energies

Review

A Meta-Analysis Review of Occupant Behaviour Models for Assessing Demand-Side Energy Consumption

Deepu Krishnan *, **Scott Kelly** and **Yohan Kim**

Institute for Sustainable Futures, University of Technology Sydney, Building 10, 235 Jones St, Ultimo, NSW 2007, Australia; scott.kelly@uts.edu.au (S.K.); yohan.kim@uts.edu.au (Y.K.)
* Correspondence: deepu.krishnan@student.uts.edu.au; Tel.: +61-451799334

Abstract: Occupant behaviour plays a significant role in shaping the dynamics of energy consumption in buildings, but the complex nature of occupant behaviour has hindered a deeper understanding of its influence. A meta-analysis was conducted on 65 published studies that used data-driven quantitative assessments to assess energy-related occupant behaviour using the Knowledge Discovery and Data Mining (KDD) framework. Hierarchical clustering was utilised to categorise different modelling techniques based on the intended outcomes of the model and the types of parameters used in various models. This study will assist researchers in selecting the most appropriate parameters and methods under various data constraints and research questions. The research revealed two distinct model categories being used to study occupant behaviour-driven energy consumption, namely (i) occupancy status models and (ii) energy-related behaviour models. Multiple studies have identified limitations on data collection and privacy concerns as constraints of modelling occupant behaviour in residential buildings. The "regression model" and its variants were found to be the preferred model types for research that models "energy-related behaviour", and "classification models" were found to be preferable for modelling "occupancy" status. There were only limited instances of data-driven studies that modelled occupant behaviour in low-income households, and there is a need to generate region-specific models to accurately model energy-related behaviour.

Keywords: occupant behaviour; occupancy; low-income households; hierarchical clustering; knowledge discovery and data mining; residential energy consumption

Citation: Krishnan, D.; Kelly, S.; Kim, Y. A Meta-Analysis Review of Occupant Behaviour Models for Assessing Demand-Side Energy Consumption. *Energies* **2022**, *15*, 1219. https://doi.org/10.3390/en15031219

Academic Editor: Hom Bahadur Rijal

Received: 6 December 2021
Accepted: 2 February 2022
Published: 7 February 2022

1. Introduction

Numerous studies have explored the behaviour of occupants within buildings and have drawn a direct connection between occupant behaviour and energy consumption. Most of these studies have focused on commercial building spaces. Only a few have focused on residential areas, and even fewer have focused on low-income and vulnerable households. Similarly, quantitative assessments of occupant behaviour and its interaction with the building features, which also influence energy conservation measures for low-income households, are rare. However, there are numerous qualitative research and intervention studies that have analysed the impact of variations in occupant behaviour on energy consumption.

Understanding the interactions between multiple parameters and more accurately predicting occupant behaviour will provide new opportunities for promoting energy conservation and improving energy efficiency. Strategies can include adapting building design in response to the effects of a specific occupant behaviour or developing technical or design solutions that influence or modify specific human behaviour. Such approaches can also be used to optimise energy consumption without compromising the indoor environmental quality and comfort of occupants.

This article explores the latest statistical modelling techniques used to quantify and understand the impacts of human interactions that influence energy consumption. Research focusing on dwellings, particularly those of low-income households, is also explored. A meta-analysis is used to identify and categorise modelling techniques used in studies. In addition, the outcomes of various statistical models and different interpretations of these outcomes are explored in detail.

The meta-analysis considers studies of residential buildings, commercial buildings, and institutional spaces. We compare studies of commercial and institutional spaces against studies of residential buildings. The inclusion of non-residential spaces helps with the exploration of different modelling techniques that are commonly used to study occupant behaviour. This literature review can be used as a reference for selecting suitable variables and specifying the types of outcomes (behaviour or occupancy modelling) researchers intend to obtain or for choosing the most appropriate modelling technique for undertaking research on occupant behaviour. In addition, the systematic approach used in this meta-analysis can be used as a framework for assimilating and organising relevant information from different studies to select suitable modelling techniques and data points for specific experiments.

1.1. Background

The factors that contribute to energy use in buildings can be broadly grouped into two categories: external factors and internal factors. Jia et al. [1] identified the building envelope, building systems, equipment, and climate as external factors, and they identified occupant behaviour, operation/maintenance, and indoor environmental conditions as internal factors. Yan et al. [2] stated that the relationship between occupant behaviour and energy consumption is primarily governed by the occupants' pursuit of environmental comfort.

It is widely believed that occupant behaviour is one of the most difficult internal factors to model when it comes to predicting building energy consumption. That is, occupant behaviour is more complicated to assess and quantify than a building's envelope and thermal properties [3]. Variables that influence occupancy and occupant interaction with building features can be classified as (i) environment-related variables, (ii) time-related variables, or (iii) other random variables that explore the psychology of occupants [4].

Buildings are often designed with the assumption that occupants are rational and well-informed about the purpose and intent of the building design [5]. Additionally, it is assumed that occupants will comply with the operational rules of the building. Clevenger and Haymaker [6] showed that uncertainty related to occupant behaviour can limit the accuracy of energy modelling and a variation by as high as 150%, as was observed for school buildings while using standard energy modelling software such as DOE–2. Similarly, in some cases, variations in occupant behaviour can account for as much as 100% of the variation in residential energy consumption [7,8]. This discrepancy between actual and modelled occupant behaviour is also attributed to the fact that many of these studies falsely consider human behaviour as deterministic. Studies emphasise there is a need for more comprehensive research that explores the relationship between behaviour and occupancy patterns [7,8]. Hence, the modelling constraints and accuracy associated with predicting occupant behaviour would benefit from further research.

Several studies have compared the effectiveness of deterministic and probabilistic models of occupant behaviour (deterministic models use physical considerations to predict an outcome, whereas a stochastic model probabilistically predicts an outcome). These studies point out that assumptions regarding occupant behaviour can lead to shortcomings in building design because there is a high likelihood that actual behaviour will deviate from assumed norms. Daniel et al. [9] showed that simulation outcomes from AccuRate (an Australian federal government-endorsed calculation engine) for occupied residential buildings significantly deviated from the actual internal conditions. The closest matches between predicted and actual outcomes were for unoccupied low-energy residential buildings. This indicates that occupant behaviour is not adequately factored into the AccuRate software, and this is likely the case for other models that use similar assumptions. Most building simulation tools use equations based on heat transfer and thermodynamic, with occupant behaviour factored in using numerical approximations of predictable and repeatable deterministic actions. Assumptions that occupant behaviours are driven by similar non-dynamic constructs have been recognised as the main limitations in most simulation tools [10,11].

The review presented by Shrestha et al. [12] on household energy-saving behaviour highlighted the importance of gender roles in energy-saving and management in households. The study further showed that occupant-specific variables such as gender, income, family composition, education, headship, age-group, habits, and other socio-economic factors significantly influence energy-saving behaviours. Furthermore, some earlier studies such as that by Schipper et al. [13] presented analyses of different family typologies and time-use schedules for residents and concluded that changes in lifestyle (duration and location of leisure time) can alter energy consumption patterns in households.

A significant reduction in energy use in residential buildings can be brought about by altering the behaviour of occupants. Hence, understanding how occupants interact with buildings is an important topic for research. For example, Pisello and Asdrubali [14] identified such measures related to behaviour change as "human-based energy retrofits", which are simple zero-cost actions that reduce energy consumption. The authors of the study claimed that energy savings as high as 239 kWh per person per annum were achieved for a village of green buildings in central Italy. Furthermore, qualitative studies of low-income households, such as those by Langevin et al. [15], Vassileva and Campillo [16], and Trombley and Halawa [17], have highlighted the importance of interventions for creating behaviour change and improving energy consumption patterns. Ouyang and Hokao [18] compared households that were trained in energy-efficient behaviour with untrained households and found that, on average, there was a potential for a more than 10% reduction in energy usage via behavioural change.

1.2. Existing Modelling Approaches and Constraints in Modelling Occupant Behaviour

Residential buildings, unlike commercial buildings, are characterised by a higher diversity of occupancy hours and activities leading to behavioural diversities, which increase the complexities of studying occupant behaviour [4]. In addition to these complexities, privacy issues and ethical concerns also act as a significant hindrance to obtaining quality data for accurate modelling. Many technologies used for data acquisition (e.g., image-based technologies, radio-based technologies, and human-in-loop methods) face restricted use for residential buildings due to privacy concerns [19]. In addition, constraint issues such as a lack of compatibility between different simulation software programs and limitations related to competencies in developing code and data-mining techniques result in a failure to deliver effective occupant behaviour models [20,21]. The fact that humans are emotional and sometimes irrational adds to the complexity of model development. These constraints significantly impact the selection and use of parameters and advanced data-driven modelling techniques.

