Does historical data still count? Exploring the applicability of smart building applications in the post-pandemic period

Xiang Xie[a,b,*], Qiuchen Lu[c], Manuel Herrera[a], Qiaojun Yu[d], Ajith Kumar Parlikad[a,b], Jennifer Mary Schooling[b,e]

[a] *Institute for Manufacturing, University of Cambridge, Cambridge, UK*
[b] *Centre for Digital Built Britain, University of Cambridge, Cambridge, UK*
[c] *The Bartlett School of Construction and Project Management, University College London, London, UK*
[d] *Artificial Intelligence Institute, Shanghai Jiao Tong University, Shanghai, China*
[e] *Centre for Smart Infrastructure and Construction, University of Cambridge, Cambridge, UK*

Abstract

The emergence of COVID-19 pandemic is causing tremendous impact on our daily lives, including the way people interact with buildings. Leveraging the advances in machine learning and other supporting digital technologies, recent attempts have been sought to establish exciting smart building applications that facilitates better facility management and higher energy efficiency. However, relying on the historical data collected prior to the pandemic, the resulting smart building applications are not necessarily effective under the current ever-changing situation due to the drifts of data distribution. This paper investigates the bidirectional interaction between human and build-

[*] Corresponding author

Email addresses: xx809@cam.ac.uk (Xiang Xie), qiuchen.lu@ucl.ac.uk (Qiuchen Lu), amh226@cam.ac.uk (Manuel Herrera), yqjllxs@sjtu.edu.cn (Qiaojun Yu), aknp2@cam.ac.uk (Ajith Kumar Parlikad), jms33@cam.ac.uk (Jennifer Mary Schooling)

ings that leads to dramatic change of building performance data distributions post-pandemic, and evaluates the applicability of typical facility management and energy management applications against these changes. According to the evaluation, this paper recommends three mitigation measures to rescue the applications and embedded machine learning algorithms from the data inconsistency issue in the post-pandemic era. Among these measures, incorporating occupancy and behavioural parameters as independent variables in machine learning algorithms is highlighted. Taking a Bayesian perspective, the value of data is exploited, historical or recent, pre- and post-pandemic, under a people-focused view.
Keywords: Post-pandemic, Smart building, Historical data, Machine learning

---

1. Introduction

Beginning in late 2019, the COVID-19 pandemic swept across the globe. As a crisis exerting a massive impact on human health, the suspension of civic and commercial activities across the world will inevitably cause huge ramifications afterwards. Despite complete lockdowns being released deliberately and gradually in many parts of the world, social distancing is still needed in short-term and medium-term to mitigate the spread of coronavirus, and additional lockdowns are being imposed periodically. This means the way people work and live has changed and will keep changing, while for many organisations/sectors some of these changes will be long term or even permanent (Kramer & Kramer, 2020). First and foremost, flextime is likely to become much more common, and perhaps even replace the 9-5 working regime altogether. These disruptive changes to life and work landscape bring the needs to reset the way we use our buildings and the opportunities to reshape the way how we manage our buildings (Megahed & Ghoneim, 2020; Ramsetty & Adams, 2020), to restore wellbeing, productivity and sustainability during the post-pandemic period.

With the overwhelming adoption of information and communication technologies (ICTs) in the built environment, the concept of the smart building has evolved as a comprehensive solution to offer convenience and comfort to their inhabitants, whilst enhancing operational efficiency (Qolomany et al., 2019; Verma et al., 2019). For example, Iqbal et al. (2018) presented a Zigbee based internet of things (IoT) architecture and proposed a Hadoop based data processing system for controlling electrical energy consumption in sustainable smart homes; Dong et al. (2019) reviewed the applications of smart building sensing systems and analysed the use in terms of energy saving, thermal comfort, visual comfort, and indoor air quality; Jia et al. (2019a) summarised the adoption of IoT technologies in smart buildings and explored their use in facilitating indoor localisation and resource tracking, energy management and facility management. The UK National Infrastructure Commission 'Data for the public good' (National Infrastructure Commission, 2017) recognised that verifiable, timely and accessible data is essential in unlocking the value of built environment. However, just like crude oil, data is valuable, but if unrefined it cannot really be used.

The massive amounts of data collected from buildings must be analysed, transformed into information, and minted to extract knowledge so that determinable insights can be acquired accordingly (Gunay et al., 2019). Machine learning (ML) is generally an appropriate strategy for building data analysis in the cases where neither prior system knowledge nor human expertise is enough to solve the problem directly. Through the continuous learning of significant quantities of quality and comprehensive data, the performance of the ML algorithms constantly improves under given model structure. Historical data that keeps snapshots of building states and inhabitants' behaviours provides supplemental insights for inferring building system dynamics using ML algorithms with less analyst intervention.

One size does not fit all. It remains to be seen whether the historical data collected and the ML solutions developed prior to the COVID pandemic still works in the post-pandemic situation. Ideally, ML algorithms must generalise from training data to the entire domain of all unseen observations so that it can make accurate extrapolations in all circumstances (Hoffer et al., 2017). However, that is not the case in reality. Faced with a post-pandemic situation that we have never seen before, the effectiveness of developed ML solutions along with the applicability of adopted historical data must be re-evaluated and re-verified. This paper attempts to give preliminary answers and examples to the following emerging questions: Is the historical data collected before the outbreak still useful post-pandemic? How do ML algorithms deal with the post-pandemic situation under ever-changing social distancing restrictions? What role do smart building applications play in making best use of data adaptively during the transition from the pre-pandemic to the post-pandemic and the future new normal?

The remainder of the paper is organised as follows. Section 2 overviews the bidirectional interaction between human and buildings, which causes dramatic change of building performance after the pandemic. Section 3 explores the applicability of typical smart building applications in the domain of facility management and energy management, and proposes migration measures that fix data inconsistency issue between pre- and post-pandemic periods. Section 4 describes a real-life case, demonstrating the impact of pandemic on building energy demand forecasting. The possibility of using transfer learning to speed up the convergence of ML enabled smart building applications are discussed in Section 5. Finally, Section 6 presents the conclusions.

2. Bidirectional relationships between humans and buildings

"We shape our buildings, thereafter they shape us". As Sir Winston Churchill said in his speech to the meeting in the House of Lords in 1943, the interaction between humans and buildings is intensive. Logically, environmental, contextual and personal factors first affect human' behaviours, then influence our built environment, and vice versa (Hong et al., 2017). Different occupancy patterns, inhabitants' lifestyle/habits, comfort preferences and associated actions lead to distinct building system performances even for buildings of the same type (Papakostas & Sotiropoulos, 1997; Leth-Petersen & Togeby, 2001; Lindén et al., 2006; Andersen et al., 2009; Maier et al., 2009). The bidirectional interactive relationships between the occupants' behaviours and building performances are the most important linkages between human and building. To be more precise, occupants' behaviours include both their presence and actions (Schweiker et al., 2018), impacting buildings in two ways: through the direct impact of their presence (heat etc.) and through their interaction with building systems (actions of turning on or off heating, ventilation and air conditioning (HVAC), lighting etc.).

Specifically,based on comprehensive analysis of human behaviours in buildings, the driving factors of human' behaviours in the built environment can be classified as: environmentally related factors, time related factors, contextual factors, physiological factors, psychological factors, social factors and random factors (Fabi et al., 2012; Inkarojrit, 2012; Stazi et al., 2017).

As can be seen in Table 1, in addition to environmental related factors, time related factors (e.g., personal habits) play a crucial role on the behaviours of inhabitants (Day et al., 2020). For instance, window opening behaviours in buildings are strongly related to the daily activities of occupants, such as sleeping, cooking and studying. Particularly, routine activities would play a decisive role in taking actions of opening and closing behaviours. In an office building, window open/close actions are usually affected by arrival and departure activities (i.e. open on arrival and close before departure) and daily working schedule (Pan et al., 2018). In a residential building, window open/close actions would not be time dependent, but rather activity dependent, such as during specific activities (e.g., cooking) (Pan et al., 2018). Generally, occupant behaviours in office buildings is more regular and constrained than in residential buildings (Day et al., 2020).

Subsequently, human behaviours can affect the building systems and their performances in return, such as energy consumption and facility services, among many others (Day et al., 2020; Laaroussi et al., 2020). For instance, to improve indoor comfort or energy efficiency, occupants may keep windows closed when the air conditioner is running or open/close the windows to adjust room comfort level and maintain indoor air quality. Human ac-

Behaviour	Drivers	Details	Affected by COVID-19
Window use	Environmental	Outdoor temperature	N
		Indoor temperature	N
		CO_2 concentration	N
	Time-related	Time of the day	**Y**
Light switching	Environmental	Work plane illuminance	N
		Illuminance	N
	Time-related	Arrivals and departures time	**Y**
Shading and blind use	Environmental	Illuminance	N
		Solar radiation	N
		Glare	N
		Outdoor and indoor temperature	N
	Time-related	Seasonal dependent	N
		Time of the day	**Y**
Air conditioning use	Environmental	Outdoor and indoor temperature	N
	Time-related	Weekday and weekend	**Y**
		Time of the day	**Y**
Thermostat use	Environmental	Outdoor and indoor temperature	N
Fans and doors	Environmental	Outdoor and indoor temperature	N

Note: ‘N’ stands for ‘No’ and ‘Y’ stands for ‘Yes’.

