2021d). In cases where the collection of learning data is less efficient, i.e., learning from a slow simulation, off-policy algorithms with higher sample efficiency can be considered. Among these algorithms, Q-learning and its variants, starting from the tabular form (Chen et al. 2018) to deep Q-network (DQN) (Wei et al. 2017), stand out and are utilized in many studies. In addition, methods like Deep Deterministic Policy Gradients (DDPG) (Gao et al. 2019) and Zap Q-learning (Raman et al. 2020) extend the Q-learning philosophy to problems with a continous action space.

Regarding the length of the control horizon and each interval, the control interval varies from 1 minute to 1 hour. A shorter interval is often more precise, however, a longer interval may be used for reasons such as: (1) limited resolution of the sample data, (2) the balance of simulation precision and speed (Zhang et al. 2019a) and (3) a longer response time of the control device to an action (Zhang et al. 2019b). The training episode is usually decided based on the hourly (Costanzo et al. 2016; Zhang et al. 2019a), daily (Ruelens et al. 2016; Nagy et al. 2018), or seasonal (Yang et al. 2015; Zhang and Lam 2018; Zhang et al. 2019b) patterns of the system operations.

4.2.2 ML-based PID control

The objective in tuning a PID loop is to adjust its output to move the process variable as quickly as possible to the setpoint while minimizing overshoot, and then holding the variable steady at the setpoint without excessive control changes. The optimized parameters (K_p, K_i, K_d) for a PID controller depend on what that controller is driving. A self-tuning PID controller can identify the process dynamics using a short period of process behavior, such as a setpoint change under the closed loop control condition, and then tuning the PID parameters based on the identified ideal parameters for both setpoint tracking and disturbance regulation characteristics; this makes it easy to set up the initial PID parameters and adapt the PID values to changes in the process dynamics (Takatsu et al. 1991). Self-tuning is critical to the automated design and implementation of PID controllers. ML techniques are increasingly studied and implemented in this field because of their pure data-driven and non-domain-knowledge-required characteristics. The ML-based PID controllers discussed in this section primarily use ML-aided tuning techniques. Zhang et al. (2011a) reviewed and researched PID control based on a back-propagation neural network (BPNN). In their paper, they showed that a NN with the ability to capture arbitrary nonlinear expressions could find the best parameters for PID control by studying system performance. The K_p, K_i, and K_d parameters learned by BPNN are then used in a self-learning PID controller. The simulation results show that the system has good static and dynamic performance.

Table 5 lists the papers that studied ML-based PID control in buildings. Most papers focus on HVAC control. Lighting control, elevator control, and fan control are also studied. In terms of ML algorithms, NNs are the most widely used algorithm for ML-based PID control in buildings. Bayesian networks (Fiducioso et al. 2019) are the only other algorithm used in these studies. Unlike RL, which is most often a black-box approach that depends only on data, ML-based PID control is a gray-box approach with a pre-determined model structure with parameters K_p, K_i, and K_d. It still depends on data, but because the model has a defined structure, it does not have to learn from scratch in the way that RL often does. The advantage of ML-based PID control over RL is its lower computation cost/time and robustness due to the pre-determined model structure. Given the more sophisticated and complicated tuning tasks in PID controllers and the requirement for robustness in

Table 5 Summary of ML-based PID control studies

No.	Reference	System type	System details	ML algorithm	Highlighted features
1	Ai et al. 2010	HVAC	Solar heating system	NN	NA
2	Attaran et al. 2016	HVAC	Humidifier and heating coil control	RBF NN combined with the epsilon constraint	Better than traditional PID
3	Fiducioso et al. 2019	HVAC	Room temperature control	Contextual Bayesian Optimization	NA
4	Song 2014	HVAC	Indoor environment	Fuzzy logic control and NN technology	Combine fuzzy and NN
5	Ursu et al. 2013	HVAC	Whole HVAC system	Fuzzy supervised neuro-control	NA
6	Zhang et al. 2011b	HVAC	Whole HVAC system	BPNN	NA
7	Sharifian et al. 2011	Building elevator	Elevator	RBF NN	Applied in elevator
8	Ghadi et al. 2016	HVAC systems and light controllers	HVAC systems and light controllers	A combination of fuzzy logic, neural controller, and PID controller	NA
9	Lee and Chen 2015	HVAC	Server fan cooling system	Combining a PID NN with fan-power-based optimization	NA
10	Dehghani and Khodadadi 2017	HVAC	Heating system	Neuro-fuzzy PID controller based on smith predictor	NA

* NA: not available

building control, ML-based PID control will be increasingly studied in the future.

4.2.3 *ML-based fuzzy control*

Fuzzy control is suitable for a system where a mathematical model is difficult to derive and dynamic characteristics are not easy to quantify (Chua et al. 2017). The goal of a fuzzy logic system (or model) is the acquisition of a knowledge (rule) base that represents the input–output function of the real system or problem that we want to model (Hussain et al. 2014). However, fuzzy-based inference mechanisms have their own limitations. As the problem complexity grows (due to the system complexity, a large amount of disparate sensor data, the number of potential faults, etc.), the number of fuzzy sets and fuzzy rules required to analyze the system performance also grows (Najafi et al. 2012). Added to this is the difficulty in adjusting and tuning fuzzy sets manually or through other approaches (Najafi et al. 2012). The objective of the learning process is to create and then fine tune the fuzzy sets and rules so as to meet user specified performance criteria for the system (Hussain et al. 2014). This process is remarkably similar to the learning process in ML, so many ML techniques are also applied in fuzzy control.

The synergy of NN technology with fuzzy logic, and the algorithms used in computational intelligence, are the basic concepts behind most ML-based fuzzy control in smart buildings (Ghadi et al. 2014). The synergistic neuro-fuzzy technique is a hybrid system consisting of NNs and fuzzy logic technology. The controller using this synergy is called a neuro-fuzzy controller, which has been widely studied in building controls. The typical diagram of a neuro-fuzzy controller is illustrated in Figure 11 of Ursu et al. (2013). Kaur presented the neuro-fuzzy controller algorithm for an air conditioning system, which combines the learning capabilities of NNs and control capabilities of fuzzy logic control. The neuro-fuzzy controller for an air conditioning system takes two inputs (temperature and humidity) and controls the compressor speed. Wang (2013) presented a fuzzy NN that has the advantage of self-adapting, self-learning and tuning on-line. In simulation, the fuzzy NN system demonstrates stability, and uncertain factors have limited impact on stability, so it demonstrates good control of the steam pressure systems of a boiler.