The techniques used to model residential energy consumption are usually classified as being top–down or bottom–up. In the top–down approach, macroeconomic variables and aggregate estimates of energy consumption are used as input variables into energy models (or subsets of the residential sector). This approach starts at the top and works its way down rather than focusing on individual dwellings or consumers. The bottom–up approach, on the other hand, calculates the energy consumption of individual residential buildings or groups of households with similar attributes. These dwellings can be further aggregated to understand and model energy consumption at a higher level of the hierarchy. Swan and Ugursal [7] further classified bottom–up models into statistical and engineering models, with statistical models being further divided into regression, conditional demand analysis, and neural network models. Engineering models can be divided into population distribution, archetype, and sample models. The behavioural models discussed in this article use statistical modelling techniques that can improve the accuracy of the modelled occupant behaviour of engineering models, which predominantly treat human behaviour as deterministic.

Many review articles have explored various methods and modelling techniques used to examine occupant behaviour and its subsequent effects on energy consumption, and they highlight need for better occupant behaviour models [22]. However, occupant behaviour has mainly been studied with the intention to improve energy modelling for commercial and institutional buildings without exclusively focusing on the data-driven modelling involved in the research. The difficulty of monitoring occupant behaviour in residential buildings has resulted in a knowledge gap. This is important because aggregate energy consumption in residential buildings is significant [4]. This review article focuses on understanding the different data-driven quantitative modelling techniques that are currently being used to understand the impacts of occupant behaviour on energy consumption using a meta-analysis. This meta-analysis is tied back to the Knowledge Discovery and Data Mining framework.

2. Scope and Methodology

Studies applying a quantitative modelling technique published between 2008 and 2020 were targeted for this review. Simple keyword searches used across multiple academic search engines identified hundreds of papers. The keywords used in this search included various combinations and derivatives of words such as "occupant"," occupancy", "behaviour"," building", "residential", " dwelling", "low-income", "household", "energy", and "energy efficiency". This process was not able to differentiate between papers of a qualitative or quantitative nature (keyword search was conducted using search engines associated with Scopus/Science direct, ASCE Library, IEEE Explore, and Google Scholar). Each paper was then individually reviewed and placed on a shortlist if a quantitative technique had been applied in the paper. During this deeper review process, the citations within each paper were also checked for other undiscovered papers. A total of fifty-four papers, consisting of eighty relevant modelling techniques, were identified. Each modelling technique was entered into a database assessed against relevant criteria. The meta-analysis was then used in a hierarchical clustering process.

The hierarchical clustering included studies for both residential buildings and commercial and institutional spaces. Studies related to commercial or institutional spaces used a larger number of data-driven and advanced modelling techniques that could be replicated in residential buildings, provided sufficient data points were available. The inclusion of non-residential spaces helped explore a wider variety of modelling techniques that had been used to study occupant behaviour.

This review process organised and classified key aspects of different data-driven modelling techniques used for modelling occupant behaviour. This review was intended to identify the key areas of study, parameters, or data collection requirements and to identify data-driven modelling techniques used for quantitative occupant behaviour modelling. Researchers can use the results of this meta-analysis to classify their research objectives in a broader field of study to identify the data collection and modelling technique requirements.

The subsequent subsections explain the Knowledge Discovery and Data Mining framework and how this framework has been used to structure this meta-analysis.

2.1. *Knowledge Discovery and Data Mining Framework*

The meta-analysis was structured based on the Knowledge Discovery and Data Mining (KDD) framework shown in Figure 1 [23]. Giving due consideration to the prevailing constraints in the behaviour modelling of occupants and the key areas of the KDD framework, we focus on three main components that can improve the knowledge base and help researchers in following a structured approach towards defining the research objective of parameter and model selection.

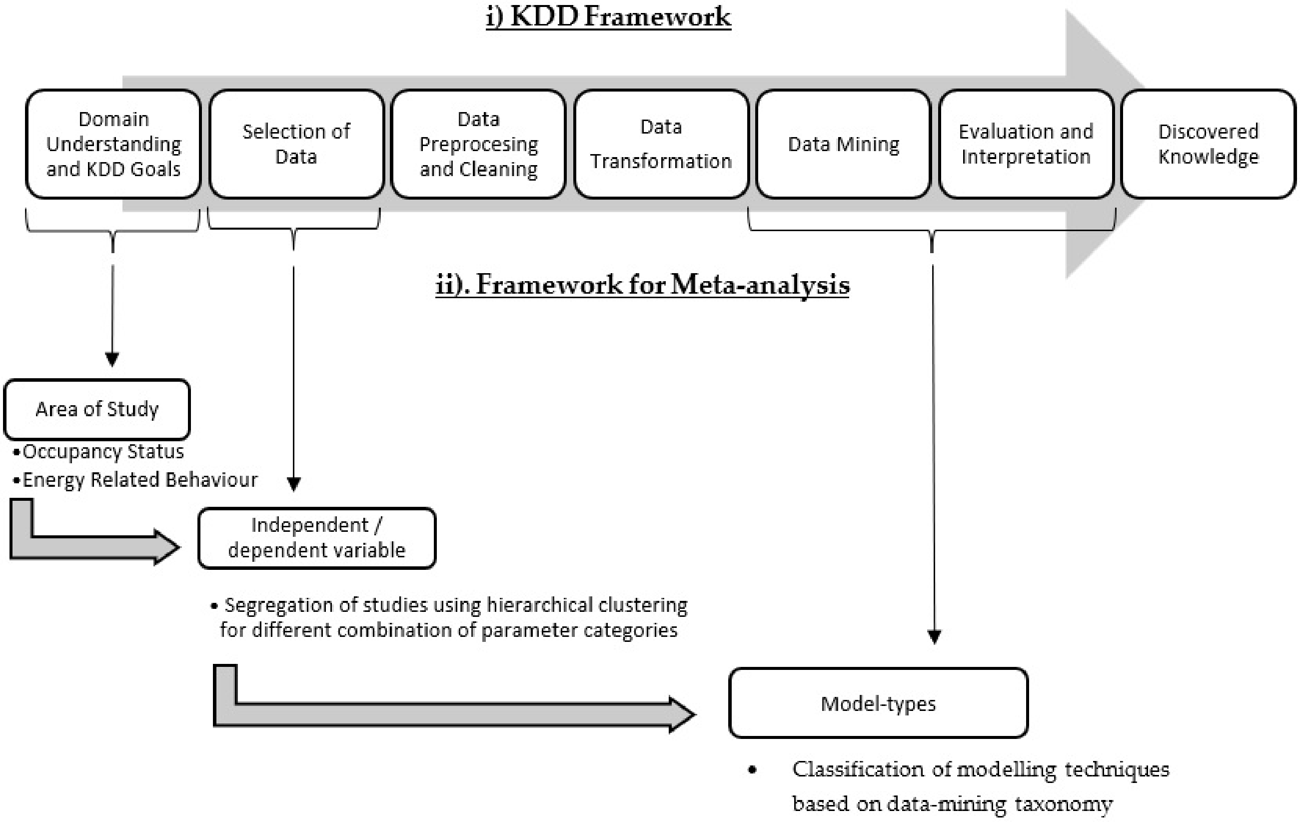

Figure 1. (**i**) Knowledge Discovery and Data-mining framework (**ii**) Framework for meta-analysis, which is derived from the KDD framework.

The KDD process starts by looking at the problem statement and environment to arrive at a clear definition of the end goals. However, in some cases, in line with the KDD process, the goals may become redefined as the process proceeds. The description of goals is followed by outlining the data required for building a knowledge base related to the given goal. The available data are evaluated, and provisions for attaining additional data are put in place. The selection of data is extremely crucial, as this forms the evidence base for constructing the model and governs the complexity of the next step, which is data pre-processing and cleaning. The pre-processing and cleaning are performed to enhance the reliability of the data. The data are then transformed, including with dimensionality reduction (including feature selection and record sampling) and attribute transformation (discretisation and functional transformation) [23].

Data transformation is influenced by the KDD goals, data-selection, and the following steps associated with modelling techniques: data mining, evaluation, and interpretation. Data mining involves the selection of the data-mining task (such as classification, regression, and clustering), the selection of the data-mining algorithm (such as a neural network or decision tree), and the implementation of the data-mining algorithm. The data-mining stage is followed by the evaluation of the model (such as assessing the accuracy of predictions) or the interpretation of relationships, patterns, and other results. The final step involves the use of domain knowledge or its incorporation into the system to address the problem statement or to achieve a specific goal [23].

2.2. Meta-Analysis of Studies Using KDD Framework

This section explains the criteria derived from the KDD framework that was used for the meta-analysis of modelling techniques used in different studies.

2.2.1. Classification of Modelling Techniques Based on the Area of Study

We used areas of study (KDD goals), the use of dependent/independent variable (selection of data), and type of model types (data mining, evaluation, and interpretation) to assimilate relevant information on the data-driven modelling of occupant behaviour (Figure 1). We identified two main KDD goals for studies on occupant behaviour based on the meta-analysis. These goals were "occupancy status" and "energy-related behaviour" and are referred to as the "area of study". These are further explained in the upcoming sections. Hence, the first level of classification for categorising modelling techniques for this meta-analysis was on the basis of the "field of study".