Table 1: Driving factors influencing occupants’ behaviours in buildings (Papakostas & Sotiropoulos, 1997; Leth-Petersen & Togeby, 2001; Lindén et al., 2006; Andersen et al., 2009; Maier et al., 2009; Stazi et al., 2017)

Category	Building Systems	Human Action
Lighting	Lighting system	Lighting operations
Opening	Blind	Blind operations
	Curtain	Curtain operations
	Windows/doors	Windows/doors operations
Heating & cooling	Heating system	Fans operations
		Air condition operations
	Cooling system	Heating radiator operations
Daily routine	Kitchen	Electric related activities (e.g., cooking, showering, TV)
	Toilet	
	Relaxing related system	Non-electric related activities (e.g., reading)
Housework	Interior cleaning	Electric related activities (e.g., washing machine)
	Laundry	
	Repair	Non-electric related activities (e.g., painting)
Daily work	Work related system	Electric related activities (e.g., computer)
		Non-electric related activities (e.g., writing)
Plant & animal	Plant care related system	Electric related activities (e.g., watering machine)
	Animal care related system	Non-electric related activities (e.g., gardening)

Table 2: A brief summary of human actions in buildings (Peng et al., 2012; Caba Heilbron et al., 2015; Laaroussi et al., 2020)

tions can be roughly classified as shown in Table 2, exerting impacts on the corresponding building systems.

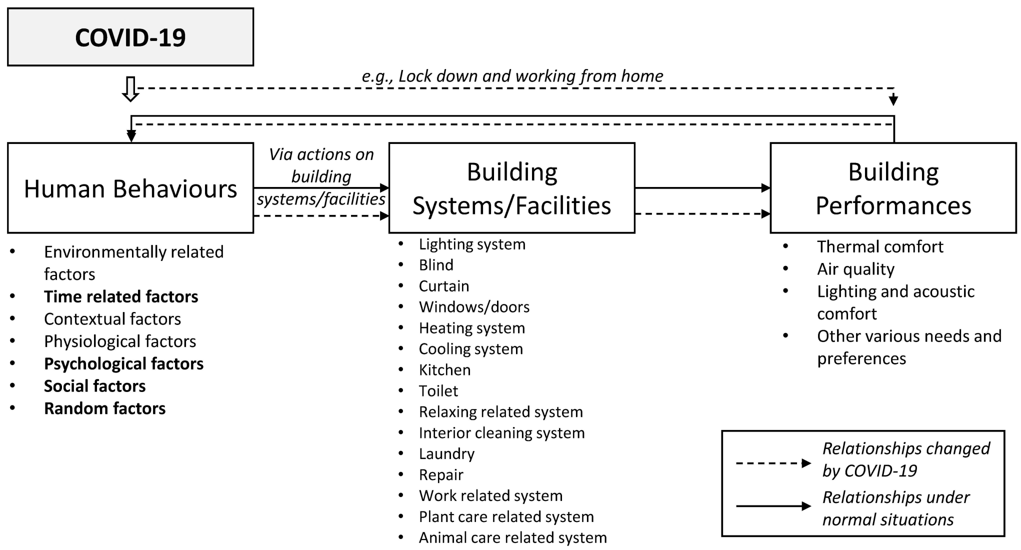


Figure 1: Bidirectional relationships of human behaviours and building performances in normal situations and affected by COVID-19

During the COVID-19 pandemic, most people changed their working patterns and had to work from home for establishing an antivirus-built environment. Residential buildings acted as both living and working spaces, while office buildings were left mostly empty or with limited accessibility (Megahed & Ghoneim, 2020). Internal, social, working and living variations can be identified as the key reasons for changing behavioural patterns observed in residential and non-residential buildings, especially during the pandemic. In detail, time related factors (working and living schedules), psychological factors (depression, anxiety and stress during lockdown), social factors (social distancing policy) and random factors give rise to the behavioural variations and further affect building systems/facilities and their performances (Figure 1). These changes bring great challenges to ML enabled smart building applications, which rely on the information and knowledge extracted from historical data. This paper seeks to address the limitations of these applications that suffer from the data inconsistency post-pandemic.

3. Machine learning for smart building applications

This section summarises the relevant literature from the perspective of ML techniques used in facility management (FM) and energy management (EM) domains (Xu et al., 2019; Kolokotsa et al., 2011), presented in Table 3. As physical assets that are built, installed, or established to serve the social and economic activities, facilities need to be monitored, operated and maintained properly for supporting and adding value to the business processes of organisations. Besides, accounting for over 40% of the total energy consumption in the world (Cao et al., 2016), building energy use needs to be measured, understood, controlled and optimised according to the spatial hierarchy of systems within buildings. Advanced data analytical tools, enabled by ML and other artificial intelligence (AI) techniques, stimulated a boom in intelligent initiatives and innovations to the FM and EM services, providing more efficient, responsive and environmentally friendly built environment to inhabitants.

Various review papers concerning FM or EM have been published since the intensive adoption of ML techniques in these domains, tracing back to Krarti (2003), which presented the applications of neural networks, fuzzy logic and genetic algorithms on building energy use prediction, envelope heat

Goal	Description
Facility management (Xu et al., 2019)	Managing facilities in the built environment at both strategic and day-to-day level to deliver operational objectives and to maintain a safe, efficient and sustainable environment
Energy management (Kolokotsa et al., 2011)	Maximising the building energy efficiency, with the ideal condition to be net-zero buildings (NZBs), while keep a satisfactory level of service at the same time

Table 3: Primary functional goals of smart buildings

transfer modeling, central plants control and fault diagnostics for building energy systems. Chicco (2012) assessed different types of clustering techniques for carrying out building electrical load pattern grouping, which guides grid demand response actions or real-time pricing. Yildiz et al. (2017) reviewed regression models for electricity load forecasting in commercial buildings, recommending multivariate linear regression for its greater user engagement and control. Miller et al. (2018) summarised the applications of unsupervised ML techniques to non-residential buildings for smart meters, portfolio analysis, operations and controls optimisation, anomaly detection and etc.

Instead of enumerating all ML techniques used in smart building applications, supervised, unsupervised, semi-supervised learning or reinforcement learning included, this paper focuses on clarifying the training needed for typical ML solutions and evaluating the applicability of ML enabled smart building applications trained with the historical data prior to the pandemic in the post-pandemic situation.

3.1. Facility management applications

Within the whole lifecycle of a building, the operation and maintenance (O&M) phase takes up the longest period, thirty to fifty years if not longer. Therefore, FM is regarded as one of the most important goals of smart buildings, significantly affecting operation, maintenance and repair cost, indoor comfort, and in the grander scheme global climate (Xu et al., 2019; Alfalah & Zayed, 2020). Generally, FM requires timely anomaly detection, control optimisation and predictive maintenance to ensure the facilities run under optimal conditions. Table 4 provides an overview of the representative publications in this category.

Author(s)	Data	Algorithm	Purpose	Applicable under pandemic
Anomaly detection				
Li & Wen (2014)	Data from typical HVAC grade sensors commonly found in AHUs	Combining wavelet transform with principle component analysis (PCA) method	Implementing automated fault detection and diagnostics in large commercial building HVAC system with operating data stored on the building automation system (BAS) central station	Yes, fluctuations caused by weather condition and internal load variations (occupancy related) are eliminated in the modeling
Fan et al. (2015)	Power consumption data from BAS for almost all electrical equipment, with duration of 1 year and sampling interval of 15-min	Quantitative association rule mining (QARM)	Utilizing rules discovered by QARM, facilitating better understanding the building operating behaviours, identifying non-typical operating conditions and detecting faulty conditions	No, rules discovered using pre-pandemic data is inapplicable in the post-pandemic period

Capozzoli et al. (2015)	Active electrical power for lighting and total active electrical power consumption data with 15 min interval, together with other independent variables (e.g., indoor/outdoor temperature)	Classification and regression tree (CART) coupled with generalized extreme studentized deviate (GESD) outliers detection	Detecting anomalous energy consumption that shows obvious difference from previous consumption with the similar boundary conditions	Yes, if the number of occupants is explicitly recorded and considered as independent variable
Li et al. (2016)	Building fault data collected by ASHRAE Research Project 1312	Information greedy feature filter (IGFF) and quadratic discriminant analysis, logistics regression, multiple support vector machine	Selecting the most informative features for building fault detection and diagnosis and chosen variables are fused together and plug into several classification techniques	No, the used fault labels do not specify the operational working conditions
Araya et al. (2017)	HVAC consumption data for a school recorded every 5 min from 2013 to 2015	Collective contextual anomaly detection using sliding window (CCAD-SW) framework	Monitoring building energy consumption with the aim to identify abnormal consumption behaviour and combining several anomaly detection classifiers to ensure diversification of anomaly classifiers	No, although calendar variables are incorporated (day, month and season), the method ignores other contextual information including occupancy