Introducing and merging advanced optimization methods in ML to the field of fuzzy control is another research direction in ML-based fuzzy control. The most common optimization algorithm in fuzzy control is the genetic algorithm, but many optimization algorithms that are ML-based or often used with ML algorithms are also applied in fuzzy control. Adaptive neuro-fuzzy inference system, ANFIS, is a specific ML algorithm that combines ML optimization algorithms with fuzzy logic. It has been widely used to predict and control building energy systems. Ardabili et al. (2020) proposed two hybrid models for HVAC control: ANFIS-particle swarm optimization and ANFIS-genetic algorithm. ANFIS controllers are also used to control hydronic sub-systems such as heating, solar energy, and other renewable energy sources. In some research, people manually combined NN with fuzzy control by using NN to predict variables first and then feeding the NN results to fuzzy control. Collotta et al. applied NN to predict indoor temperatures and then applied the predicted values to a fuzzy logic control unit for on/off switching of the HVAC system. Papantoniou et al. applied NN to predict outdoor and indoor air temperatures that are then used for real-time HVAC setpoint control using fuzzy techniques. Table 6 lists the papers related to ML-based fuzzy control.

Table 6 Summary of ML-based fuzzy control studies

No.	Reference	Building type	System type	System details	ML algorithm	Highlighted features
1	Kaur and Kaur 2012	NA*	HVAC	Control compressor speed	Neuro fuzzy controller	NA
2	Kaur and Kaur 2012; Ursu et al. 2013	Commercial building	HVAC	Flow rate of air and chilled/ heated water in the coil	Fuzzy supervised neural control	Consider both energy and thermal comfort
3	Collotta et al. 2014; Esen et al. 2008	NA	HVAC	On/off switching of the HVAC system	NN predicts indoor temperatures that are used for a fuzzy logic control unit	NA
4	Ardabili et al. 2016	NA	HVAC	Control temperature and relative humidity values	Fuzzy and predictive (RBF) controllers	Simulations
5	Papantoniou et al. 2015	Commercial building	HVAC	Control temperature setpoint	NN predicts outdoor and indoor air temperature that are used for real time control using fuzzy techniques	NA
6	Ardabili et al. 2020	Industrial building	HVAC	Predictions for future controls	Adaptive neuro-fuzzy inference system-particle swarm optimization and adaptive neuro-fuzzy inference system-genetic algorithm	NA
7	Yu and Dexter 2010	NA	HVAC	Room temperature setpoint and tank water temperature setpoint	Fuzzy controller tuned by RL	Simulations
8	Du and Li 2010	NA	HVAC	Water flow rate	Self-learning fuzzy control method based on RBF NNs	NA
9	Wang 2013	NA	Boiler	Steam pressure control	Fuzzy NN	NA
10	Ali et al. 2014	Commercial building	HVAC	Supply air pressure control	Adaptive neuro-fuzzy controllers	Simulations

* NA: not available

4.2.4 Challenges and opportunities of RL

Applying RL in real-world building control problems is still challenging. First, deep RL approaches in general are data intensive. Though this is reasonable and necessary due to the requirements of adequate exploration and exploitation to learn the optimal control law, it poses a challenge in real-world applications. To address this, the following approaches are currently under investigation: (1) develop data-driven or physics-informed building model learning approaches that can automatically learn the thermal dynamics from building operation data and deliver that model in the form of a building simulator used to train the RL agent; (2) use transfer learning (Pinto et al. 2022b) to jumpstart the RL controller training from one of the following four sources: (a) an existing RL controller from a similar building, see Zhang et al. (2020c) as an example, (b) a building controller learned via imitation learning, which imitates the existing building controller, (c) a building controller learned via offline RL (i.e., train RL policy from a static dataset without exploration, see examples in Pinto et al. (2022a) and Fujimoto et al. (2019)), which not only imitates but also improves the existing building controller, and (d) a building controller learned from other operation scenarios and weather conditions, see Lissa et al. (2020) as an example. Another drawback of RL when compared with MPC is that it cannot handle some types of constraints directly (e.g., state related constraints). One common workaround in RL is like the penalty method used in constrained optimization, where a constraints violation penalty is included in the reward function. This, though in general an effective solution, does not provide theoretical guarantee that the constraints will not be violated and determining the penalty coefficient is tricky. Safe RL approaches, e.g., Paternain et al. (2019), can help address this challenge by solving a constrained Markov decision process to deliver a control policy that guarantees required safety levels (in the form of chance constraints). Finally, RL training can be sensitive to hyper-parameters and there are usually multiple hyper-parameters for an RL algorithm, related to general learning and the specific configuration for the algorithm. Though there are general rules of thumb, choosing hyper-parameters that deliver the best performance is still challenging, especially for different control problems (buildings with different floor plan, control variables and constraints) in which the optimal combination of hyper-parameters may differ. Therefore, future studies are required to investigate and find approaches that aid real-world practitioners in efficiently determining the best performing hyper-parameters when training RL controllers for new buildings.

5 Conclusions

5.1 Future trends and challenges

From our point of view, future studies on the application of ML for building control will move from simply applying specific ML algorithms to developing a more comprehensive and generic workflow that leads to ML controllers for real buildings. To be more specific, we predict the following trends: (1) the focus will move from ML algorithm to sensor/data related studies, including metadata schema (or semantic data) and smart sensors in MLC; (2) there will be increased exploration of novel methodologies to consider and combine domain knowledge into the MLC development; and (3) more research will focus on practical topics such as data engineering and computation cost for ML controller development in real buildings.

Real building applications, reduction of engineering cost, and automation are the three major challenges of building MLC studies. Specifically, the challenges include: (1) engineering cost reduction and automation are the major advantages of MLC, but to embed MLC algorithms into building automation systems is challenging—more advanced hardware is required to meet the computation requirements; (2) the automation mentioned in (1) can be realized by making the developed MLC algorithms compatible with various energy system types, building types, and weather conditions; it is challenging to develop a generic and extendible framework with automated tuning and modeling workflow; and (3) it is challenging to test the algorithm extendibility on multiple real-building testbeds because testbeds with fault-free sensors and high-quality data are rare.

5.2 Potential directions of future research

Although MLC in building operation is carefully reviewed in this paper, some important research topics are barely covered, which leaves many research and industry opportunities in the future. The potential directions of future research are identified by: (1) insights and conclusions of the literature reviewed in each sub-section, (2) the authors' research and industry experience, and (3) the emerging research trends and topics in the field of building energy efficiency.