2.2.2. Categorisation of Modelling Techniques Based on Area of Study and Independent Variables

For the second level of classification, a sparse matrix was used to list different variables used for different modelling techniques. The compiled sparse matrix had 80 rows, one row for each of the modelling techniques, and 12 columns, one column for each parameter category (which included both independent variable and dependent variables for training the models), and the presence or absence of the parameters in a study was recorded as either 1 ("present") or 0 ("absent"). In addition to these 12 columns of parameter categories, the area of study (i.e., energy-related behaviour or occupancy status) was used as an input for categorising the 80 modelling techniques (Figure 2). That is, each modelling technique was considered as a data point that was identified by its variables and the area of study/outcome of the model. The clustering process, which was used for categorising different modelling techniques, is represented in Figure 2.

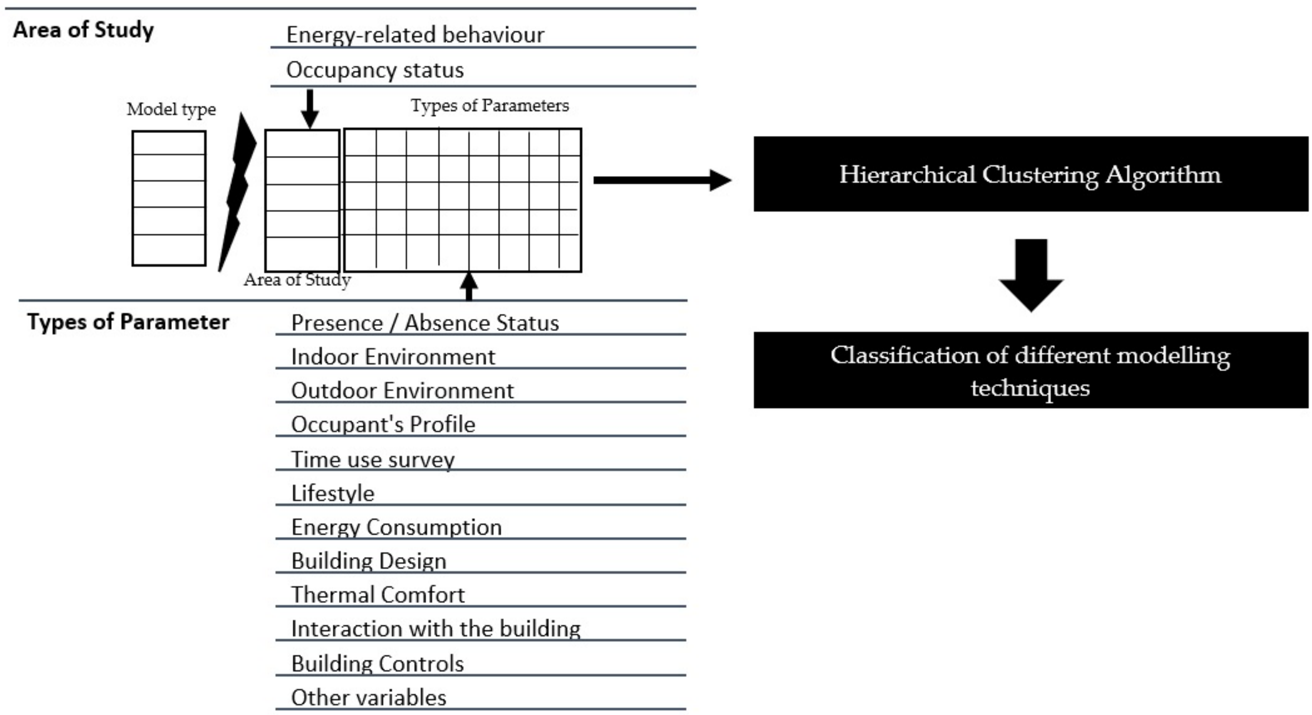

Figure 2. Clustering of different modelling techniques based on the "Area of Study" and "Types of Parameters.".

The result of clustering is represented as a dendrogram and was generated using Euclidean distance as the dissimilarity structure and Ward's method as the agglomeration method [24–27]. Ward's method states that the distance between two clusters, A and B, is the increase in the sum of squares as we merge them; that is, for two clusters A and B:

$$\nabla(A,B) = \frac{n_A n_B}{n_A + n_B}|m_A - m_B|^2 \tag{1}$$

where mj is the centre of the cluster j, nj is the number of points in it, and delta (∇) is the merging cost of combining clusters A and B. The agglomerative approach of clustering is a bottom–up approach that starts with each of the nodes as a single cluster. The algorithm groups the smaller clusters into a larger cluster and calculates the corresponding distances. The resulting clusters are shown using a tree dendrogram (Figure 3), which shows the Euclidean distance (y) between different clusters that are formed at various levels of the hierarchy.

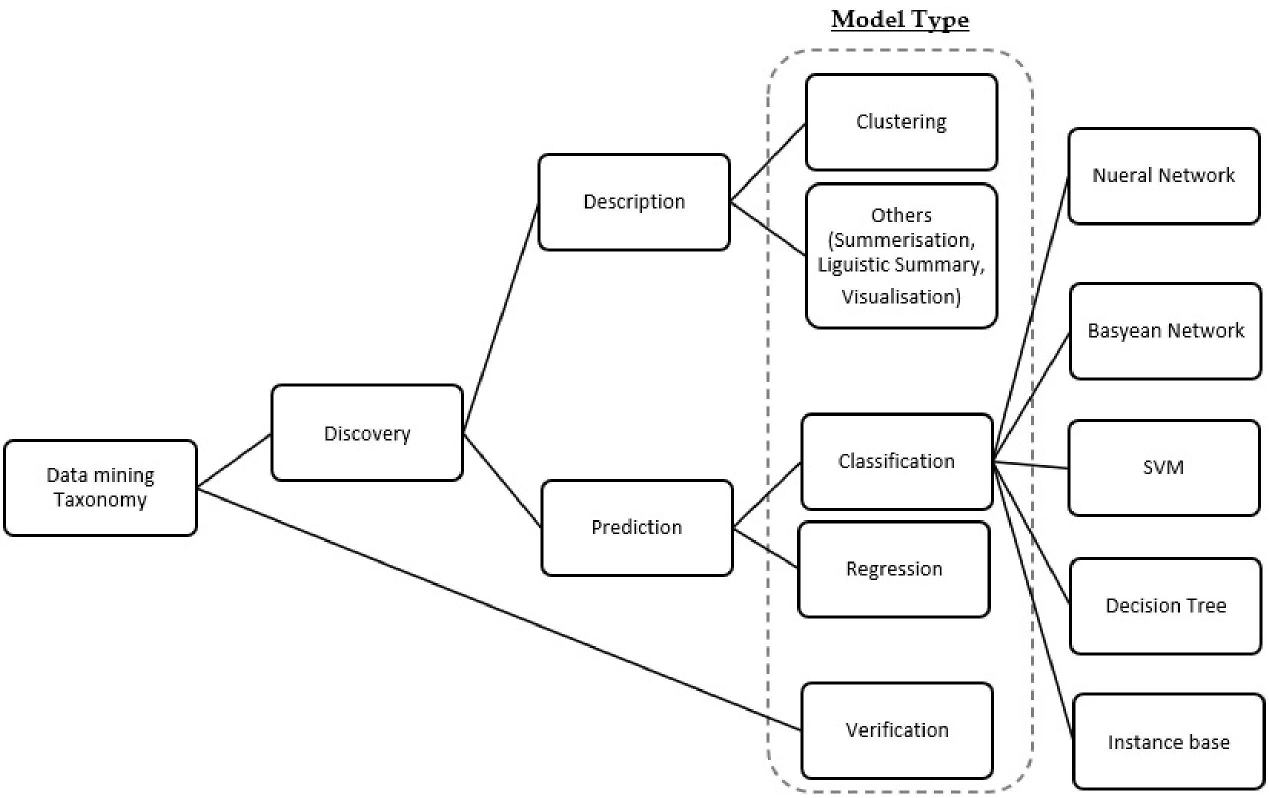

Figure 3. Data mining taxonomy and model type classification for segregating different modelling techniques.

2.2.3. Classification of Data-Driven Modelling Techniques

The data-mining taxonomy shown in Figure 3 has been used to classify the modelling techniques that were identified for this meta-analysis [23]. The data-mining methods can be grouped into either the verification-oriented approach where the system focuses on verifying a hypothesis or the discovery-oriented approach where the system finds new rules and patterns in the data. The verification-oriented approach comprises more traditional statistical methods such as the t-test and ANOVA, whereas, the discovery-oriented techniques are based on inductive learning and involve the construction of a model, either explicitly or implicitly, by generalising the observation from a training sample.

The discovery approach is further divided into predictive modelling and descriptive modelling, which are also known as supervised and unsupervised learning, respectively, in machine learning terminology. The descriptive approach assesses a sample without a target attribute or dependent variable. The supervised learning or predictive approach focuses on discovering and generalising the relationship between the input parameters (dependent variables) and target attribute (independent variables). The predictive approach is further divided into regression and classification models. The classification model based on the classifier, architecture, and algorithmic approach can be further categorised into other sub-categories (Figure 3). Hence, as the third level of classification for this meta-analysis, the modelling techniques in different studies were classified into five model type categories: clustering, classification, regression, verification, and others.