Touzani et al. (2019)	Simulated hourly energy consumption data generated using EnergyPlus for two types of DOE reference buildings	Non-routine event (NRE) detection using the CORT dissimilarity metric	Detecting NRE caused changes in building energy use, which contributes to the assessment of anomalous energy use behaviour	Yes, the short-term and temporary NREs can still be detected using limited data during the post-pandemic
Lu et al. (2020)	Averaged vibration frequency data for centrifugal pump with a sampling time of 1 hour	Bayesian online changepoint detection (BOCPD)	Realizing a continuous anomaly detection of building assets, filtering contextual anomalies through cross-referencing with external operation information	Yes, the method only depends on the temporal correlations of the found changepoints and the operational condition variations
Control optimisation				
May-Ostendorp et al. (2013)	Offline model predictive control (MPC) data conducted on a US office building prototype	Generalized linear models (GLM), classification and regression trees (CART), and adaptive boosting based rule extraction	Extracting rules from offline MPC results for a mixed mode building operated during the cooling season	No, the extraction is based on the original offline MPC optimisation which is generated prior to the pandemic

Domahidi et al. (2014)	Simulated data under 6 scenarios using BACLab simulator with a step of 1 hour and a prediction horizon of 24 h	Automated rule based control synthesis procedure for binary decisions using support vector machines (SVMs) and adaptive boosting (AdaBoost)	Proposing a novel way of synthesizing rule based controllers for energy efficient building control, which learn from simulation data with advanced control formulations	No, the controller is designed to work under the same working condition as the pre-pandemic
Qiu et al. (2020)	Measured weather and measured system cooling load data from a real HVAC system in a metro station in Guangzhou	Q-learning, a typical model-free reinforcement learning	Proposing a model-free optimal control method that is able to function and evolve simultaneously during the system operation period	Yes, the reinforcement learning allows the controller to learn from the continuous interactions with the environment through a trial and error process during the post-pandemic
Predictive maintenance				
Cauchi et al. (2017a)	Data from major overhaul every 20 years and inspections on a weekly basis	Fault maintenance trees (FMTs), modelled in the form of continous time Markov chains (CTMCs)	Incorporating maintenance and degradation models, and serving as a planning platform for balancing total costs and dependability of a system	No, under the reduced working loads, the mean times to failure (MTTF) for components are different from the anticipated values during the pre-pandemic

Cauchi et al. (2017b)	Biomass boiler input and output power data with duration of 1 year and sampling interval of 15-min	Model-based dynamic programming algorithm that computes the optimal maintenance action for cleaning or replacing the boiler	Proposing realistic predictive maintenance strategies that minimise the total operational costs of the boiler, the cleaning costs and the discomfort costs	No, only the data during the post-pandemic should be used for the predictive maintenance scheduling
Yang et al. (2018)	Work-order data collected in routine operations and maintenance	Failure mode and effects analysis (FMEA) method for HVAC prognostics	Generating an FMEA from a building cluster's work-orders to perform HVAC fault isolation and prognostics, ultimately estimating mean time between failures in future	No, only the data during the post-pandemic should be used in the FEMA modeling

Table 4: Publications from the facility management category

Anomaly detection: Anomaly detection for smart buildings focuses on revealing anomalous situations occurring within buildings, their subsystems and components, indicating the equipment faults or improper operations. The choice of the ML techniques used to be flexible, relying on the available data. However, the applicability of these ML enabled applications post-pandemic is compromised due to the drifts of data distribution, except for three situations. First, if a data preprocessing procedure removes the influence of changing operational conditions from raw data, and the ML techniques are used to extract occupancy-insensitive features, the solution should be applicable under the post-pandemic situation (Li & Wen, 2014). Second, if occupancy or relevant quantities (e.g., carbon dioxide concentration) are explicitly taken as independent variable in the ML algorithm, these solutions are likely to be compatible for post-pandemic scenarios (Capozzoli et al., 2015). Last, changepoint detection or non-routine event detection (Touzani et al., 2019; Lu et al., 2020) still works, because an obvious change in the statistical properties of data before and after the lockdown can be detected and the post-pandemic baseline can be modelled accordingly.

Control optimisation: Control optimisation aims at regulating the operational performance of building systems to efficient and sustainable status. Model predictive control (MPC) formulates the building dynamics into a mathematical model and selects an optimised control strategy accordingly (Maddalena et al., 2020). To realise a simple implementation, synthesised rule-based controllers extract decision rules from advanced control MPC schemes (May-Ostendorp et al., 2013; Domahidi et al., 2014). Relying on the modelled building dynamics prior to the pandemic, these controllers are inapplicable during the post-pandemic period considering the shifted occupancy behaviour (Gholamzadehmir et al., 2020). On the other hand, reinforcement learning (RL) is a promising candidate under the pandemic situation, because essentially, a RL based controller continuously adapts over time according to its interaction with buildings.And the initial learning phase is maintained moderate with existing pre-pandemic knowledge.

Predictive maintenance: Predictive maintenance designs to forecast the trend of facility performance degradation and deduce the optimal maintenance policy that minimises the overall cost (operational, maintenance, repair and etc.). Data driven or empirical based degradation modeling, which is the core of the predictive maintenance, predicts the remaining useful life using historical operational data or work-order data (Cauchi et al., 2017b; Yang et al., 2018). Apparently, degradation behaviour, sometimes involving interdependencies between associated facilities, is unpredictable under unseen working conditions. Therefore, it is necessary to accumulate post-pandemic data before predictive maintenance applications can be conducted again.

3.2. Energy management applications

Inefficient energy management, especially in aging buildings, will heighten the negative impacts to the environment and inevitably accelerate global warming and climate change at the macro level. In response to this challenge, ML techniques have been widely used to support building energy management and improve building energy efficiency (Molina-Solana et al., 2017), focusing on the topics including energy demand forecasting (Amasyali & El-Gohary, 2018), energy demand disaggregation (Armel et al., 2013), and energy demand response (Antonopoulos et al., 2020). Table 5 provides an overview of the representative publications in this category.

Author(s)	Data	Algorithm	Purpose	Applicable under pandemic
Energy demand forecasting				
Yu et al. (2010)	Residential energy consumption data along with ten predictor variables about indoor temperature, building envelop, appliance types and occupancy	Decision tree method	Providing accurate predictive models with interpretable flowchart-like tree structures that enable users to quickly extract useful information	Probably, the model should be valid if individual inhabitant's consumption behaviour is consistent before and after the pandemic
Edwards et al. (2012)	Residential energy consumption data collected from three different homes located in west Knox County, Tennessee	Least squares support vector machine (LS-SVM)	Predicting next hour residential building consumption	No, the method explicitly assumes occupancy pattern is consistent with typical energy usage patterns of American households
Zhang et al. (2016)	Daily energy consumption data consists of weekdays (Monday to Friday) from 1-Jan-13 to 31-Dec-13 and half-hourly consumption data from 02-June-2012 00:00 to 11-June-2012 23:30	Weighted support vector regression (SVR) with differential evolution optimisation	Forecasting both half-hourly and daily energy consumption without manually changing any model parameter	No, the original electricity consumption series is modelled directly without considering the behavioural factors
Energy demand disaggregation				

Ji et al. (2015)	Hourly energy consumption data in four commercial buildings (two office buildings and two shopping malls) in Shanghai	Fourier series model (FSM) based method	Calculating the lighting-plug, power and HVAC terminal end-use hourly electricity consumption in commercial buildings	No, the disaggregation assumes that lighting and equipment energy use vary periodically in daily and annual cycles in commercial buildings, which is not true during the post-pandemic
Niu et al. (2018)	Hourly energy consumption data and on-site weather data of Atlantic Fleet Drill Hall building at Great Lakes, IL, USA	Fourier series based decomposition method	Decoupling the HVAC electricity consumption from the total building electricity consumption	No, the HVAC related consumption, which is polynomially proportional to the outside air temperature, may change after the pandemic
Zhou et al. (2019)	Hourly data on appliance power consumption in commercial buildings in Shanghai and the meteorological data from the monitoring station of the China Meteorological Administration	Finite mixture model (FMM) based method	Identifying the behaviour pattern and the consumption of each appliance and disaggregating the total power consumption into the appliance-level power consumption	Can be, if the occupant behaviour is explicitly taken as influential factor in the FMM of power consumption of specific types of end-uses

Zhao et al. (2020)	The REFIT dataset (Murray et al., 2017) and the appliance manufacturer information	Optimisation based, graph signal processing (GSP) based and convolutional neural network (CNN) based disaggregation method	Load disaggregation for a specific range of appliances	No, survey needs to be redone, asking inhabitants to fill another one-off appliance questionnaires, e.g., the use frequency of appliances after the pandemic
Energy demand response				
Pallonetto et al. (2019)	Simulation data (using BCVTB) for a detached bungalow-type house within one month with 15-min resolution	Machine learning based predictor module and optimisation modules	Designing, developing and testing of an energy management system to provide demand response capabilities for residential buildings	No, the complicated predictive controls rely on multiple interactional predictor components, which is nearly impossible to generalise to unseen situation

Table 5: Publications from the energy management category

Energy demand forecasting: Measuring, modeling and forecasting the energy demand of buildings is crucial to realise smart buildings. Instead of relying on thermodynamic principles, most of ML-based energy demand forecasting learns from historical data (time-series prediction) with partial knowledge of on-site physical information (regression prediction). Solutions like Yu et al. (2010), which explicitly consider the generalisation for occupancy scheduling, are likely to be applicable post-pandemic. However, besides occupancy scheduling, occupant behaviour, like plugging, positively impacted the energy demand (Kim et al., 2020). It remains to be seen whether and to what extent changing occupant behaviour before and after the pandemic would lead to notable deviations between the predicted and the actual consumption levels.