From the research perspective:

- **Smart and connected communities.** The research topic of smart and connected communities and cities is increasingly studied and is a big trend in building energy studies. MLC can play a key role in the community or urban scale for multiple buildings and t`he interactions among building and energy systems. However, the implementation and study of MLC in the larger scale needs more attention and has great research potential in the future.
- **Grid-interactive efficient buildings.** The research topics of building demand flexibility and grid-interactive efficient buildings are also increasingly studied. More review and technical studies are expected to apply MLC in understanding and coordinating the operation of buildings and the grid.
- **Fault detection and diagnostics.** Equipment faults and control errors are pervasive in today's commercial buildings (Zhang et al. 2020a). It is challenging for building controls to maintain performance in the presence of building faults. It is even more challenging for MLC, where the uncertainty of the impact of faults on building control is even harder to quantify because of the data-driven characteristics of MLC (Bae et al. 2021; Zhang et al. 2021a). As a result, how to make MLC work better with building faults have great research potential.
- **Performance evaluation**. Almost all papers claimed improved performance by using their proposed MLC, but the conclusions are from different building, different energy system, different MLC algorithms, different test scenarios, and different performance metrics. If the MLC performs good in one building, will it also be adaptive to other buildings? Most papers did not answer this question. High-quality test data and testbeds are greatly needed for the development of algorithms in this field, which is key to quantify and evaluate MLC performance. Although we already have standard virtual control testbeds such as BOPTEST (Blum et al. 2019), more real building testbeds, which can be accessed remotely (e.g., through cloud services), with high-quality labeled test data are needed for a better development and performance evaluation of MLC studies.
- **Selection of the most suitable MLC**. It is hard to conclude which MLC technique is suitable for a particular application. However, future studies should guide people to select MLC for different applications by comparing different MLC techniques for test problems and then recommending best suited techniques for the particular problem. Many researchers combine different ML techniques (hybrid) to overcome their deficiencies. No single MLC has all the desirable features, so most of the methods can complement one another resulting in better control. In most cases, hybrid techniques and meta modeling provided better results and should obtain more attention from researchers.

From the industry perspective:

- **Robustness and reliability**. Most of the studies were validated by simulations or on small-scale HVAC systems. Practical validation needs to be performed on commercial HVAC systems, controlling them on a real-time basis. The simulation performance might not be reliable and robust for the stated goals, as the method might work for some buildings but could not be generalized to others. Methods to evaluate these techniques in terms of not only accuracy but also robustness to real-world uncertainty are needed. For example, the performance of MLC under significant environmental changes (such as climate change and natural disasters) or changes in thermal comfort requirements is a good research topic. In addition, cybersecurity studies in building control are also worth investigating to improve robustness and reliability.
- **Adaptability and transferability**. The lack of model adaptability and transferability limits a model trained with one data-rich building to be used in another building with limited data. Scalable MLC is one of the most important research directions from the industry perspective. For example, MPC is very mature and well-studied in academia, but it is barely applied in the building industry and rule-based control is still more prevalent. The use of ML to further empower the automation of developing MLC in buildings is a promising research direction in the future. Metadata-schema-based MLC is one research direction that would improve adaptability and transferability.
- **Technoeconomic analysis of MLC deployment in buildings**. The cost of applying MLC includes development, operation, and maintenance costs. The authors estimate that a large proportion of cost comes from operation because MLC requires local or cloud computing infrastructure with high computational capability to support ML or MLC. This is either a one-time investment (for local computing) or continuous cost (for cloud computing). The cloud computing controls and optimizes building operation by using a shared and dynamic infrastructure which relies on the advancements in internet connection speed and computational capability. The cost and benefits of MLC are barely studied in existing literature, but this research is essential to MLC application in industry.
- **Accelerating MLC training and inference.** The speeds for MLC training and inference are both essential. How fast MLC can be trained or developed determines the feasibility of MLC. The speed for training is especially important for RL, as discussed in Section 4.2.4. How fast MLC can infer will determine whether MLC can run in real-time. To realize that, we should either improve the computational capability (e.g., by using high performance computing), or reduce the computational burden (e.g., by using computationally efficient models and surrogate models). Both the training and inference speed for MLC are worth investigating.
- **Model interpretability.** It is a challenge for modelers and users of MLC to fully understand the inference mechanism learned, thus jeopardizing trust in the MLC decision. The models developed are typically of low interpretability, which is the nature of black-box models. Interpretable machine learning is an emerging subject in the field of big data analytics, which aims to provide methods and tools to enhance model interpretability without sacrificing model complexity (Doshi-Velez and Kim 2017). However, thus far this topic has only been studied in ML energy modeling (Fan et al. 2019b; Manfren et al. 2022), thermal comfort modeling (Zhang et al. 2021c), and fault detection (Madhikermi et al. 2019; Li et al. 2021). Interpretable MLC is worth investigating in the future to improve the credibility of MLC and persuade the industry to widely apply MLC.
- **Identify suitable applications.** MLC is not going to resolve all the challenges of building controls. Identifying the control applications that are suitable or worth consideration for applying MLC is critical to motivating building control vendors and the building industry to widely apply MLC. MLC is originally motivated by problems involving complex control tasks where it may be difficult or impossible to model the system and develop a useful control law. Instead, people leverage experience and data to learn effective controllers. To put it in another way, complex control tasks with complex development of models and control laws, but with abundant data, are the places where industry should apply MLC. As we mentioned earlier, smart and connected communities, grid-interactive efficient buildings, fault detection and diagnostics, would benefit from MLC. In addition, MPC would be useful in buildings where the modeling of components is very complex, time-consuming, or unscalable. However, more specific applications need to be identified to motivate control vendors and the building industry to widely apply MLC.

5.3 Summary

This paper systematically reviewed MLC in building energy systems. We utilized the SSSS module to exhaust relevant literature. We summarized the existing review papers on MLC and identified that most review papers do not include papers with the exact scope of MLC in building energy systems; a review that specifically and systematically focuses on MLC in buildings energy systems is missing. Then, technical papers were reviewed in terms of two major categories of applications of ML in the control process: building system and component modeling for control, and learning the control process. The key conclusions from the review of technical papers are:

(1) Building system and component modeling for control is the most widely studied ML topic in building control. However, there is less research into further utilization and integration of the ML models to control.

(2) RL is one of the most recent research topics in the field of building control. RL papers discuss diversified building system types, levels of control, ML algorithms, data sources, sample time/control interval/control horizon, action space and reward structure.