The insights from this meta-analysis of different modelling techniques will be helpful for classifying studies based on modelling outcomes (occupancy status or energy-related behaviour), the complexity of an experiment based on its data collection requirements, and the nature/complexity of modelling techniques. The comprehensive list of parameters presented in Appendix B can also help in evaluating the options for data collection for a given research structure or for better understanding the ethics implications for an experiment. For example, the options available for capturing the presence or absence of occupants can be evaluated and compared using this review to select the best-suited alternative based on the existing ethical considerations for a study.

This meta-analysis will add value to specific steps of the quantitative study of occupant behaviour. These steps are: (1) defining the research objective, (2) structuring the experiment and defining the scope for the data collection for the analysis, and (3) selecting the model type. In addition, the meta-analysis also presents a structured and scalable approach that can be used to identify, assimilate, and organise relevant information for evaluating different statistical modelling techniques. The algorithmic approach used in this meta-analysis (hierarchical clustering) makes it scalable for studying a larger number of modelling techniques.

3. Results and Discussions

This section summarises the insights from the meta-analysis approach described in the previous section. The analysis classified the 80 modelling techniques from 65 studies using three levels of classification: (1) field of study, (2) the use of dependent/independent variables, and (3) model type.

3.1. Defining the Area of Study: Difference between Occupancy Status and Energy-Related Behaviour

Variations in occupant behaviour are the result of two variables: occupancy and behaviour. The behaviour of an occupant relates to how they interact with the building in relation to energy consumption. This behaviour is referred to as "energy-related behaviour" in the upcoming sections. Energy-related behaviour mainly refers to the interaction of occupants with different features of the building, the habitual behaviours of the occupants, and personality traits or lifestyles that influence energy consumption in the household. Occupancy is defined as the presence and absence of occupants in a building over time [1]. Occupancy in a building is referred to as "occupancy status" in the upcoming sections. Occupancy status could be further extended to include the number of occupants, the physical distribution of occupants and other static information related to the occupants, e.g., age, ethnicity, or level of education.

We contend that occupancy status and energy-related behaviour need to be treated as separate entities, as this enables these two distinct concepts to be explored in greater detail. This is done because the methods that need to be applied for studying occupancy status and energy-related behaviour are very different. This difference is particularly evident when it comes to model types, data requirements, data acquisition, and the application of resultant models in simulation tools. To this end, we distinguished between these two concepts when reviewing and assessing the literature and the methods that were applied.

In residential buildings, occupancy status profile may not show frequent changes over time. A study of time-use data in Spain identified three peaks in the occupancy status of dwellings that coincided with morning, noon, and evening [28]. Additionally, occupancy levels for a dwelling usually remain static unless ownership changes [29]. However, changes in factors such as lifestyle, income, health, and comfort can drive occupancy status and trigger changes in behaviour that alter their interactions with features of a dwelling, resulting in shifts in energy consumption patterns.

Based on the discussion in this section, modelling techniques are classified based on the intended outcome of the model, which can either be 'occupancy status' or the 'energy-related behaviour' of the occupant (Figure 2). However, there have been instances where 'occupancy status' or 'energy-related behaviour' were explicitly modelled for achieving a separate research objective. For example, the primary objective of the study of Perez-Fargallo et al. [30] was to develop a thermal comfort model. This involved modelling energy-related behaviour in terms of occupant's preferences and acceptance of thermal conditions in low-income households. This study was included in this review and was classified as a research on energy-related behaviour. The results of the classification of studies based on the area of study are provided in Appendix A.

3.2. Identifying Data or Parameters Categories Required for Occupancy Status or Energy-Related Behaviour Models

This section explores different parameters used in modelling occupant behaviour and categorises them in generic groups based on the nature of information these parameters are adding to the model.

3.2.1. Classification of Data Used for Occupant Behaviour

A wide range of parameters are used in different modelling approaches; hence we grouped the parameters into different categories based on the purpose and nature of the data provided by them. For example, data from Wi-Fi connections, occupancy schedules, time-use surveys, video recordings, photographs, and motion sensors were grouped into a category called "presence and absence status" because these parameters were used to ascertain the presence of occupants during a specific time period or under particular conditions. The estimation of occupancy using CO_2 concentration or other models (such as DeST (Designer's Simulation Toolkit)) were not included in "presence and absence status" because they are not direct measurements. All the parameters and their respective parameter categories for the 54 reviewed articles are provided in Appendix B.

We identified 12 parameter categories that could be used to summarise the data used in research focused on modelling occupancy status or energy-related behaviour. These categories are: (1) presence and absence status, (2) indoor environment, (3) outdoor environment, (4) occupant's profile, (5) time-use survey, (6) lifestyle, (7) energy consumption, (8) building design, (9) thermal comfort, (10) interaction with the building, and (11) building controls. The data that did not fall in these categories were placed into a separate 12th category called "other parameters". A list of the input data that were grouped under each of the above-mentioned categories is shown in Appendix B.

The "indoor and outdoor environment" parameter categories include temperature, relative humidity, air velocity, air pressure, CO_2 concentration, and lighting conditions. The "occupant's profile" category captures details regarding the occupant, such as age and gender; in some cases, these represent consolidated demographic data for a cohort. More information about the occupant is captured using the "time-use survey" data, which usually detail routine activities done by an occupant. Similarly, parameters under the "lifestyle" category capture details such as perceived behavioural control, attitude towards energy conservation, degree of physical activity, frugality, family type, and bill consciousness. "Energy consumption" is another important parameter category that captures details related to gas-consumption readings and electricity-consumption data, which are consolidated values or (in some cases) appliance-specific.

"Interaction with building features" is an important parameter used in studies that have modelled "energy-related behaviour" and covers data such as the opening/closing of the windows, the use of lighting systems, the use of fans, and preferred thermostat settings. The data under this category were found to mainly be collected through surveys or by passive data collection techniques using sensors and audio-visual tools. "Building controls" capture details related to automated controls or pre-set schedules in a building. These could be data related to plug load controls or automated or pre-set schedules for equipment such as washing machines, water heaters, HVAC systems, and lighting systems. Data related to "thermal comfort" comprise a separate category. Data under this category were found to usually be collected through surveys that cover the level of perceived thermal comfort by an occupant and their requirements for thermal comfort.

The parameter categories (Appendix B) have been used in different contexts, i.e., for training (either as dependent or independent variables) or interpreting a model. For example, the parameters in the "building design" category were used as independent variables in models for studying 'energy-related' behaviour" [31,32] and, in some cases, as a supplementary/explanatory data for case selection or for interpreting results. [30,33,34]. Similarly, some parameters in "presence and absent status" were used as dependent variables in studies on "occupancy status" [35]; it was also used as an independent variable while modelling "energy-related behaviour" [31,32].

3.2.2. Parameters Use in Modelling Occupancy Status and Energy-Related Behaviour

Figures 4 and 5 summarise the meta-analysis of modelling techniques based on the independent/dependent variables and the area of study. Figure 4 is a dendrogram that shows the level hierarchy of the clusters identified by the clustering algorithm. Each of the numbers at the lowest level of the dendrogram (Figure 4) represents the individual modelling technique, and the number denotes the identification assigned to each of the modelling techniques during the clustering process. The parameters used by modelling techniques in clusters C and D are related to the internal environment and the status of occupancy (presence/absence) (Figure 4). Clusters A and B contain modelling techniques that use a broader range of parameters and have predominantly been used for modelling energy-related behaviour.

The techniques in cluster B frequently use parameters that explore the outdoor/indoor environment, presence or absence status, and the interaction of an occupant with the building features, hence focusing on more dynamic and measured parameters. The models in cluster A more frequently use static parameters compared to the models in cluster B. The use of parameters related to occupant's profile (e.g., age and occupation), lifestyle-related data, and building design are more frequently seen in cluster A. The modelling techniques used in clusters C and D generate models that predict or study the occupancy status profile (Figure 5). Techniques in cluster C use the internal environment to predict occupancy status (presence/absence), whereas the techniques in cluster D map presence/absence or time-use patterns against timestamps. The colour of bars represents one of the two modelling outcomes, namely 'occupancy status (blue)' and 'energy-related behaviour (red)'. The y-axis of the bar charts in Figure 5 is the parameter categories representing different variables listed in Appendix B. The x-axis shows "frequency of use", representing the number of modelling techniques (within a cluster) that use variables in that parameter category.

Techniques for modelling energy-related behaviour use a wider variety of parameters. However, the types of variables used for behaviour modelling depend upon the type of behaviour that is being modelled. For example, the modelling of behaviour in the context of energy use would require variables related to energy consumption, occupant profile, and data from a time-use survey [3]. Modelling approaches aimed at modelling other energy-related behaviours, such as understanding the agreeableness to behaviour change and the impact of normative feedback, would require total energy consumption and occupant profiles [36]. Studies that target intentions for energy conservation would focus more on certain variables that are listed under the "lifestyle (behaviour-specific data)" parameter category in Appendix B [37,38]. Additionally, if interaction with a building feature is the outcome, then the variables used would mainly be in the categories of "internal environment", "external environment", and "interaction with building" (also refer to Appendix B) [39,40]. Furthermore, models for studying energy-related behaviour are more complicated than models for studying occupancy status. However, the computational power required for both the outcomes also depends on the data points and the size of the cohort. Furthermore, it is interesting to note that none of these studies considered co-benefits such as the health and social benefits resulting from an energy-related behaviour.