Energy demand disaggregation: As far back as Hart (1992), energy demand disaggregation was proposed to provide fine-grained energy feedback by individual end-uses, which can potentially reduce domestic electricity consumption up to 4.5% compared to aggregated feedback (Kelly & Knottenbelt, 2016). Disaggregating total building energy consumption usually relies on the periodic pattern of specific end-uses (Ji et al., 2015) or their correlations with external conditions (Niu et al., 2018; Zhou et al., 2019). Unless occupants and their behaviour are explicitly considered in the end-use models, the ML solutions are not likely to be applicable under the post-pandemic situation. For appliance-level energy disaggregation, because the frequency and duration of use would be different from before, another round of appliance surveys is needed to gather updated appliance usage information.

Energy demand response: In order to reduce the investment in energy generation under peak demand, demand-side response aims to minimise consumption at times of high demand. With the penetration of Renewable Energy Systems (RESs), adapting energy demand further assists in reducing grid frequency instability. Essentially, the demand response actuates balancing strategies that coordinate the requirements and needs between the energy retailer and the customers. In the post-pandemic period, the price scheme from the retailer and the consumers' demand will not be the same as they used to be. Considering the increased complexity caused by multiple participants' behavioural sophistication (Panait & Luke, 2005), the intelligent agents needs to be completely retrained after understanding the emerging post-pandemic situation.

4. Smart building applications and their applicability post-pandemic

In this paper, the potential of using ML in smart buildings is summarised in four forms: component status recognition, system behaviour modeling and control optimisation as well as intra-system/inter-system coordination. Considering the scale of typical buildings, the computational resources that could be allocated to each building/system/component are usually constrained. As shown in Figure 2, the hierarchical design of smart building applications makes data refinement and information processing stepwise and affordable computationally.

Rather than layering according to the spatial granularity, the hierarchy for smart building applications is defined based on the analytical granularity. For instance, if the total building energy consumption is analysed for anomaly detection, it means that the entire building is regarded as one com-

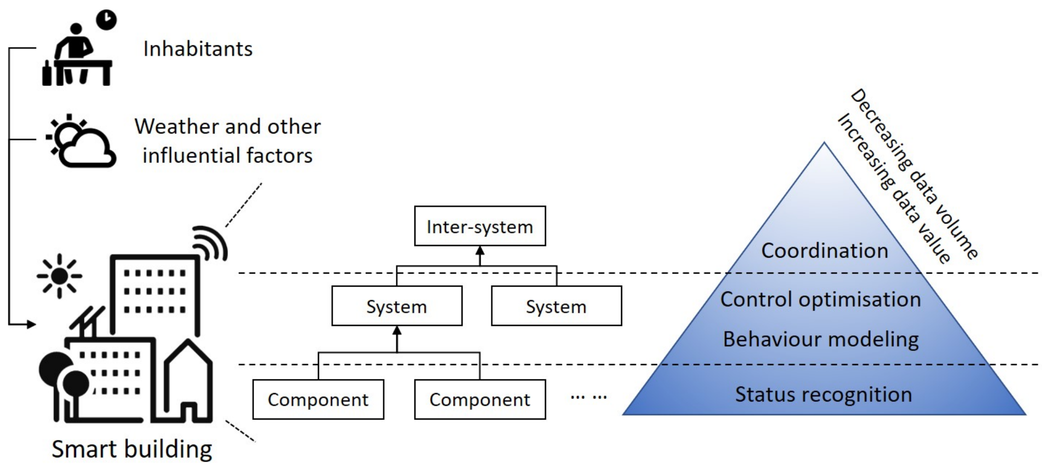


Figure 2: Typical forms of smart building applications

ponent. The overview of ML used in FM and EM applications gives us some clues in terms of their applicability in the post-pandemic situation. For applications at a lower level, one option is to eliminate the human presence/behaviour related components from raw data and only focus on the analysis of occupant-insensitive components (Li & Wen, 2014). However, the decomposition is rather empirical and leads to the loss of large amounts of information, which makes it unsuitable for complicated system modeling and analysis. For applications at a higher level, adaptive ML algorithms deserve a place. Basically, these algorithms treat historical data as a "starting point". Non-routine event (NRE) detection (Touzani et al., 2019) or changepoint detection (CPD) (Lu et al., 2020) are typical solutions in this category, which tracks the evolution of inspected time-series over time to achieve continuous learning. In terms of control optimisation, reinforcement learning is another example, which continuously interacts with the changing environment to gradually maximise the potential reward from a suboptimal start point

(Jia et al., 2019b; Wang & Hong, 2020). However, these options do not provide universal answers to the applicability in the post-pandemic period. Instead, explicitly incorporating occupancy and other behavioural parameters (associated with actions listed in Table 2) as independent variables in ML solutions could be a promising approach. A wide variety of modern sensors like thermal sensors and camera have been developed to detect occupancy accordingly (Roselyn et al., 2019). Taking energy demand forecasting applications as example, to make it work in practice, this approach should be in line with a probabilistic vision of a building energy model, since the uncertain nature associated with occupant behaviour and/or estimations of occupant density are embedded in the model. Rather than deterministic/fixed values for these variables, it is convenient to use probability distributions capturing their variability, and their consequent impact in the building energy model (Stewart et al., 2016). The fundamental basis for these considerations lies in Bayesian statistics which have already shown its capacity to quantify uncertainties in both building energy models and occupant density or other behavioural parameters (Tian et al., 2016). In particular, Bayesian networks have shown suitability for dealing with uncertainty in a plethora of cases. A key feature of the Bayesian network is the graphical representation of the mathematical model over the corresponding random variables (O'Neill & O'Neill, 2016). For instance, in Barthelmes et al. (2017), a Bayesian network is used to capture underlying complicated relationships between various influencing factors and window opening/closing behaviour of occupants in residential buildings. This leads to a better understanding of the correlation structure between the involved variables. Amayri et al. (2019) use Bayesian networks to estimate the number of occupants through its relationship to a number of variables collected by a series of sensors. Similarly, this paper remarks on the outstanding correlation between occupancy and energy consumption. As a consequence, building energy analysis, within a post-pandemic framework, should highlight the changed level of occupancy caused by the enforced social-distancing policies. And the pre-pandemic trained energy models need to be verified in new scenarios of occupant density and behaviour. Once again, a Bayesian perspective will be able to address this paradigm. Specifically, Bayesian structural time series, introduced in the works of Brodersen et al. (2015) and Scott & Varian (2014) among others, are capable to estimate the causal effect of any intervention in a time series regression. The intervention process can be understood as a change to a procedure or policy, exogenous to the modelled time series but having an impact on its outcome, in this case, occupant density and occupant behaviour. The causal effect of such an intervention is estimated by a comparison between the predicted outcome under the hypothesis of no-intervention in a post-intervention period and the actual observed time series in such a period.

5. Case study

Building energy consumption, one of the most important aspects in defining the usage of buildings, is selected in the case study to illustrate the impact of the pandemic on building electricity usage and potential energy demand forecasting applications. It has been verified in Santin et al. (2009), Cvetković et al. (2020) and Chen et al. (2020) that occupant behaviours inside a building, their presence included, significantly affect the total energy use (e.g., around 4.2% of the variation in energy use for space heating), and in specific scenario analysed, the consumption of natural gas can increase by 21.26%, electricity by 58.39% compared with pre-pandemic in the residential sector.

In this paper, the electricity consumption data from a university building in West Cambridge site of University of Cambridge is used to examine the pre- and post-pandemic data inconsistency issue for potential energy demand forecasting applications. The Alan Reece building, shown in Figure 3, is a three-story building standing over a 40,000 square foot comprehensive area. It includes spaces with diverse uses, such as teaching, office, research, laboratory, canteen and etc. The electricity consumption (kWh per half an hour) and local ambient temperature of the Alan Reece building from 13th October 2018 to 11th October 2020 is used, partially shown in Figure 4. However, no specific occupancy data is available to be incorporated.