(3) ML-powered traditional control (e.g., ML-based self-tuning for PID and ML-based fuzzy control) is also a promising research direction which tends to automatically improve traditional controls.

Based on the results of the review, we identified gaps in the existing research and promising directions for future research. We predict that future efforts will focus on overcoming the challenges of moving MLC to real buildings on a more comprehensive scale. Some key steps include automating processes, reducing engineering costs, acquiring high quality datasets, minimizing computational costs, and proving and improving the reliability of the approaches. Research will also be needed in improving the integration of metadata schema and smart sensors with ML algorithms.

Acknowledgements

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Building Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Declaration of interests

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

Author contributions statement

All authors contributed to the study conception and design. Liang Zhang: Conceptualization; Formal analysis; Investigation; Methodology; Writing—original draft, review & editing. Zhelun Chen: Conceptualization; Formal analysis; Investigation; Visualization; Writing—original draft, review & editing. Xiangyu Zhang: Investigation; Writing—original draft, review & editing. Amanda Pertzborn: Writing—review & editing. Xin Jin: Writing—review & editing. All authors read and approved the final manuscript."

References

Afroz Z, Shafiullah G, Urmee T, et al. (2018). Modeling techniques used in building HVAC control systems: A review. *Renewable and Sustainable Energy Reviews*, 83: 64–84.

Ahmad MW, Mourshed M, Yuce B, et al. (2016). Computational intelligence techniques for HVAC systems: A review.*Building Simulation*, 9: 359–398.

Ahmad MW, Hippolyte J-L, Mourshed M, et al. (2017a). Random forests and artificial neural network for predicting daylight illuminance and energy consumption. In: Proceedings of the 15th International IPBSA Building Simulation Conference, San Francisco, USA.

Ahmad MW, Mourshed M, Rezgui Y (2017b). Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy and Buildings*, 147: 77–89.

Ai T, Yu J-q, Liu Y-f, et al. (2010). Study on neural network self-tuning PID control for temperature of active solar house heating system. In: Proceedings of 2010 2nd International Workshop on Intelligent Systems and Applications, Wuhan, China.

Alawadi S, Mera D, Fernández-Delgado M, et al. (2022). A comparison of machine learning algorithms for forecasting indoor temperature in smart buildings. *Energy Systems*, 13: 689–705.

Ali MM, Khan MA, Shafi M, et al. (2014). Heat ventilation & air-conditioning system with self-tuning fuzzy PI controller. *International Journal of Modern Engineering Research (IJMER)*, 4(9): 24–35.

Ardabili SF, Mahmoudi A, Gundoshmian TM (2016). Modeling and simulation controlling system of HVAC using fuzzy and predictive (radial basis function, RBF) controllers. *Journal of Building Engineering*, 6: 301–308.

Ardabili S, Beszedes B, Nadai L, et al. (2020). Comparative analysis of single and hybrid neuro-fuzzy-based models for an industrial heating ventilation and air conditioning control system. In: Proceedings of International Conference on Computing and Communication Technologies (2010 RIVF), Ho Chi Minh City, Vietnam.

Atabay D, Herzog S, Sanger F, et al. (2013). Self-adapting building models and optimized hvac scheduling for demand side management. In: Proceedings of the 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013).

Attaran SM, Yusof R, Selamat H (2016). A novel optimization algorithm based on epsilon constraint-RBF neural network for tuning PID controller in decoupled HVAC system. *Applied Thermal Engineering*, 99: 613–624.

Ayoub M (2020). A review on machine learning algorithms to predict daylighting inside buildings. *Solar Energy*, 202: 249–275.

Bae Y, Bhattacharya S, Cui B, et al. (2021). Sensor impacts on building and HVAC controls: A critical review for building energy performance. *Advances in Applied Energy*, 4: 100068.

Barrett E, Linder S (2015). Autonomous HVAC control: A reinforcement learning approach. In: Proceedings of Joint European Conference on Machine Learning and Knowledge Discovery in Databases.

Beccali M, Bonomolo M, Ciulla G, et al. (2018). Assessment of indoor illuminance and study on best photosensors' position for design and commissioning of Daylight Linked Control systems. A new method based on artificial neural networks. *Energy*, 154: 466–476.

Blum D, Jorissen F, Huang S, et al. (2019). Prototyping the BOPTEST framework for simulation-based testing of advanced control strategies in buildings. Lawrence Berkeley National Laboratory.

Chaudhuri T, Soh YC, Li H, et al. (2017). Machine learning based prediction of thermal comfort in buildings of equatorial Singapore. In: Proceedings of 2017 IEEE International Conference on Smart Grid and Smart Cities (ICSGSC), Singapore.

Chen Z, Zhang Y, Zhang T (2015). An intelligent control approach to home energy management under forecast uncertainties. In: Proceedings of 2015 IEEE 5th International Conference on Power Engineering, Energy and Electrical Drives (POWERENG), Riga, Latvia.

Chen Y, Norford LK, Samuelson HW, et al. (2018). Optimal control of HVAC and window systems for natural ventilation through reinforcement learning. *Energy and Buildings*, 169: 195–205.

Chen Z, Jiang C (2018). Building occupancy modeling using generative adversarial network. *Energy and Buildings*, 174: 372–379.

Chen B, Cai Z, Bergés M (2019). Gnu-RL: A precocial reinforcement learning solution for building HVAC control using a differentiable MPC policy. In: Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, New York, USA.

Cheng M-Y, Cao M-T (2014). Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. *Applied Soft Computing*, 22: 178–188.

Chow TT, Zhang GQ, Lin Z, et al. (2002). Global optimization of absorption chiller system by genetic algorithm and neural network. *Energy and Buildings*, 34: 103–109.

Chua KH, Lim YS, Morris S (2017). A novel fuzzy control algorithm for reducing the peak demands using energy storage system. *Energy*, 122: 265–273.

Cifuentes J, Marulanda G, Bello A, et al. (2020). Air temperature forecasting using machine learning techniques: a review. *Energies*, 13: 4215.

Cole WJ, Powell KM, Hale ET, et al. (2014). Reduced-order residential home modeling for model predictive control. *Energy and Buildings*, 74: 69–77.

Collotta M, Messineo A, Nicolosi G, et al. (2014). A dynamic fuzzy controller to meet thermal comfort by using neural network forecasted parameters as the input. *Energies*, 7: 4727–4756.

Costa A, Keane MM, Torrens JI, et al. (2013). Building operation and energy performance: Monitoring, analysis and optimisation toolkit. *Applied Energy*, 101: 310–316.