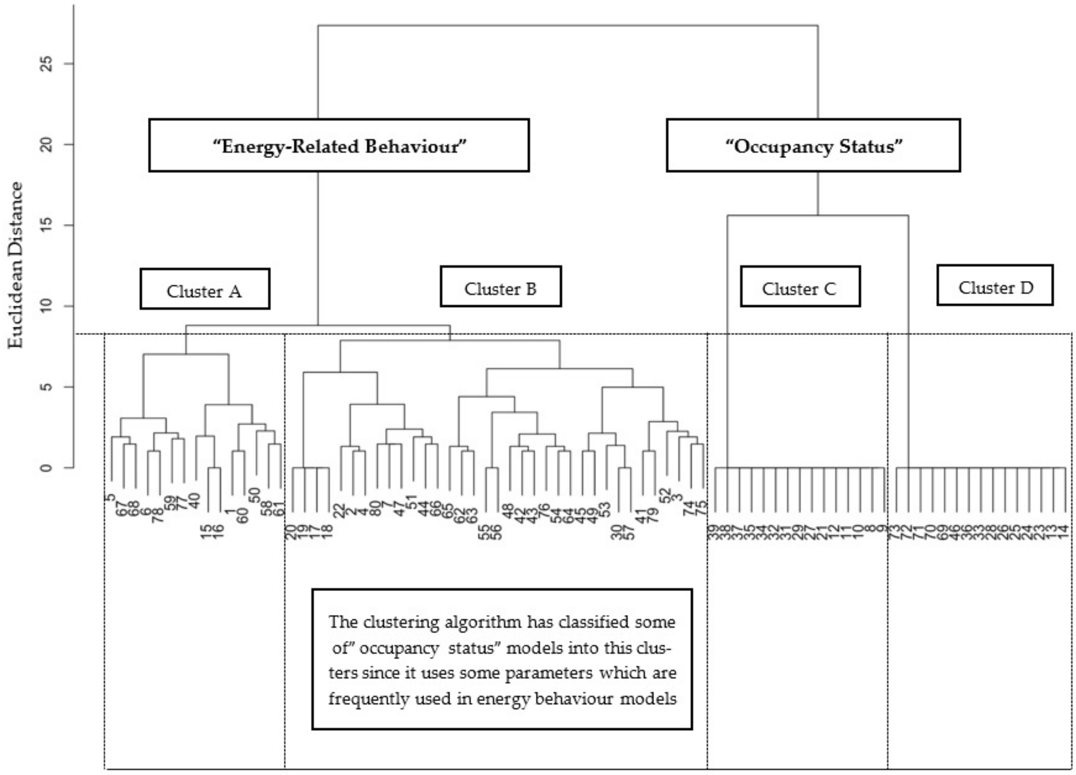

Figure 4. Dendrogram showing the classification of modelling techniques based on the parameters used in models.

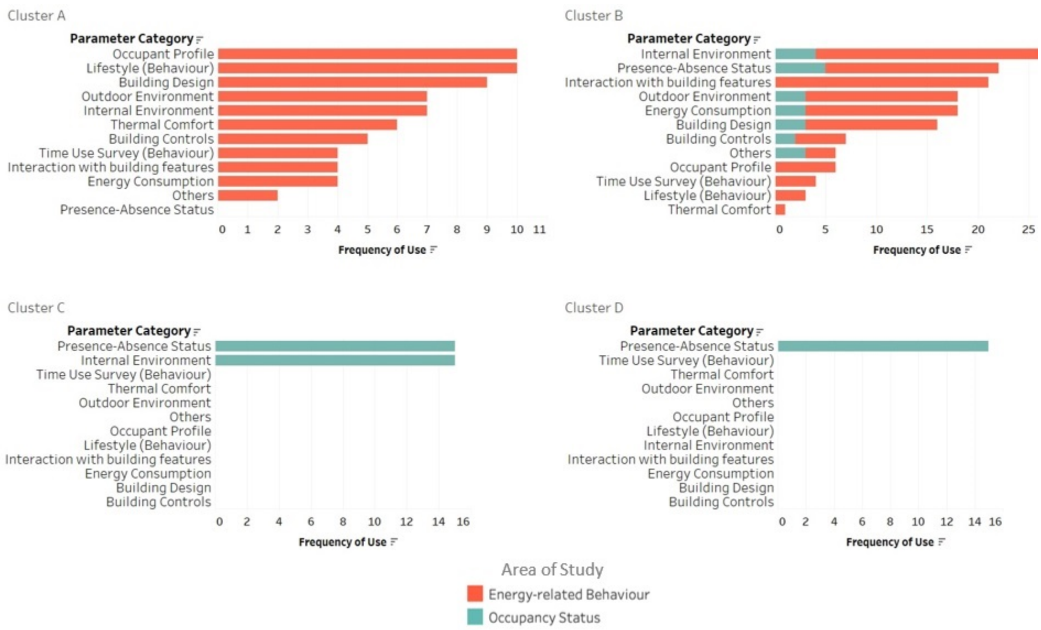

Figure 5. List of parameters used in different clusters along with their frequency of use.

3.2.3. Importance of Developing Region-Specific Models

The nature of the parameters that are used for behaviour modelling can vary from one geographical region to another. Figure 4 shows that the main parameter groups that are used in behaviour modelling are "interaction with the building features", "weather data (outdoor environment)", "occupant's profile", "energy consumption", "building design", and "behaviour data related to lifestyle". These parameters are usually specific to a region, and as such, it is evident that geographic location and demographic features affect behaviour models. Additionally, there have independent studies such as that by Lenzen [41] that have corroborated the claim that the resource endowments of a region, historical events, socio-cultural norms, behavioural attributes, and current market conditions determine the household energy consumption in a country. In addition to this, energy tariff structures, policies in the power sector, and the penetration of rooftop photovoltaic solutions can directly or indirectly impact energy consumption in the buildings of a particular country or a state [42,43]. Hence, we can infer that for a country or region, indigenous research into occupant behaviour is crucial. Furthermore, due to the varied climatic conditions in a geographically vast country, models need to be region-specific to generate an accurate prediction of occupant behaviour for energy modelling. To ensure the accuracy of models over time, these region-specific models need to be continuously updated to capture the impacts of changes in policies and the introduction of new technologies.

3.3. Types of Modelling Techniques

Regression is the most preferred model type used in categories A and B, followed by model types such as classification and clustering. These categories (A and B) mainly include research that model energy-related behaviour. The studies in cluster C and D model occupancy status, and classification is the most frequently used model type.

The modelling techniques utilised in all studies included in this meta-analysis included model types that use a discovery approach (as per data-mining taxonomy). Furthermore, the majority of these studies used predictive modelling using regression or classification model types, with a few studies also following a descriptive approach using clustering/other model types. Additionally, it was observed that most of the recent studies, especially the ones exploring occupancy status in commercial spaces (mainly in category C and D), used the neural networks and other advanced supervised learning approaches (classification model type). For studies focusing on dwellings, instances of advance supervised learning could be seen in experiments (mainly in category B) that explored the interaction of occupants with building features. However, the use of advanced unsupervised learning was found to be more prevalent in studies focusing on commercial and institutional buildings. The descriptive approach of data mining that uses methods such as clustering or other mathematical models was mainly found in exploratory studies on occupant behaviour.

Modelling techniques that are employed to study or model occupancy status and some energy-related behaviours use classification model types powered by machine learning algorithms. Among these classification model types, the K-nearest neighbours algorithm (KNN) [44–46], which uses the Euclidian distance of a specific data-point from the centroid of a group of data-points, has been one of the most frequently used models for predicting occupancy status through classification. Apart from KNN models, classification models based on support vector machine (SVMs) [45,47], which are highly efficient for non-linear classification (using kernel), have also been frequently used. Models that use the Markov chain [48–52] or artificial neural networks (ANNs) [45,47,49,50,53] are also common. Similarly, different variants of regression models [20,29,32–34,39,40,54–62] and combinations of regression models that incorporate concepts such as the Theory of Planned Behaviour [37,38] and other classification models including reinforcement learning (RL) [63–65] have typically been used for modelling energy-related behaviour. In addition, Markov models have also been used in multiple studies on energy-related behaviour. Models that neglect to incorporate occupancy status as part of their assessment are missing an important explanatory information for modelling energy consumption.

Most studies aiming to understand occupancy status were found to have monitored a particular building or space rather than examining the occupants themselves. On the other hand, studies aiming to understand occupant behaviour could be classified as a building-space monitoring study or have been conducted on a fixed number of occupants or cohorts. A fixed cohort can be a fixed number of households or a sample of occupants. Studies focused on interaction with specific building features have usually studied a fixed cohort in a living space or building.