During the period, the building went through a few distinct operational stages. Before 20th March (first vertical line), the building was under normal operation. Approximately 180 staff and PhD students regularly worked in the building, with roughly 60% occupation rate daily on weekdays considering their holidays due to research/teaching/working patterns. Meanwhile, a variable number of undergraduate and postgraduate students used the building as well. In response to the intensive transmission of the coronavirus, the University moved into its "red" phase on 18th March, and the Alan Reece building was shut down from 5pm on 20th March. Within the first phase of the lockdown, until 22nd June (second vertical line), an extremely limited number of COVID-related research activities were allowed on sporadic days, with the presence of around 5 to 6 people typically. Later on, the building

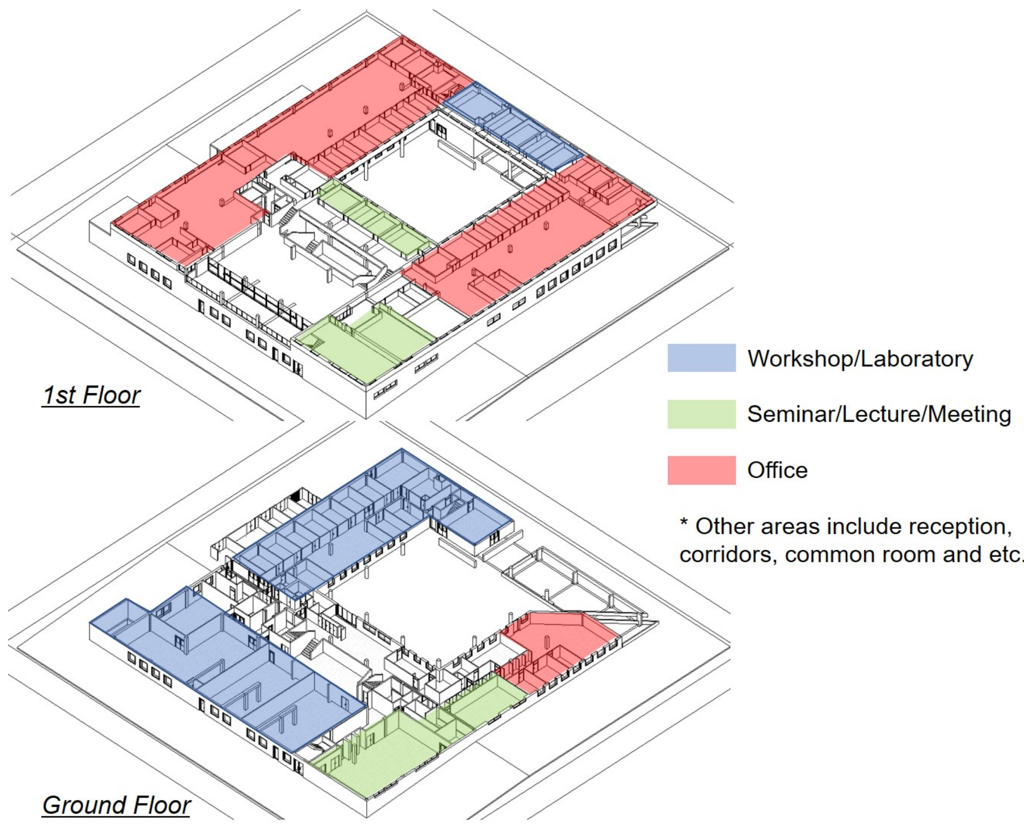


Figure 3: Layout of the Alan Reece Building

gradually restored its functionalities by admitting more people in the next two stages (until early-August and 5th October, third and forth vertical lines), with less than 10 and 20 people admitted on working days respectively. While with the Michaelmas term starting from 5th October, groups of students were also allowed back into the building, leading to around 50 people at any one time during the weekdays.

The electricity consumed in the Alan Reece building can be break down into several end-uses, i.e., air conditioning, space heating/cooling, water heating, lighting, refrigeration, appliance and other plug loads. Three indepen-

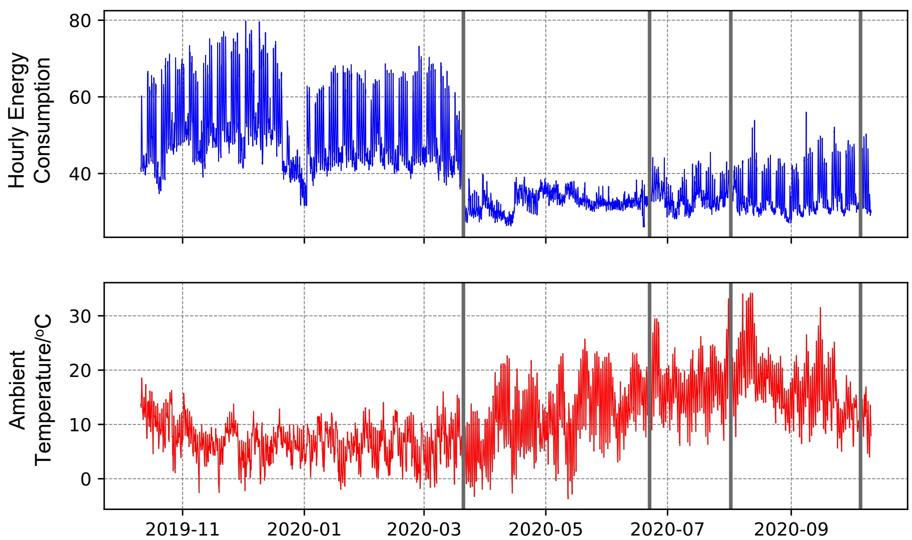


Figure 4: Hourly building electricity consumption ($kWh \cdot hh^{-1}$) and ambient temperature

dent air handling units (AHUs) are installed to regulate indoor air quality, while variable refrigerant flow (VRF) heat pump system is installed to provide simultaneous heating and cooling to different areas within the building. The appliances include personal equipment (e.g., laptop, mobile phone), experimental equipment (e.g., 3D printer, lathe) and large equipment (e.g., elevator). Besides temperature and other ambient parameters, the level of occupancy has a decisive influence on most of these end-uses. During the pandemic, for AHUs and VRF system, the decrement in occupancy reduces the HVAC load accordingly, and for water heating, lighting, refrigeration and appliances, their usage frequency and duration drop significantly as well. Statistically speaking, compared to the same periods (April to September) in 2019, the electricity consumption of the reference building saw a decrease of

36.7% in average.

The causal impact model developed by Google (Brodersen et al., 2015) is used to reveal the intervention of behavioural change caused by COVID-19 pandemic and the consequent lockdown, which directly affect the building occupancy and the energy use. The model is based on Bayesian structural time series, a regression state-space model that predicts the daily energy consumption response in case of no lockdown (no intervention) taking place. The resulting model is called counterfactual and its predictions are compared to the observed time series after the intervention (post-pandemic period) to infer its effect. The counterfactual model is, therefore, of pivotal importance for the success of the causal impact model and its computation is typically based on the combination of: the historical time series model prior to the lockdown, the relationship to such a time series with any exogenous, predictive variables no being impacted by the lockdown (e.g., ambient temperature), and any available prior information about the possible results of the lockdown facing a similar contextual change (e.g., building occupancy level).

Figure 5 contains 3 panels. The top panel shows the daily energy consumption and a counterfactual prediction for the post-lockdown period. The second and third panel show the difference between the observed energy consumption and the counterfactual predictions. The difference is shown cumulatively over time in the case of the bottom panel.

A consequence of the analysis shown in Figure 5 is the effect of the lockdown, and the corresponding drop in the building occupancy level. The predictive model trained prior to the pandemic, relying on historical energy consumption, calendar variable and ambient temperature cannot keep its ac-