Costanzo GT, Iacovella S, Ruelens F, et al. (2016). Experimental analysis of data-driven control for a building heating system. *Sustainable Energy, Grids and Networks*, 6: 81–90.

Cox SJ, Kim D, Cho H, et al. (2019). Real time optimal control of district cooling system with thermal energy storage using neural networks. *Applied Energy*, 238: 466–480.

Cui C, Wu T, Hu M, et al. (2016). Short-term building energy model recommendation system: A meta-learning approach. *Applied Energy*, 172: 251–263.

Dai X, Liu J, Zhang X (2020). A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings. *Energy and Buildings*, 223: 110159.

Dehghani A, Khodadadi H (2017). Designing a neuro-fuzzy PID controller based on Smith predictor for heating system. In: Proceedings of 2017 17th International Conference on Control, Automation and Systems (ICCAS), Jeju, R.O. Korea.

Ding X, Du W, Cerpa A (2019). OCTOPUS: Deep reinforcement learning for holistic smart building control. In: Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, New York, USA.

Dong B, Lam KP (2014). A real-time model predictive control for building heating and cooling systems based on the occupancy behavior pattern detection and local weather forecasting. *Building Simulation*, 7: 89–106.

Doshi-Velez F, Kim B (2017). Towards a rigorous science of interpretable machine learning. arXiv: 1702.08608.

Du D, Li X (2010). A self-learning fuzzy control method based on RBF neural networks. In: Proceedings of 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), Taiyuan, China.

Esen H, Inalli M, Sengur A, et al. (2008). Artificial neural networks and adaptive neuro-fuzzy assessments for ground-coupled heat pump system. *Energy and Buildings*, 40: 1074–1083.

Fan C, Wang J, Gang W, et al. (2019a). Assessment of deep recurrent neural network-based strategies for short-term building energy predictions. *Applied Energy*, 236: 700–710.

Fan C, Xiao F, Yan C, et al. (2019b). A novel methodology to explain and evaluate data-driven building energy performance models based on interpretable machine learning. *Applied Energy*, 235: 1551–1560.

Fan C, Xiao F, Zhao Y (2017). A short-term building cooling load prediction method using deep learning algorithms. *Applied Energy*, 195: 222–233.

Fazenda P, Veeramachaneni K, Lima P, et al. (2014). Using reinforcement learning to optimize occupant comfort and energy usage in HVAC systems. *Journal of Ambient Intelligence and Smart Environments*, 6: 675–690.

Fiducioso M, Curi S, Schumacher B, et al. (2019). Safe contextual Bayesian optimization for sustainable room temperature PID control tuning. arXiv: 1906.12086.

Finck C, Li R, Zeiler W (2019). Economic model predictive control for demand flexibility of a residential building. *Energy*, 176: 365–379.

Florita AR, Henze GP (2009). Comparison of short-term weather forecasting models for model predictive control. *HVAC&R Research*, 15: 835–853.

Fujimoto S, Meger D, Precup D (2019). Off-policy deep reinforcement learning without exploration. In: Proceedings of the 36th International Conference on Machine Learning.

Gao G, Li J, Wen Y (2019). Energy-efficient thermal comfort control in smart buildings via deep reinforcement learning. arXiv: 1901.04693.

Garnier A, Eynard J, Caussanel M, et al. (2014). Low computational cost technique for predictive management of thermal comfort in non-residential buildings. *Journal of Process Control*, 24: 750–762.

Ghadi YY, Rasul MMG, Khan MKK (2014). Recent developments of advanced fuzzy logic controllers used in smart buildings in subtropical climate. *Energy Procedia*, 61: 1021–1024.

Ghadi YY, Rasul MG, Khan MMK (2016). Design and development of advanced fuzzy logic controllers in smart buildings for institutional buildings in subtropical Queensland. *Renewable and Sustainable Energy Reviews*, 54: 738–744.

Guo Y, Wang J, Chen H, et al. (2018). Machine learning-based thermal response time ahead energy demand prediction for building heating systems. *Applied Energy*, 221: 16–27.

Hagerman J (2014). Buildings-to-grid technical opportunities: introduction and vision. EERE Publication and Product Library.

Han M, May R, Zhang X, et al. (2019). A review of reinforcement learning methodologies for controlling occupant comfort in buildings. *Sustainable Cities and Society*, 51: 101748.

Han M, May R, Zhang X, et al. (2020). A novel reinforcement learning method for improving occupant comfort via window opening and closing. *Sustainable Cities and Society*, 61: 102247.

Hidalgo-León R, Litardo J, Urquizo J, et al. (2019). Some factors involved in the improvement of building energy consumption: A brief review. In: Proceedings of 2019 IEEE Fourth Ecuador Technical Chapters Meeting (ETCM), Guayaquil, Ecuador.

Hong T, Wang Z, Luo X, et al. (2020). State-of-the-art on research and applications of machine learning in the building life cycle. *Energy and Buildings*, 212: 109831.

Hosoz M, Ertunç HM, Bulgurcu H (2007). Performance prediction of a cooling tower using artificial neural network. *Energy Conversion and Management*, 48: 1349–1359.

Hu J, Olbina S (2011). Illuminance-based slat angle selection model for automated control of split blinds. *Building and Environment*, 46: 786–796.

Huang H, Chen L, Hu E (2015a). A neural network-based multi-zone modelling approach for predictive control system design in commercial buildings. *Energy and Buildings*, 97: 86–97.

Huang H, Chen L, Hu E (2015b). A new model predictive control scheme for energy and cost savings in commercial buildings: An airport terminal building case study. *Building and Environment*, 89: 203–216.

Hussain S, Gabbar HA, Bondarenko D, et al. (2014). Comfort-based fuzzy control optimization for energy conservation in HVAC systems. *Control Engineering Practice*, 32: 172–182.

Idowu S, Saguna S, Åhlund C, et al. (2016). Applied machine learning: forecasting heat load in district heating system. *Energy and Buildings*, 133: 478–488.

Jain A, Smarra F, Mangharam R (2017). Data predictive control using regression trees and ensemble learning. In: Proceedings of 2017 IEEE 56th Annual Conference on Decision and Control (CDC), Melbourne, Australia.

Javed A, Larijani H, Ahmadinia A, et al. (2014). Comparison of the robustness of RNN, MPC and ANN controller for residential heating system. In: Proceedings of 2014 IEEE Fourth International Conference on Big Data and Cloud Computing, Sydney, Australia.