Comparing the Accuracy of Modelling Techniques

Some studies on occupancy and energy-related behaviour in this review compared the accuracy of different models. RMSE (root mean square error) [45,47,49], MAE (mean average error) [45,49], MAPE (mean average percentage error) [45], and R-squared values [60] were mostly used to identify which models were more accurate. The accuracy of the modelling techniques varied in different studies. Classification models such as support vector machines were found to be more accurate than ANN and KNN models for modelling occupancy status using Wi-Fi data, whereas the ANN model showed the best performance when Wi-Fi data were combined with environment data [45]. Linear regression was found to be superior to SVM classification and support vector regression (SVR) in modelling occupant-group schedules in office buildings using appliance-specific power consumption data [64].

The Markov chain model was found to be superior to the ANN and SVR models for the short-term prediction of occupant numbers, whereas it was slightly less accurate for predicting the presence and absence of occupants for 15 min, 30 min, and 1 h forecasts [49]. At the same time, a nearest neighbour (NN) model using a customised distance function along with association rule outperformed the Markov chain model with respect to modelling occupant coordination and generalisation of behaviour patterns while using time-use data for households [66]. ANN and SVR showed higher accuracies in predicting occupancy counts using CO_2 concentration compared to dynamic physical models [47]. Another study by Candanedo [60] on low-energy buildings compared multiple linear regression, support vector machine (using radial kernel), random forest, and gradient boosting machine (GBM); compared to the other two model types, GBM and the random forest classification model showed improved RMSE and variance (R squared) for the predictions. It was also observed that specific studies combined multiple modelling techniques to increase the accuracy of their outcomes.

3.4. Studies on Residential Buildings

Most of the studies on energy-related behaviour in dwellings were recent (post−2013) and focused on areas of study related to temperature preference, time use, energy-saving practices, interaction with the building, and other actions that influence energy consumption. Furthermore, studies on residential buildings mainly used survey data to obtain estimations of occupancy status rather than using passive data collection tools. Additionally, some studies used existing models to estimate parameters such as energy use [29,38], occupancy [29,31], building features [29], and other specific details for modelling energy-related behaviour. When it comes to low-income households, there is considerable space for more detailed data-driven research that focuses on occupant behaviour.

Some behaviour models for residential buildings mainly used regression modelling and other descriptive statistical approaches. On the other hand, some studies combined statistical modelling with certain theoretical approaches to explore the role of behaviour in energy conservation. For example, the Theory of Planned Behaviour was used in combination with hierarchical regression in multiple instances to study practices and motivations pertaining to energy conservation [37,38].

In comparison to the behavioural modelling techniques used for institutional and commercial buildings, modern data-intensive methods such as deep learning, reinforcement learning, and machine learning have been less frequently used for research into residential buildings. While regression remains the most preferred model type, clustering and classification have also been used. It should be noted that this review considered the modelling techniques of a limited number of studies, so the frequency and the preferences concerning the use of a specific model cannot be inferred with absolute certainty. However, in general, a preference for regression model type was observed.

It has been noted that there were only a limited number of papers that explicitly explored energy-related behaviour or occupancy status in low-income households. Although many qualitative studies have focused on energy consumption-related behaviours in low-income households, there have been very few studies that adopted a quantitative approach. Hence, there is a need for data-driven models to design effective initiatives for optimising energy consumption in low-income households.

The energy-related behaviours of low-income and other vulnerable households may be different from the behaviours of typical households. The disposable income of a household plays a significant role in energy-related behaviour. The use of appliances in low-income households is different from their use in high-income households. Specifically, the use of appliances such as dishwashers, air conditioners, and microwaves is higher for wealthier households [67]. Cayla [68] showed that income has a vital impact on energy consumption. Low-income households exhibit capital constraints in the purchase of appliances and in their use; the research also revealed that income influences the use of space heating and thermal comfort inside dwellings.

Studies in the USA have related occupants' profiles, incomes and other vulnerability indices to health outcomes. These studies further stated that deaths due to exposure to extreme weather are more prominent in aged populations, with ethnicity, location, and income being other determining factors [69]. Additionally, it is important to note that the study of energy use and occupant behaviour in low-income households has been less explored because these households are difficult to reach and are typically not modernised for facilitating the retrieval of quality data for analysis (e.g., availability of smart meters to collect utility data) [70].

Residential energy consumption influences and, in some instances, drives the living standards of occupants. Occupancy status and energy-related behaviour can therefore be seen as the two main pillars of human behaviour that have significant impacts on the energy consumption of buildings. Specific sub-populations of society such as vulnerable households and low-income households are more impacted by changes in energy consumption patterns since household income is a determining factor in energy-related expenditure [71]. Furthermore, James and Ambrose [72] observed that in low-income households, an approach that focused on both building retrofits and behavioural change resulted in a reduction of energy consumption by 18% as opposed to a much lower 11.6% reduction when only building retrofits were performed. Accurate energy modelling that focuses on understanding human behaviour and the factors behind improved energy consumption for vulnerable low-income households can pave the way for more tailor-made and practical approaches for optimising energy consumption. This can further add to the eradication of energy poverty, improvements in living standards, and the maximising of certain co-benefits such as thermal comfort.

4. Limitations

This meta-analysis specifically considered studies that involved data-driven modelling techniques to model or analyse occupant behaviour. There are less data-intensive and qualitative studies of occupant behaviour that approached this area from different dimensions, but these qualitive approaches were not covered in this meta-analysis and were beyond the scope of this review article.

While efforts were made to cover a variety of quantitative studies on occupant behaviour, the studies presented in this meta-analysis should not be seen as an exhaustive representation of all the available quantitative studies on this topic. The constraints around the keyword search mentioned in Section 2 comprise an important factor that feeds into this limitation.

5. Conclusions

There is an urgent need for more data-intensive occupant behaviour models for low-income households. Privacy issues related to data collection, as well as the ethical considerations related to these issues, need to be addressed to ensure that quality data are available for quantitative behaviour modelling. Other important findings are as follows:

- The review process identified two high-level research goals in studies exploring occupant behaviour: the modelling of "occupancy status" and "energy-related behaviour". Studies on occupancy status were found to deal with the presence and absence status of the occupant, whereas studies on energy-related behaviour were found to explore specific behavioural traits, lifestyles, and interactions of occupants with the building that influence energy consumption.
- A detailed list of different parameters or data that were used in modelling occupant behaviour is presented in Appendix B. These parameters were grouped into 12 categories. We used these parameter categories to further group modelling techniques into four separate categories. Category D uses time-series fluctuation in the occupant number to model the "occupancy status" in a building. Studies in category C predict the "occupancy status" based on the changes in the indoor environment of a building. The categories A and B mainly contain models that study "energy-related behaviour" and use a more significant number of parameters. The models in category B frequently use dynamic and measurable parameters, and the modelling techniques in category A frequently use static occupant-related parameters.
- The is a need for region-specific studies (e.g., for Australia) for developing customised behaviour models, as there are many parameters that vary depending on geography, demographics, and other macroeconomic factors.
- This study will assist in the selection of appropriate data-mining approaches and model types for studies on occupant behaviour based on the category-specific and goal-specific description of model types.

Furthermore, the applications of machine learning, reinforcement learning, ANN, and other advanced computational algorithms need further exploration in regard to their use in behaviour modelling in dwellings. These approaches can contribute to the development of more accurate, dynamic, and intuitive models that will enable researchers to better understand occupancy status and energy-related behaviour in dwellings.

Author Contributions: This manuscript was conceptualised and drafted by D.K. under the supervision of S.K. and Y.K. The manuscript was further reviewed and edited by the co-authors. All authors have read and agreed to the published version of the manuscript.

Funding: This study is a part of doctoral research supported by an Australian Government Research Training Program Scholarship and Research Training Program (RTP) Fees Offset Scholarship funded by the Australian Government. Ethical approval for this doctoral research was received from the University of Technology Sydney (UTS) Human Research Ethics Committee (Reference number: ETH18-3000).

Data Availability Statement: This study does not report or make scientific inferences based o datasets.

Acknowledgments: The study was conducted in tandem with the Victoria Healthy Home Project funded by Victoria Government. Ethical approval for the project was received from Victorian Department of Human Services (DHS) Human Research Ethics Committee (Reference number: 04/17), University of Technology (UTS) Sydney Human Research Ethics Committee (Reference number: ETH18-2273). This trial is registered with the Australian and New Zealand Clinical Trials

Registry (ACTRN 12618000160235). We acknowledge the support from the Victoria Healthy Homes Project Team* for developing this paper. * Who's in the Healthy Homes Team (VHHP Team)? CHERE (Centre for Health Economic Research and Evaluation): Rosalie Viney, Kees van Gool, Stephen Goodall, Katie Page, Patsy Kenny, Margaret Campbell, Phil Haywood, Brendan Mulhern, Lutfun Hossain, Thomas Longden, Dan Liu. ISF: Kerryn Wilmot, Yohan Kim, Scott Kelly, Jay Falletta, Deepu Krishnan. SV: Toby Cumming, Matthew Soeberg. AEF: Daniel Chadwick.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. List of Studies Used in the Meta-Analysis.