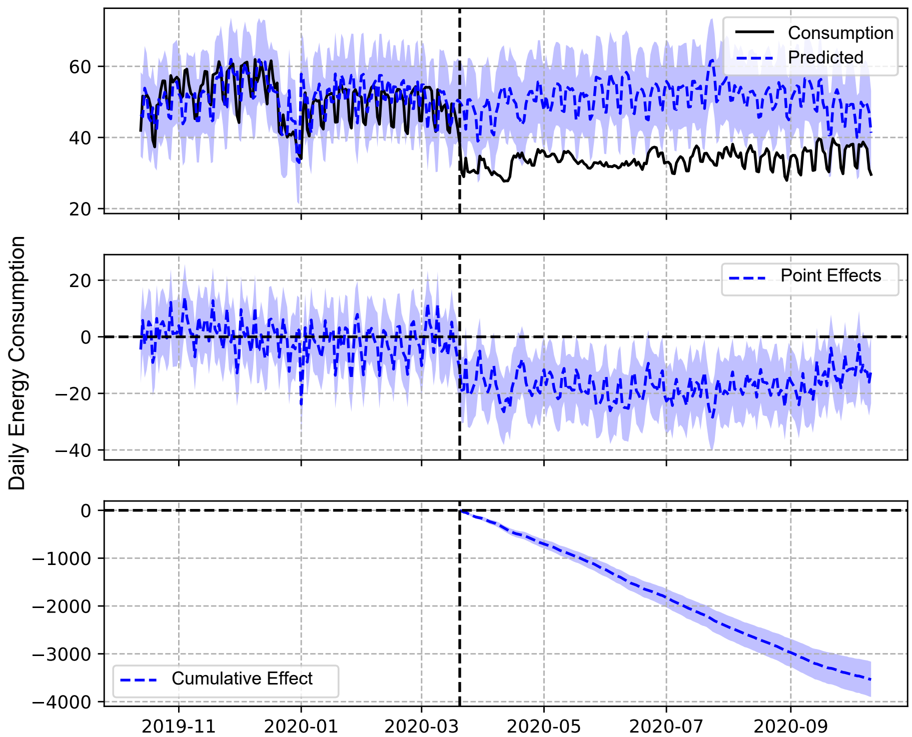
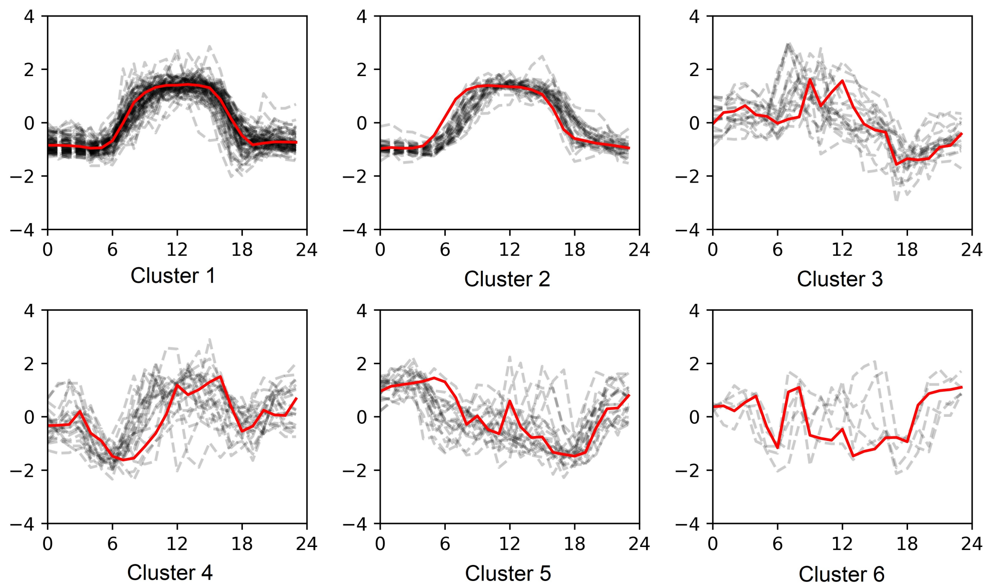
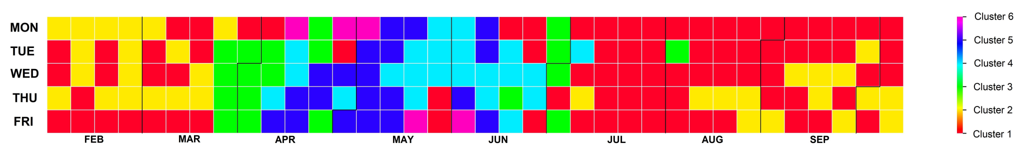


Figure 5: Causal effect of lockdown on the daily energy consumption ($kWh \cdot hh^{-1}$) at the IfM building. Top panel: time series of energy pre- and post-pandemic (observed vs. predicted). Middle panel: pointwise causal effect (difference counterfactual predictions and observed). Bottom panel: cumulative effect of the lockdown on the energy consumption.

curacy and it is largely biased for the post-pandemic period, when the occupancy level is significantly lower. During the post-pandemic period, the daily energy consumption has an average value of approximately 34 $kWh \cdot hh^{-1}$. By contrast, in the absence of a lockdown, we would have expected an average consumption of 51 $kWh \cdot hh^{-1}$. The 95% credible interval of this counterfactual prediction is [49.03, 52.76]. Therefore, a rough estimation of the causal effect the lockdown has on the energy consumption is -17 $kWh \cdot hh^{-1}$ with a corresponding 95% credible interval estimation of [-19.05, -15.32].



(a) TDUE profiles



(b) Distribution of the TDEU profiles

Figure 6: Results of DEU profiles clustering

To better describe the building energy consumption characteristics, 24-dimensional daily energy usage (DEU) profiles are segmented from the time series building electricity consumption data (Li et al., 2018). Clustering analysis is adopted to identify typical daily energy usage (TDEU) profiles and highlight the drifting of the weekday DEU profiles with time. A Gaussian mixture model (GMM) based cluster analysis is used to cluster DEU profiles (Li et al., 2018) and the median of all DEU profiles in the same cluster is considered as TDEU profile of this cluster. Figure 6 (a) illustrates the DEU clustering results, in which the red curves represents the 6 TDEU profiles identified while the grey curves are all corresponding DEU profiles belonging to that cluster. Figure 6 (b) shows the distribution of the TDEU profiles in a calendar view. Notable DEU drifting can be observed during the transition between different stages, particularly at the end of March and the end of June. Although the amount of electricity consumed after August is still much lower than the pre-pandemic level, the cluster of DEU almost gets back in the swing, suggesting that various types of activities reoccurred in the building with restricted intensity. It is fair to say that the historical data before 20th March can be used if the activities, which are the consequences of occupant behaviours, are properly recorded.

The evidences above validate the hypothesis that the occupancy level and other behavioural factors are informative in energy demand forecasting, and more broadly, enable the applicability of smart building applications to be restored their applicability within frequently-changing scenarios.

6. Discussion

As stated in Schooling et al. (2020), the fundamental purpose of the wider built environment and the infrastructure embedded is to provide a platform for human flourishing, to better serve people and society. The people-focused view is shared among buildings as well. However, more than simply providing comfortable and efficient living environment to people, the authors believe that a people-focused view also means to monitor, analyse, comprehend and sometimes influence the interactive behaviour of people with buildings, which is in accordance with the conclusion drawn in Alfalah & Zayed (2020) and Laaroussi et al. (2020). Learning from this crisis brought by the global pandemic, the development of smart building applications must be based on ML techniques that are robust to societal variations, and stand on a social-technical basis. Particularly, under this crisis, social-distancing and lockdown practices are introduced in a localised and adaptive manner (Rahman et al., 2020), and occupancy density is to be regulated to avoid long time exposure and prevent COVID transmission (Sun & Zhai, 2020). Accordingly, these social parameters need to be monitored and taken into consideration during the deployment of smart building applications (Ahmed et al., 2021).

Data is the soul of the digitalisation and intelligentisation of the buildings. However, we have to recognise that data comes with costs. Data generation, transmission, processing and even storage are quite expensive, particularly for the data involved with occupants' presence/behaviour. The interruption caused by the COVID-19 pandemic is likely to cause enormous loss regarding the applicability of historical data as the training basis, if the occupancy data was not properly collected. So how should we deal with existing ML algorithms, that come without the proper reference to occupants' status?

Learning from the past, let us take a glance at the existing measures taken to cope with the limited data availability problem. If historical data of buildings is limited but there are similar buildings with significant quantities of historical data, in such cases, transfer learning, transferring knowledge and experience learned from similar buildings to empower the ML with reasoning ability and fast convergence, is an important approach to tackle the problems. For instance, Mocanu et al. (2016) focus on cross-building transfer learning, combining reinforcement learning with deep belief network, and using data from other buildings to predict energy consumption for buildings with limited historical data; Ribeiro et al. (2018) propose a cross-building transfer learning method named Hephaestus, which is based on time series multi-feature regression with seasonal and trend adjustment for cross-building energy forecasting.

Great changes have taken place in people's behavioural habits, and data from buildings post-pandemic phase may drift to a distinguishing distribution, which invalidates much of the historical data without the reference to occupancy status. In the cases where the data collected during the novel post-pandemic situation is insufficient to train a ML model, an effective approach is to transfer useful data to the target building from other source buildings with similar purpose and functions (Liu et al., 2017). Specifically, for filtering appropriate knowledge to be transferred to the target building, we can adjust the transferability weight for each data sample from source buildings according to their similarity to the data samples from the target building. Through the fusion of weighted multi-source building data, the size of the available training samples for the target building increases, thus potentially improving the convergence speed and model accuracy of the ML enabled smart building applications.

7. Conclusion

The coronavirus pandemic has brought astonishing upheavals to the world, and of course to buildings with diverse uses as well. With the pervasive digital transformation of buildings, a diverse selection of smart building applications have been developed to sophisticatedly extract and infer knowledge from data and support corresponding decision-making processes, especially in the domains of facility management and energy management. These approaches have suffered from the changed interactive pattern between humans and buildings during the pandemic, including but not limited to the variations of occupancy and occupants' behaviour. As a result, current smart building applications, which heavily rely on a certain volume of pre-pandemic data to feed into machine learning algorithms, might fail. To reveal the impact of pandemic on smart building applications, the interactive relationships between human and buildings are described, and an evaluation of the applicability is presented for typical ML enabled smart building applications trained with historical data, most of which has been collected prior to the pandemic. Six categories of applications were reviewed in this paper, including anomaly detection, control optimisation, predictive maintenance, energy demand forecasting, energy demand disaggregation and energy demand response.