Jin Y, Yan D, Kang X, et al. (2021). Forecasting building occupancy: A temporal-sequential analysis and machine learning integrated approach. *Energy and Buildings*, 252: 111362.

Jordan MI, Mitchell TM (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349: 255–260.

Kamel E, Sheikh S, Huang X (2020). Data-driven predictive models for residential building energy use based on the segregation of heating and cooling days. *Energy*, 206: 118045.

Karunamurthy K, Chatterjee D, Pattanashetti S, et al. (2020). Prediction of thermal performance of cooling tower of a chiller plant using machine learning. *IOP Conference Series*: *Earth and Environmental Science*, 573: 012029.

Kaur A, Kaur A (2012). Development of neuro fuzzy controller algorithm for air conditioning system. *International Journal of Engineering Science and Technology (IJEST)*, 4: 1667–1671.

Kazmi H, Mehmood F, Lodeweyckx S, et al. (2018). Gigawatt-hour scale savings on a budget of zero: Deep reinforcement learning based optimal control of hot water systems. *Energy*, 144: 159–168.

Kim Y, Park C (2014). Nonlinear predictive control of chiller system using gaussian process model. In: Proceedings of the 2nd Asia Conference of International Building Performance Simulation Association, Nagoya, Japan.

Kim W, Jeon Y, Kim Y (2016). Simulation-based optimization of an integrated daylighting and HVAC system using the design of experiments method. *Applied Energy*, 162: 666–674.

Kim J-H, Seong N-C, Choi W (2019). Modeling and optimizing a chiller system using a machine learning algorithm. *Energies*, 12: 2860.

Kumar R, Aggarwal RK, Sharma JD (2013). Energy analysis of a building using artificial neural network: A review. *Energy and Buildings*, 65: 352–358.

Kumar R, Khatri SK, Diván MJ (2021). Data center air handling unit fan speed optimization using machine learning techniques. In: Proceedings of 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India.

Kwok SSK, Lee EWM (2011). A study of the importance of occupancy to building cooling load in prediction by intelligent approach. *Energy Conversion and Management*, 52: 2555–2564.

Lazos D, Sproul AB, Kay M (2014). Optimisation of energy management in commercial buildings with weather forecasting inputs: A review. *Renewable and Sustainable Energy Reviews*, 39: 587–603.

Lee C, Chen R (2015). Optimal self-tuning PID controller based on low power consumption for a server fan cooling system. *Sensors*, 15: 11685–11700.

Lee YM, Horesh R, Liberti L (2015). Optimal HVAC control as demand response with on-site energy storage and generation system. *Energy Procedia*, 78: 2106–2111.

Li S, Ren S, Wang X (2013). HVAC room temperature prediction control based on neural network model. In: Proceedings of 2013 Fifth International Conference on Measuring Technology and Mechatronics Automation, Hong Kong, China.

Li B, Xia L (2015). A multi-grid reinforcement learning method for energy conservation and comfort of HVAC in buildings. In: Proceedings of 2015 IEEE International Conference on Automation Science and Engineering (CASE), Gothenburg, Sweden.

Li Z, Dong B (2018). Short term predictions of occupancy in commercial buildings—Performance analysis for stochastic models and machine learning approaches. *Energy and Buildings*, 158: 268–281.

Li G, Yao Q, Fan C, et al. (2021). An explainable one-dimensional convolutional neural networks based fault diagnosis method for building heating, ventilation and air conditioning systems. *Building and Environment*, 203: 108057.

Liang W, Quinte R, Jia X, et al. (2015). MPC control for improving energy efficiency of a building air handler for multi-zone VAVs. *Building and Environment*, 92: 256–268.

Lissa P, Schukat M, Barrett E (2020). Transfer learning applied to reinforcement learning-based HVAC control. *SN Computer Science*, 1: 127.

Liu S, Henze GP (2006). Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory. Part 2: Results and analysis. *Energy and Buildings*, 38: 148–161.

Liu H, Lee S, Kim M, et al. (2013). Multi-objective optimization of indoor air quality control and energy consumption minimization in a subway ventilation system. *Energy and Buildings*, 66: 553–561.

Lusis P, Khalilpour KR, Andrew L, et al. (2017). Short-term residential load forecasting: Impact of calendar effects and forecast granularity. *Applied Energy*, 205: 654–669.

Ma J, Qin J, Salsbury T, et al. (2012). Demand reduction in building energy systems based on economic model predictive control. *Chemical Engineering Science*, 67: 92–100.

Ma N, Aviv D, Guo H, et al. (2021). Measuring the right factors: A review of variables and models for thermal comfort and indoor air quality. *Renewable and Sustainable Energy Reviews*, 135: 110436.

Madhikermi M, Malhi AK, Främling K (2019). Explainable artificial intelligence based heat recycler fault detection in air handling unit. In: Proceedings of International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems.

Manfren M, James PA, Tronchin L (2022). Data-driven building energy modelling–An analysis of the potential for generalisation through interpretable machine learning. *Renewable and Sustainable Energy Reviews*, 167: 112686.

Manjarres D, Mera A, Perea E, et al. (2017). An energy-efficient predictive control for HVAC systems applied to tertiary buildings based on regression techniques. *Energy and Buildings*, 152: 409–417.

Mason K, Grijalva S (2019). A review of reinforcement learning for autonomous building energy management. *Computers & Electrical Engineering*, 78: 300–312.

Mayne DQ (2014). Model predictive control: Recent developments and future promise. *Automatica*, 50: 2967–2986.

Mba L, Meukam P, Kemajou A (2016). Application of artificial neural network for predicting hourly indoor air temperature and relative humidity in modern building in humid region. *Energy and Buildings*, 121: 32–42.

Merabti S, Draoui B, Bounaama F (2016). A review of control systems for energy and comfort management in buildings. In: Proceedings of 2016 8th International Conference on Modelling, Identification and Control (ICMIC).

Mirinejad H, Welch KC, Spicer L (2012). A review of intelligent control techniques in HVAC systems. In: Proceedings of 2012 IEEE Energytech, Cleveland, OH, USA.

Mnih V, Badia AP, Mirza M, et al. (2016). Asynchronous methods for deep reinforcement learning. In: Proceedings of International Conference on Machine Learning.

Monfet D, Corsi M, Choinière D, et al. (2014). Development of an energy prediction tool for commercial buildings using case-based reasoning. *Energy and Buildings*, 81: 152–160.