Reference	Title	Outcome/Area of Study
[36]	Personality Traits and Energy Conservation	Energy-related behaviour
[63]	A Stochastic Model of Integrating Occupant Behaviour into Energy Simulation with Respect to Actual Energy Consumption in High-Rise Apartment Buildings	Energy-related behaviour
[73]	Simulating the Human–Building Interaction: Development and Validation of an Agent-Based model of Office Occupant behaviours	Energy-related behaviour
[20]	Integrating Building Performance Simulation in Agent-Based Modelling Using Regression Surrogate Models: A Novel Human-In-The-Loop Energy Modelling Approach	Energy-related behaviour
[5]	Designing Buildings for Real Occupants: An Agent-Based Approach	Energy-related behaviour
[30]	Development of a New Adaptive Comfort Model for Low-Income Housing in the Central-South of Chile	Energy-related behaviour
[74]	Development of an Occupancy Prediction Model Using Indoor Environmental Data Based on Machine Learning Techniques	Occupancy status
[44]	Occupancy Determination Based on Time Series of CO_2 Concentration, Temperature and Relative Humidity	Occupancy status
[45]	Occupancy Prediction through Machine Learning and Data Fusion of Environmental Sensing and Wi-Fi Sensing in Buildings	Occupancy status
[35]	Modelling and Predicting Occupancy Profile in Office Space with a Wi-Fi Probe-based Dynamic Markov Time-Window Inference Approach	Occupancy status
[48]	A High-Resolution Domestic Building Occupancy Model for Energy Demand Simulations	Occupancy status
[66]	Accurate Household Occupant Behaviour Modelling Based on Data-Mining Techniques	Energy-related behaviour
[60]	Data-Driven Prediction Models of Energy Use of Appliances in a Low-Energy House	Energy-related behaviour
[75]	A Novel Feature Selection Framework with Hybrid Feature-Scaled Extreme Learning Machine (HFS-ELM) for Indoor Occupancy Estimation	Occupancy status
[76]	Application of Mobile Positioning Occupancy Data for Building Energy Simulation: An Engineering Case Study	Occupancy status
[49]	Short-Term Predictions of Occupancy in Commercial Buildings—Performance Analysis for Stochastic Models and Machine Learning Approaches	Occupancy status

Table A1. *Cont.*

Reference	Title	Outcome/Area of Study
[77]	Occupancy Estimation with Environmental Sensing via Non-Iterative LRF Feature Learning in Time and Frequency Domains	Occupancy status
[78]	Understanding Occupancy Pattern and Improving Building Energy Efficiency through Wi-Fi-Based Indoor Positioning	Occupancy status
[79]	Method for Room Occupancy Detection Based on Trajectory of Indoor Climate Sensor Data	Occupancy status
[61]	Analysis of Occupants' Behaviour Related to the Use of Windows in German Households	Energy-related behaviour
[80]	Indoor Occupancy Estimation from Carbon Dioxide Concentration	Occupancy status
[81]	Detection of Occupancy Profile Based on Carbon Dioxide Concentration Pattern Matching	Occupancy status
[50]	Occupancy Prediction through Markov-Based Feedback Recurrent Neural Network (M-FRNN) Algorithm with Wi-Fi Probe Technology	Occupancy status
[82]	Occupancy Estimation from Environmental Parameters Using Wrapper and Hybrid Feature Selection	Occupancy status
[51]	A Methodology Based on Hidden Markov Models for Occupancy Detection and a Case Study in a Low-Energy Residential Building	Occupancy status
[46]	Modelling Occupancy Distribution in Large Spaces with Multi-Feature Classification Algorithm	Occupancy status
[47]	Predicting Occupancy Counts Using Physical and Statistical CO_2-Based Modelling Methodologies	Occupancy status
[3]	Modelling Energy Consumption in Residential Buildings: A Bottom–Up Analysis Based on Occupant Behaviour Pattern Clustering and Stochastic Simulation	Energy-related behaviour
[83]	Extracting Typical Occupancy Data of Different Buildings from Mobile Positioning Data	Occupancy status
[84]	Modelling and Analysing Occupant Behaviours in Building Energy Analysis Using an Information Space Approach	Energy-related behaviour
[85]	Spatial-Temporal Event-Driven Modelling for Occupant Behaviour Studies Using Immersive Virtual Environments	Energy-related behaviour
[86]	A Simulation Approach to Estimate Energy Savings Potential of Occupant Behaviour Measures	Energy-related behaviour
[87]	Methodology for Detection of Occupant Actions in Residential Buildings Using Indoor Environment Monitoring Systems	Energy-related behaviour
[88]	Non-Intrusive Occupancy Monitoring for Energy Conservation in Commercial Buildings	Occupancy status
[64]	Occupant Behaviour and Schedule Modelling for Building Energy Simulation through Office Appliance Power Consumption Data Mining	Energy-related behaviour
[65]	LightLearn: An Adaptive and Occupant-Centred Controller for Lighting Based on Reinforcement Learning	Energy-related behaviour
[89]	Analysis of User Behaviour Profiles and Impact on The Indoor Environment in Social Housing of Mild Climate Countries	Energy-related behaviour
[90]	Inference of Thermal Preference Profiles for Personalized Thermal Environments with Actual Building Occupants	Energy-related behaviour
[91]	Occupant Behaviour in Building Energy Simulation: Towards a Fit-For-Purpose Modelling Strategy	Energy-related behaviour

Table A1. *Cont.*

Reference	Title	Outcome/Area of Study
[92]	Air-Conditioning Usage Conditional Probability Model for Residential Buildings	Energy-related behaviour
[33]	Window Opening Behaviour of Occupants in Residential Buildings in Beijing	Energy-related behaviour
[31]	A Preliminary Research on the Derivation of Typical Occupant Behaviour Based on Large-Scale Questionnaire Surveys	Energy-related behaviour
[39]	Comparison of Theoretical and Statistical Models of Air-Conditioning-Unit Usage Behaviour in a Residential Setting Under Japanese Climatic Conditions	Energy-related behaviour
[34]	Window Opening Behaviour Modelled from Measurements in Danish Dwellings	Energy-related behaviour
[93]	Verification of Stochastic Behavioural Models of occupants' Interactions with Windows in Residential Buildings	Energy-related behaviour
[94]	Clustering Household Energy-Saving Behaviours by Behavioural Attribute	Energy-related behaviour
[37]	Thermal Comfort or Money Saving? Exploring Intentions to Conserve Energy among Low-Income Households in the United States	Energy-related behaviour
[38]	How Do Socio-Demographic and Psychological Factors relate to Households' Direct and Indirect Energy Use and Savings?	Energy-related behaviour
[62]	Factors Influencing Energy-Saving Behaviour of Urban Households in Jiangsu Province	Energy-related behaviour
[32]	The Effect of Occupancy and Building Characteristics on Energy Use for Space and Water Heating in Dutch Residential Stock	Energy-related behaviour
[54]	On Uses of Energy in Buildings: Extracting Influencing Factors of Occupant Behaviour by Means of a Questionnaire Survey	Energy-related behaviour
[29]	Sensitivity Analysis of the Effect of Occupant Behaviour on the Energy Consumption of Passive House Dwellings	Energy-related behaviour
[95]	Behavioural Patterns and User Profiles Related to Energy Consumption for Heating	Energy-related behaviour
[28]	Analysis and Modelling of Active Occupancy of the Residential Sector in Spain: An Indicator of Residential Electricity Consumption	Occupancy status
[59]	Air-Conditioning Use Behaviours when Elevated Air Movement Is Available	Energy-related behaviour
[58]	Development of Integrated Occupant-Behavioural Stochastic Model Including the Fan Use in Japanese Dwellings	Energy-related behaviour
[96]	Linking Energy–Cyber–Physical Systems with Occupancy Prediction and Interpretation through Wi-Fi Probe-Based Ensemble Classification	Occupancy status
[97]	Carbon Dioxide-Based Occupancy Estimation Using Stochastic Differential Equations	Occupancy status
[52]	A Markov-Switching Model for Building Occupant Activity Estimation	Occupancy status
[98]	How Do Urban Residents Use Energy for Winter Heating at Home? A Large-Scale Survey in the Hot Summer and Cold Winter Climate Zone in the Yangtze River Region	Energy-related behaviour
[57]	Do Preferred Thermostat Settings Differ by Sex?	Energy-related behaviour
[56]	Contextualising Adaptive Comfort Behaviour within Low-Income Housing of Mumbai, India	Energy-related behaviour
[55]	Data-Driven Occupant Action Prediction to Achieve an Intelligent Building	Energy-related behaviour
[53]	Prediction of Occupancy Level and Energy Consumption in Office Building Using Blind System Identification and Neural Networks	Occupancy status
[99]	A Scalable Bluetooth Low Energy Approach to Identify Occupancy Patterns and Profiles in Office Spaces	Occupancy status