This paper suggests three measures to mitigate the data inconsistency issue for practical smart building applications in the post-pandemic era. For relatively simple analysis, eliminating the effect of occupants' behaviour by decomposing occupancy-insensitive features is effective, with the cost of losing partial information. Alternatively, adaptive ML algorithms, using which the evolution of building systems is tracked over time, are immune from the after-effect of the pandemic. However, to provide a universal answer, it is recommended to incorporate occupancy and other behavioural parameters as independent variables in the conventional ML algorithm. To this end, Bayesian ML models, including Bayesian networks, deserve a place due to their natural capability to deal with the uncertainty within occupancy related variables. Through incorporating these variables, smart building applications can take full advantage of data, both pre- and post-pandemic, under a people-focused view.

Acknowledgements

This research forms part of the Centre for Digital Built Britain's (CDBB) work at the University of Cambridge within the Construction Innovation Hub (CIH). The Construction Innovation Hub is funded by UK Research and Innovation through the Industrial Strategy Fund.

References

Ahmed, I., Ahmad, M., Rodrigues, J. J., Jeon, G., & Din, S. (2021). A deep learning-based social distance monitoring framework for COVID-19. *Sustainable Cities and Society*, *65*, 102571.

Alfalah, G., & Zayed, T. (2020). A review of sustainable facility management research. *Sustainable Cities and Society*, *55*, 102073.

Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, *81*, 1192–1205.

Amayri, M., Ploix, S., Kazmi, H., Ngo, Q.-D., & Safadi, E. A. E. (2019). Estimating occupancy from measurements and knowledge using the bayesian network for energy management. *Journal of Sensors*, *2019*, 7129872.

Andersen, R. V., Toftum, J., Andersen, K. K., & Olesen, B. W. (2009). Survey of occupant behaviour and control of indoor environment in danish dwellings. *Energy and Buildings*, *41*, 11–16.

Antonopoulos, I., Robu, V., Couraud, B., Kirli, D., Norbu, S., Kiprakis, A., Flynn, D., Elizondo-Gonzalez, S., & Wattam, S. (2020). Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. *Renewable and Sustainable Energy Reviews*, *130*, 109899.

Araya, D. B., Grolinger, K., ElYamany, H. F., Capretz, M. A., & Bitsuamlak, G. (2017). An ensemble learning framework for anomaly detection in building energy consumption. *Energy and Buildings*, *144*, 191–206.

Armel, K. C., Gupta, A., Shrimali, G., & Albert, A. (2013). Is disaggregation the holy grail of energy efficiency? The case of electricity. *Energy Policy*, *52*, 213–234.

Barthelmes, V. M., Heo, Y., Fabi, V., & Corgnati, S. P. (2017). Exploration of the bayesian network framework for modelling window control behaviour. *Building and Environment*, *126*, 318–330.

Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., Scott, S. L. et al. (2015). Inferring causal impact using bayesian structural time-series models. *The Annals of Applied Statistics*, *9*, 247–274.

Caba Heilbron, F., Escorcia, V., Ghanem, B., & Carlos Niebles, J. (2015). Activitynet: A large-scale video benchmark for human activity understanding. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 961–970).

Cao, X., Dai, X., & Liu, J. (2016). Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy and buildings*, *128*, 198–213.

Capozzoli, A., Lauro, F., & Khan, I. (2015). Fault detection analysis using data mining techniques for a cluster of smart office buildings. *Expert Systems with Applications*, *42*, 4324–4338.

Cauchi, N., Hoque, K. A., Abate, A., & Stoelinga, M. (2017a). Efficient probabilistic model checking of smart building maintenance using fault maintenance trees. In *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments* (pp. 1–10).

Cauchi, N., Macek, K., & Abate, A. (2017b). Model-based predictive maintenance in building automation systems with user discomfort. *Energy*, *138*, 306–315.

Chen, C.-f., de Rubens, G. Z., Xu, X., & Li, J. (2020). Coronavirus comes home? energy use, home energy management, and the social-psychological factors of COVID-19. *Energy research & social science*, *68*, 101688.

Chicco, G. (2012). Overview and performance assessment of the clustering methods for electrical load pattern grouping. *Energy*, *42*, 68–80.

Cvetković, D., Nešović, A., & Terzić, I. (2020). Impact of people's behavior on the energy sustainability of the residential sector in emergency situations caused by COVID-19. *Energy and Buildings*, *230*, 110532.

Day, J. K., McIlvennie, C., Brackley, C., Tarantini, M., Piselli, C., Hahn, J., O'Brien, W., Rajus, V. S., De Simone, M., Kjærgaard, M. B. et al. (2020). A review of select human-building interfaces and their relationship to human behavior, energy use and occupant comfort. *Building and Environment*, (p. 106920).

Domahidi, A., Ullmann, F., Morari, M., & Jones, C. N. (2014). Learning decision rules for energy efficient building control. *Journal of Process Control*, *24*, 763–772.

Dong, B., Prakash, V., Feng, F., & O'Neill, Z. (2019). A review of smart building sensing system for better indoor environment control. *Energy and Buildings*, *199*, 29–46.

Edwards, R. E., New, J., & Parker, L. E. (2012). Predicting future hourly residential electrical consumption: A machine learning case study. *Energy and Buildings*, *49*, 591–603.

Fabi, V., Andersen, R. V., Corgnati, S., & Olesen, B. W. (2012). Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models. *Building and Environment*, *58*, 188–198.

Fan, C., Xiao, F., & Yan, C. (2015). A framework for knowledge discovery in massive building automation data and its application in building diagnostics. *Automation in Construction*, *50*, 81–90.

Gholamzadehmir, M., Del Pero, C., Buffa, S., Fedrizzi, R. et al. (2020). Adaptive-predictive control strategy for HVAC systems in smart buildings–A review. *Sustainable Cities and Society*, (p. 102480).

Gunay, H. B., Shen, W., & Newsham, G. (2019). Data analytics to improve building performance: A critical review. *Automation in Construction*, *97*, 96–109.

Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, *80*, 1870–1891.

Hoffer, E., Hubara, I., & Soudry, D. (2017). Train longer, generalize better: closing the generalization gap in large batch training of neural networks. In *Advances in Neural Information Processing Systems* (pp. 1731–1741).

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

Inkarojrit, V. (2012). *Balancing comfort: occupants' control of window blinds in private offices*. Ph.D. thesis University of California, Berkeley.

Iqbal, J., Khan, M., Talha, M., Farman, H., Jan, B., Muhammad, A., & Khattak, H. A. (2018). A generic internet of things architecture for controlling electrical energy consumption in smart homes. *Sustainable cities and society*, *43*, 443–450.

Ji, Y., Xu, P., & Ye, Y. (2015). HVAC terminal hourly end-use disaggregation in commercial buildings with fourier series model. *Energy and Buildings*, *97*, 33–46.

Jia, M., Komeily, A., Wang, Y., & Srinivasan, R. S. (2019a). Adopting internet of things for the development of smart buildings: A review of enabling technologies and applications. *Automation in Construction*, *101*, 111–126.

Jia, R., Jin, M., Sun, K., Hong, T., & Spanos, C. (2019b). Advanced building control via deep reinforcement learning. *Energy Procedia*, *158*, 6158–6163.

Kelly, J., & Knottenbelt, W. (2016). Does disaggregated electricity feedback reduce domestic electricity consumption? A systematic review of the literature. *arXiv preprint arXiv:1605.00962*, .

Kim, M. K., Kim, Y.-S., & Srebric, J. (2020). Impact of correlation of plug load data, occupancy rates and local weather conditions on electricity consumption in a building using four back-propagation neural network models. *Sustainable Cities and Society*, *62*, 102321.

Kolokotsa, D., Rovas, D., Kosmatopoulos, E., & Kalaitzakis, K. (2011). A roadmap towards intelligent net zero-and positive-energy buildings. *Solar Energy*, *85*, 3067–3084.

Kramer, A., & Kramer, K. Z. (2020). The potential impact of the Covid-19 pandemic on occupational status, work from home, and occupational mobility. *Journal of Vocational Behavior*, *119*, 103442.

Krarti, M. (2003). An overview of artificial intelligence-based methods for building energy systems. *Journal of Solar Energy Engineering*, *125*, 331–342.

Laaroussi, Y., Bahrar, M., El Mankibi, M., Draoui, A., & Si-Larbi, A. (2020). Occupant presence and behavior: a major issue for building energy performance simulation and assessment. *Sustainable Cities and Society*, (p. 102420).

Leth-Petersen, S., & Togeby, M. (2001). Demand for space heating in apartment blocks: measuring effects of policy measures aiming at reducing energy consumption. *Energy Economics*, *23*, 387–403.