Mosaico G, Saviozzi M, Silvestro F, et al. (2019). Simplified state space building energy model and transfer learning based occupancy estimation for HVAC optimal control. In: Proceeding of 2019 IEEE 5th International forum on Research and Technology for Society and Industry (RTSI), Florence, Italy.

Mosavi A, Salimi M, Faizollahzadeh Ardabili S, et al. (2019). State of the art of machine learning models in energy systems, a systematic review. *Energies*, 12: 1301.

Mustafaraj G, Lowry G, Chen J (2011). Prediction of room temperature and relative humidity by autoregressive linear and nonlinear neural network models for an open office. *Energy and Buildings*, 43: 1452–1460.

Nagy A, Kazmi H, Cheaib F, et al. (2018). Deep reinforcement learning for optimal control of space heating. arXiv: 1805.03777.

Naidu DS, Rieger CG (2011). Advanced control strategies for HVAC&R systems—An overview: part II: soft and fusion control. *HVAC&R Research*, 17: 144–158.

Najafi M, Auslander DM, Bartlett PL, et al. (2012). Application of machine learning in the fault diagnostics of air handling units. *Applied Energy*, 96: 347–358.

O'Neill Z, O'Neill C (2016). Development of a probabilistic graphical model for predicting building energy performance. *Applied Energy*, 164: 650–658.

Oldewurtel F, Ulbig A, Parisio A, et al. (2010). Reducing peak electricity demand in building climate control using real-time pricing and model predictive control. In: Proceedings of 49th IEEE Conference on Decision and Control (CDC), Atlanta, GA, USA.

Oldewurtel F, Parisio A, Jones CN, et al. (2012). Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, 45: 15–27.

Overgaard A, Nielsen BK, Kallesøe CS, et al. (2019). Reinforcement learning for mixing loop control with flow variable eligibility trace. In: Proceedings of 2019 IEEE Conference on Control Technology and Applications (CCTA), Hong Kong, China.

Papantoniou S, Kolokotsa D, Kalaitzakis K (2015). Building optimization and control algorithms implemented in existing BEMS using a web based energy management and control system. *Energy and Buildings*, 98: 45–55.

Parisio A, Rikos E, Glielmo L (2014). A model predictive control approach to microgrid operation optimization. *IEEE Transactions on Control Systems Technology*, 22: 1813–1827.

Park JY, Dougherty T, Fritz H, et al. (2019a). LightLearn: an adaptive and occupant centered controller for lighting based on reinforcement learning. *Building and Environment*, 147: 397–414.

Park S, Ahn KU, Hwang A, et al. (2019b). Artificial Neural Network Models for Optimal Start and Stop of Chiller and AHU. *Journal of the Architectural Institute of Korea Structure and Construction*, 35: 45–52.

Park S, Ahn KU, Hwang S, et al. (2019c). Machine learning *vs*. hybrid machine learning model for optimal operation of a chiller. *Science and Technology for the Built Environment*, 25: 209–220.

Paternain S, Calvo-Fullana M, Chamon LFO, et al. (2019). Safe policies for reinforcement learning via primal-dual methods. arXiv: 1911.09101.

Peng KS, Morrison CT (2016). Model Predictive Prior Reinforcement Learning for a Heat Pump Thermostat. In: Proceedings of IEEE International Conference on Automatic Computing: Feedback Computing.

Pinto G, Messina R, Li H, et al. (2022a). Sharing is caring: an extensive analysis of parameter-based transfer learning for the prediction of building thermal dynamics. *Energy and Buildings*, 276: 112530.

Pinto G, Wang Z, Roy A, et al. (2022b). Transfer learning for smart buildings: A critical review of algorithms, applications, and future perspectives. *Advances in Applied Energy*, 5: 100084.

Potočnik P, Vidrih B, Kitanovski A, et al. (2019). Neural network, ARX, and extreme learning machine models for the short-term prediction of temperature in buildings.*Building Simulation*, 12: 1077–1093.

Raman NS, Devraj AM, Barooah P, et al. (2020). Reinforcement learning for control of building HVAC systems. 2020 American Control Conference (ACC), Denver, CO, USA.

Ruano A, Pesteh S, Silva S, et al. (2016). PVM-based intelligent predictive control of HVAC systems. *IFAC-PapersOnLine*, 49: 371–376.

Ruelens F, Claessens BJ, Vandael S, et al. (2017). Residential demand response of thermostatically controlled loads using batch reinforcement learning. *IEEE Transactions on Smart Grid*, 8: 2149–2159.

Ryu SH, Moon HJ (2016). Development of an occupancy prediction model using indoor environmental data based on machine learning techniques. *Building and Environment*, 107: 1–9.

Sala-Cardoso E, Delgado-Prieto M, Kampouropoulos K, et al. (2018). Activity-aware HVAC power demand forecasting. *Energy and Buildings*, 170: 15–24.

Salimi S, Hammad A (2019). Critical review and research roadmap of office building energy management based on occupancy monitoring. *Energy and Buildings*, 182: 214–241.

Schulman J, Wolski F, Dhariwal P, et al. (2017). Proximal policy optimization algorithms. arXiv: 1707.06347.

Seyedzadeh S, Rahimian FP, Glesk I, et al. (2018). Machine learning for estimation of building energy consumption and performance: a review. *Visualization in Engineering*, 6: 5.

Sharifian MBB, Mirlo A, Tavoosi J, Sabahi M (2011). Self-adaptive RBF neural network PID controller in linear elevator. In: Proceedings of 2011 International Conference on Electrical Machines and Systems, Beijing, China.

Smarra F, Jain A, de Rubeis T, et al. (2018). Data-driven model predictive control using random forests for building energy optimization and climate control. *Applied Energy*, 226: 1252–1272.

Song Y (2014). Intelligent PID controller bassed on fuzzy logic control and neural network technology for indoor environment quality improvement. PhD Thesis, University of Nottingham, UK.

Swider DJ, Browne MW, Bansal PK, Kecman V (2001). Modelling of vapour-compression liquid chillers with neural networks. *Applied Thermal Engineering*, 21: 311–329.

Takatsu H, Kawano T, Kitano K (1991). Intelligent self-tuning PID controller. *IFAC Proceedings Volumes*, 24: 11–15.

Tang F, Kusiak A, Wei X (2014). Modeling and short-term prediction of HVAC system with a clustering algorithm. *Energy and Buildings*, 82: 310–321.

Touzani S, Prakash AK, Wang Z, et al. (2021). Controlling distributed energy resources via deep reinforcement learning for load flexibility and energy efficiency. *Applied Energy*, 304: 117733.