Li, D., Zhou, Y., Hu, G., & Spanos, C. J. (2016). Optimal sensor configuration and feature selection for AHU fault detection and diagnosis. *IEEE Transactions on Industrial Informatics*, *13*, 1369–1380.

Li, K., Ma, Z., Robinson, D., & Ma, J. (2018). Identification of typical building daily electricity usage profiles using gaussian mixture model-based clustering and hierarchical clustering. *Applied energy*, *231*, 331–342.

Li, S., & Wen, J. (2014). A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform. *Energy and Buildings*, *68*, 63–71.

Lindén, A.-L., Carlsson-Kanyama, A., & Eriksson, B. (2006). Efficient and inefficient aspects of residential energy behaviour: What are the policy instruments for change? *Energy policy*, *34*, 1918–1927.

Liu, X., Liu, Z., Wang, G., Cai, Z., & Zhang, H. (2017). Ensemble transfer learning algorithm. *IEEE Access*, *6*, 2389–2396.

Lu, Q., Xie, X., Parlikad, A. K., & Schooling, J. M. (2020). Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance. *Automation in Construction*, *118*, 103277.

Maddalena, E. T., Lian, Y., & Jones, C. N. (2020). Data-driven methods for building control-a review and promising future directions. *Control Engineering Practice*, *95*, 104211.

Maier, T., Krzaczek, M., & Tejchman, J. (2009). Comparison of physical performances of the ventilation systems in low-energy residential houses. *Energy and Buildings*, *41*, 337–353.

May-Ostendorp, P. T., Henze, G. P., Rajagopalan, B., & Corbin, C. D. (2013). Extraction of supervisory building control rules from model predictive control of windows in a mixed mode building. *Journal of Building Performance Simulation*, *6*, 199–219.

Megahed, N. A., & Ghoneim, E. M. (2020). Antivirus-built environment: Lessons learned from Covid-19 pandemic. *Sustainable Cities and Society*, (p. 102350).

Miller, C., Nagy, Z., & Schlueter, A. (2018). A review of unsupervised statistical learning and visual analytics techniques applied to performance analysis of non-residential buildings. *Renewable and Sustainable Energy Reviews*, *81*, 1365–1377.

Mocanu, E., Nguyen, P. H., Kling, W. L., & Gibescu, M. (2016). Unsupervised energy prediction in a smart grid context using reinforcement cross-building transfer learning. *Energy and Buildings*, *116*, 646–655.

Molina-Solana, M., Ros, M., Ruiz, M. D., Gómez-Romero, J., & Martín-Bautista, M. J. (2017). Data science for building energy management: A review. *Renewable and Sustainable Energy Reviews*, *70*, 598–609.

Murray, D., Stankovic, L., & Stankovic, V. (2017). An electrical load measurements dataset of united kingdom households from a two-year longitudinal study. *Scientific data*, *4*, 1–12.

National Infrastructure Commission (2017). Data for the public good. `https://www.nic.org.uk/publications/data-public-good/`.

Niu, F., O'Neill, Z., & O'Neill, C. (2018). Data-driven based estimation of HVAC energy consumption using an improved fourier series decomposition in buildings. *Building Simulation*, *11*, 633–645.

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

Pallonetto, F., De Rosa, M., Milano, F., & Finn, D. P. (2019). Demand response algorithms for smart-grid ready residential buildings using machine learning models. *Applied Energy*, *239*, 1265–1282.

Pan, S., Xiong, Y., Han, Y., Zhang, X., Xia, L., Wei, S., Wu, J., & Han, M. (2018). A study on influential factors of occupant window-opening behavior in an office building in china. *Building and environment*, *133*, 41–50.

Panait, L., & Luke, S. (2005). Cooperative multi-agent learning: The state of the art. *Autonomous agents and multi-agent systems*, *11*, 387–434.

Papakostas, K., & Sotiropoulos, B. (1997). Occupational and energy behaviour patterns in greek residences. *Energy and Buildings*, *26*, 207–213.

Peng, C., Yan, D., Wu, R., Wang, C., Zhou, X., & Jiang, Y. (2012). Quantitative description and simulation of human behavior in residential buildings. *Building simulation*, *5*, 85–94.

Qiu, S., Li, Z., Li, Z., Li, J., Long, S., & Li, X. (2020). Model-free control method based on reinforcement learning for building cooling water systems: validation by measured data-based simulation. *Energy and Buildings*, (p. 110055).

Qolomany, B., Al-Fuqaha, A., Gupta, A., Benhaddou, D., Alwajidi, S., Qadir, J., & Fong, A. C. (2019). Leveraging machine learning and big data for smart buildings: A comprehensive survey. *IEEE Access*, *7*, 90316–90356.

Rahman, M. A., Zaman, N., Asyhari, A. T., Al-Turjman, F., Bhuiyan, M. Z. A., & Zolkipli, M. (2020). Data-driven dynamic clustering framework for mitigating the adverse economic impact of Covid-19 lockdown practices. *Sustainable Cities and Society*, *62*, 102372.

Ramsetty, A., & Adams, C. (2020). Impact of the digital divide in the age of COVID-19. *Journal of the American Medical Informatics Association*, *27*, 1147–1148.

Ribeiro, M., Grolinger, K., ElYamany, H. F., Higashino, W. A., & Capretz, M. A. (2018). Transfer learning with seasonal and trend adjustment for cross-building energy forecasting. *Energy and Buildings*, *165*, 352–363.

Roselyn, J. P., Uthra, R. A., Raj, A., Devaraj, D., Bharadwaj, P., & Kaki, S. V. D. K. (2019). Development and implementation of novel sensor fusion algorithm for occupancy detection and automation in energy efficient buildings. *Sustainable Cities and Society*, *44*, 85–98.

Santin, O. G., Itard, L., & Visscher, H. (2009). The effect of occupancy and building characteristics on energy use for space and water heating in dutch residential stock. *Energy and buildings*, *41*, 1223–1232.

Schooling, J., Burgess, G., & Enzer, M. (2020). Flourishing systems: Re-envisioning infrastructure as a platform for human flourishing.

Schweiker, M., Carlucci, S., Andersen, R. K., Dong, B., & O'Brien, W. (2018). Occupancy and occupants' actions. In *Exploring Occupant Behavior in Buildings* (pp. 7–38). Springer.

Scott, S. L., & Varian, H. R. (2014). Predicting the present with bayesian structural time series. *International Journal of Mathematical Modelling and Numerical Optimisation*, *5*, 4–23.

Stazi, F., Naspi, F., & D'Orazio, M. (2017). A literature review on driving factors and contextual events influencing occupants' behaviours in buildings. *Building and Environment*, *118*, 40–66.

Stewart, R., Urban, M., Duchscherer, S., Kaufman, J., Morton, A., Thakur, G., Piburn, J., & Moehl, J. (2016). A bayesian machine learning model for estimating building occupancy from open source data. *Natural Hazards*, *81*, 1929–1956.

Sun, C., & Zhai, Z. (2020). The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission. *Sustainable cities and society*, *62*, 102390.

Tian, W., Yang, S., Li, Z., Wei, S., Pan, W., & Liu, Y. (2016). Identifying informative energy data in bayesian calibration of building energy models. *Energy and Buildings*, *119*, 363–376.

Touzani, S., Ravache, B., Crowe, E., & Granderson, J. (2019). Statistical change detection of building energy consumption: Applications to savings estimation. *Energy and Buildings*, *185*, 123–136.

Verma, A., Prakash, S., Srivastava, V., Kumar, A., & Mukhopadhyay, S. C. (2019). Sensing, controlling, and IoT infrastructure in smart building: a review. *IEEE Sensors Journal*, *19*, 9036–9046.

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

Xu, J., Lu, W., Xue, F., & Chen, K. (2019). 'Cognitive facility management': Definition, system architecture, and example scenario. *Automation in Construction*, *107*, 102922.

Yang, C., Shen, W., Chen, Q., & Gunay, B. (2018). A practical solution for HVAC prognostics: Failure mode and effects analysis in building maintenance. *Journal of Building Engineering*, *15*, 26–32.

Yildiz, B., Bilbao, J. I., & Sproul, A. B. (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.

Yu, Z., Haghighat, F., Fung, B. C., & Yoshino, H. (2010). A decision tree method for building energy demand modeling. *Energy and Buildings*, *42*, 1637–1646.

Zhang, F., Deb, C., Lee, S. E., Yang, J., & Shah, K. W. (2016). Time series forecasting for building energy consumption using weighted support vector regression with differential evolution optimization technique. *Energy and Buildings*, *126*, 94–103.

Zhao, B., Ye, M., Stankovic, L., & Stankovic, V. (2020). Non-intrusive load disaggregation solutions for very low-rate smart meter data. *Applied Energy*, *268*, 114949.

Zhou, Y., Shi, Z., Shi, Z., Gao, Q., & Wu, L. (2019). Disaggregating power consumption of commercial buildings based on the finite mixture model. *Applied Energy*, *243*, 35–46.