Ursu I, Nastase I, Caluianu S, et al. (2013). Intelligent control of HVAC systems. Part I: Modeling and synthesis. *Incas Bulletin*, 5: 103–118.

Vázquez-Canteli JR, Nagy Z (2019). Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Applied Energy*, 235: 1072–1089.

Verrilli F, Srinivasan S, Gambino G, et al. (2017). Model predictive control-based optimal operations of district heating system with thermal energy storage and flexible loads. *IEEE Transactions on Automation Science and Engineering*, 14: 547–557.

Wagiman KR, Abdullah MN, Hassan MY, et al. (2020). Lighting system control techniques in commercial buildings: Current trends and future directions. *Journal of Building Engineering*, 31: 101342.

Wahid F, Kim D (2016). A prediction approach for demand analysis of energy consumption using K-nearest neighbor in residential buildings. *International Journal of Smart Home*, 10: 97–108.

Wang L (2013). The application of fuzzy neural network to boiler steam pressure control. *International Journal of Computer Science Issues (IJCSI)*, 10: 704–707.

Wang Z, Hong T, Piette MA (2019). Data fusion in predicting internal heat gains for office buildings through a deep learning approach. *Applied Energy*, 240: 386–398.

Wang Z, Hong T (2020). Reinforcement learning for building controls: The opportunities and challenges. *Applied Energy*, 269: 115036.

Wei X, Kusiak A, Li M, et al. (2015). Multi-objective optimization of the HVAC (heating, ventilation, and air conditioning) system performance. *Energy*, 83: 294–306.

Wei T, Wang Y, Zhu Q (2017). Deep reinforcement learning for building HVAC control. In: Proceedings of the 54th Annual Design Automation Conference 2017, Austin, TX, USA.

Wei W, Ramalho O, Malingre L, et al. (2019). Machine learning and statistical models for predicting indoor air quality. *Indoor Air*, 29: 704–726.

Wu Z, Li N, Peng J, et al. (2018). Using an ensemble machine learning methodology-Bagging to predict occupants' thermal comfort in buildings. *Energy and Buildings*, 173: 117–127.

Xu C, Chen H, Wang J, et al. (2019). Improving prediction performance for indoor temperature in public buildings based on a novel deep learning method. *Building and Environment*, 148: 128–135.

Yang Z, Li N, Becerik-Gerber B, et al. (2014). A systematic approach to occupancy modeling in ambient sensor-rich buildings. *SIMULATION*, 90: 960–977.

Yang L, Nagy Z, Goffin P, et al. (2015). Reinforcement learning for optimal control of low exergy buildings. *Applied Energy*, 156: 577–586.

Yang S, Wan MP, Chen W, et al. (2020). Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization. *Applied Energy*, 271: 115147.

Yang S, Wan MP, Chen W, et al. (2021). Experiment study of machine-learning-based approximate model predictive control for energy-efficient building control. *Applied Energy*, 288: 116648.

Yildiz B, Bilbao JI, Sproul AB (2017). A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*, 73: 1104–1122.

Yousaf A, Asif RM, Shakir M, et al. (2021). A novel machine learning-based price forecasting for energy management systems. *Sustainability*, 13: 12693.

Yu Z, Dexter A (2010). Online tuning of a supervisory fuzzy controller for low-energy building system using reinforcement learning. *Control Engineering Practice*, 18: 532–539.

Yu Z, Haghighat F, Fung BCM, et al. (2010). A decision tree method for building energy demand modeling. *Energy and Buildings*, 42: 1637–1646.

Zhang Y, Zhao D, Zhang J (2011a). Research on PID controller based on the BP neural network. In: Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology, Harbin, China.

Zhang Y, Zuo W-h, Lin T (2011b). Application of BP neural network PID control in air-conditioning system. *Computer Simulation*, 28(4): 149–151. (in Chinese)

Zhang L (2018). Data-driven whole building energy forecasting model for data predictive control. PhD Thesis, Drexel University, USA.

Zhang Z, Lam KP (2018). Practical implementation and evaluation of deep reinforcement learning control for a radiant heating system. In: Proceedings of the 5th Conference on Systems for Built Environments, Shenzen, China.

Zhang L, Wen J (2019). A systematic feature selection procedure for short-term data-driven building energy forecasting model development. *Energy and Buildings*, 183: 428–442.

Zhang C, Kuppannagari SR, Kannan R, et al. (2019a). Building HVAC scheduling using reinforcement learning via neural network based model approximation. In: Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, New York, USA.

Zhang Z, Chong A, Pan Y, et al. (2019b). Whole building energy model for HVAC optimal control: A practical framework based on deep reinforcement learning. *Energy and Buildings*, 199: 472–490.

Zhang L, Frank S, Kim J, et al. (2020a). A systematic feature extraction and selection framework for data-driven whole-building automated fault detection and diagnostics in commercial buildings. *Building and Environment*, 186: 107338.

Zhang X, Chintala R, Bernstein A, et al. (2020b). Grid-interactive multi-zone building control using reinforcement learning with global-local policy search. arXiv: 2010.06718.

Zhang X, Jin X, Tripp C, et al. (2020c). Transferable reinforcement learning for smart homes. In: Proceedings of the 1st International Workshop on Reinforcement Learning for Energy Management in Buildings & Cities.

Zhang L (2021). Data-driven building energy modeling with feature selection and active learning for data predictive control. *Energy and Buildings*, 252: 111436.

Zhang L, Wen J (2021). Active learning strategy for high fidelity short-term data-driven building energy forecasting. *Energy and Buildings*, 244: 111026.

Zhang L, Leach M, Bae Y, et al. (2021a). Sensor impact evaluation and verification for fault detection and diagnostics in building energy systems: A review. *Advances in Applied Energy*, 3: 100055.

Zhang L, Wen J, Li Y, et al. (2021b). A review of machine learning in building load prediction. *Applied Energy*, 285: 116452.

Zhang W, Wen Y, Tseng KJ, et al. (2021c). Demystifying thermal comfort in smart buildings: An interpretable machine learning approach. *IEEE Internet of Things Journal*, 8: 8021–8031.

Zhang X, Biagioni D, Cai M, et al. (2021d). An edge-cloud integrated solution for buildings demand response using reinforcement learning. *IEEE Transactions on Smart Grid*, 12: 420–431.

Zhou Y, Li D, Spanos CJ (2015). Learning optimization friendly comfort model for HVAC model predictive control. In: Proceedings of 2015 IEEE International Conference on Data Mining Workshop (ICDMW), Atlantic City, NJ, USA.