Appendix B

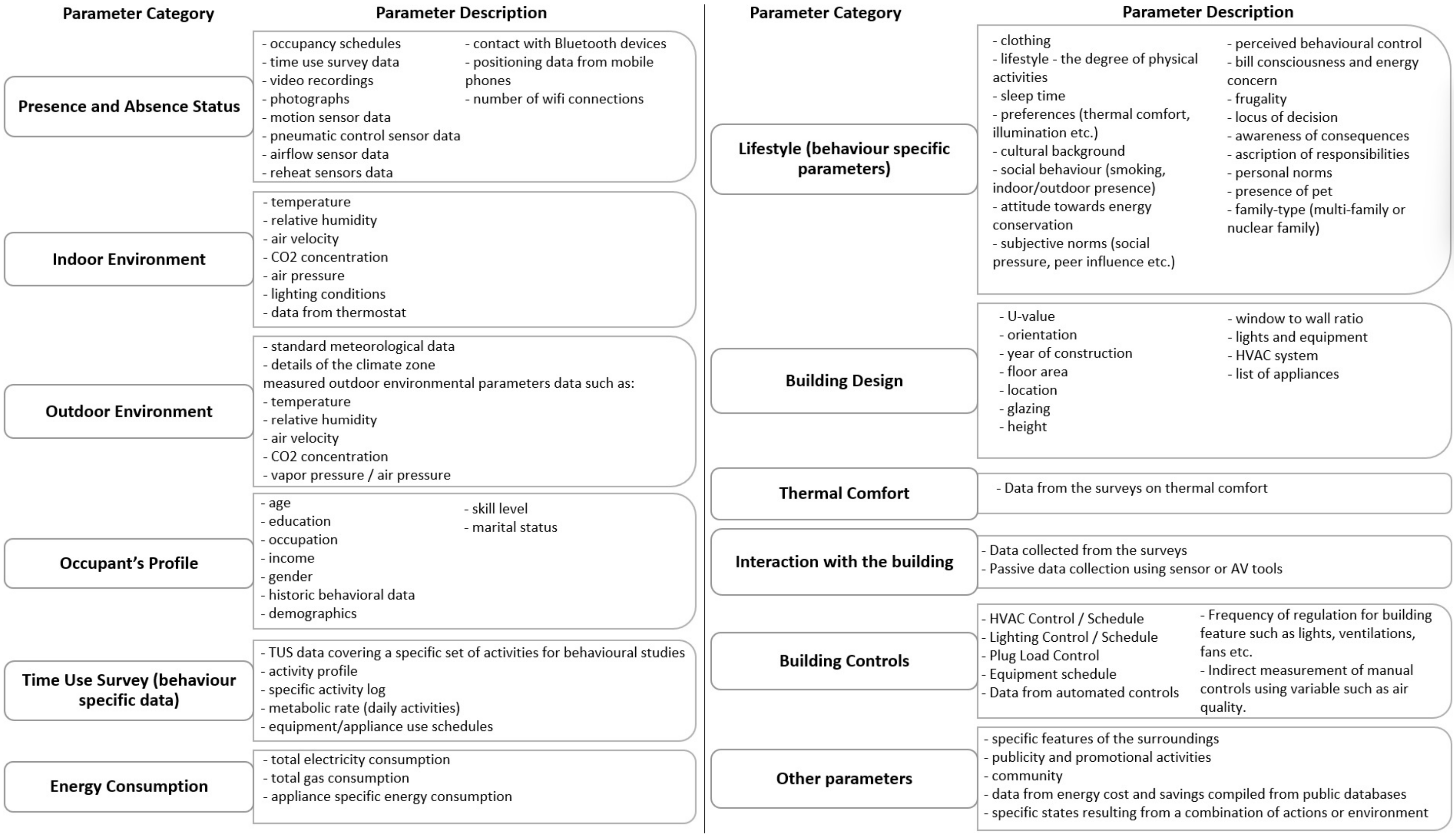

Figure A1. Categorisation of Parameters Used in Different Modelling Techniques.

References

1. Jia, M.; Srinivasan, R.S.; Raheem, A.A. From Occupancy to Occupant Behavior: An Analytical Survey of Data Acquisition Technologies, Modeling Methodologies and Simulation Coupling Mechanisms for Building Energy Efficiency. *Renew. Sustain. Energy Rev.* **2017**, *68*, 525–540. [CrossRef]
2. Yan, D.; Hong, T.; Dong, B.; Mahdavi, A.; D'Oca, S.; Gaetani, I.; Feng, X. IEA EBC Annex 66: Definition and Simulation of Occupant Behavior in Buildings. *Energy Build.* **2017**, *156*, 258–270. [CrossRef]
3. Diao, L.; Sun, Y.; Chen, Z.; Chen, J. Modeling Energy Consumption in Residential Buildings: A Bottom-up Analysis Based on Occupant Behavior Pattern Clustering and Stochastic Simulation. *Energy Build.* **2017**, *147*, 47–66. [CrossRef]
4. Balvedi, B.F.; Ghisi, E.; Lamberts, R. A Review of Occupant Behaviour in Residential Buildings. *Energy Build.* **2018**, *174*, 495–505. [CrossRef]
5. Andrews, C.J.; Yi, D.; Krogmann, U.; Senick, J.A.; Wener, R.E. Designing Buildings for Real Occupants: An Agent-Based Approach. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2011**, *41*, 1077–1091. [CrossRef]
6. Clevenger, C.M.; Haymaker, J. The Impact of the Building Occupant on Energy Modeling Simulations. *Jt. Int. Conf. Comput. Decis. Mak. Civ. Build. Eng.* **2006**, 1–10. [CrossRef]
7. Swan, L.G.; Ugursal, V.I. Modeling of End-Use Energy Consumption in the Residential Sector: A Review of Modeling Techniques. *Renew. Sustain. Energy Rev.* **2009**, *13*, 1819–1835. [CrossRef]
8. Seryak, J.; Kissock, K. Occupancy and Behavioral Affects on Residential Energy Use. In *Proceedings of the Solar Conference*; American Solar Energy Society: Boulder, CO, USA; American Institute of Architects: Washington, DC, USA, 2003; pp. 717–722.
9. Daniel, L.; Soebarto, V.; Williamson, T. House Energy Rating Schemes and Low Energy Dwellings: The Impact of Occupant Behaviours in Australia. *Energy Build.* **2015**, *88*, 34–44. [CrossRef]
10. Fabi, V.; Andersen, R.; Corgnati, S. Description of Occupant Behaviour in Building Energy Simulation: State-of-Art and Concepts for Improvements. *Proc. Build. Simul.* **2011**, 2882–2889. [CrossRef]
11. Goldstein, R.; Tessier, A.; Khan, A. Customizing the Behavior of Interacting Occupants Using Personas. In Proceedings of the Fourth National Conference of IBPSA-USA, New York, NY, USA, 11–13 August 2010; pp. 252–259.
12. Shrestha, B.; Tiwari, S.R.; Bajracharya, S.B.; Keitsch, M.M.; Rijal, H.B. Review on the Importance of Gender Perspective in Household Energy-Saving Behavior and Energy Transition for Sustainability. *Energies* **2021**, *14*, 7571. [CrossRef]
13. Schipper, L.; Bartlett, S.; Hawk, D.; Vine, E. Linking Life-Styles and Energy Use: A Matter of Time? *Annu. Rev. Energy* **1989**, *14*, 273–320. [CrossRef]
14. Pisello, A.L.; Asdrubali, F. Human-Based Energy Retrofits in Residential Buildings: A Cost-Effective Alternative to Traditional Physical Strategies. *Appl. Energy* **2014**, *133*, 224–235. [CrossRef]
15. Langevin, J.; Gurian, P.L.; Wen, J. Reducing Energy Consumption in Low Income Public Housing: Interviewing Residents about Energy Behaviors. *Appl. Energy* **2013**, *102*, 1358–1370. [CrossRef]
16. Vassileva, I.; Campillo, J. Increasing Energy Efficiency in Low-Income Households through Targeting Awareness and Behavioral Change. *Renew. Energy* **2014**, *67*, 59–63. [CrossRef]
17. Trombley, J.; Halawa, E. Can Further Energy Reductions Be Achieved through Behaviour Changes in Low Income Households? *Energy Procedia* **2017**, *121*, 230–237. [CrossRef]
18. Ouyang, J.; Hokao, K. Energy-Saving Potential by Improving Occupants' Behavior in Urban Residential Sector in Hangzhou City, China. *Energy Build.* **2009**, *41*, 711–720. [CrossRef]
19. Wagner, A.; Brien, W.O. *Exploring Occupant Behavior in Buildings*; Springer: Cham, Switzerland, 2018; ISBN 978-3-319-61463-2.
20. Papadopoulos, S.; Azar, E. Integrating Building Performance Simulation in Agent-Based Modeling Using Regression Surrogate Models: A Novel Human-in-the-Loop Energy Modeling Approach. *Energy Build.* **2016**, *128*, 214–223. [CrossRef]
21. Menassa, C.C.; Asce, A.M.; Kamat, V.R.; Asce, A.M.; Lee, S.; Asce, A.M.; Azar, E.; Asce, S.M.; Feng, C.; Asce, S.M.; et al. Conceptual Framework to Optimize Building Energy Consumption by Coupling Distributed Energy Simulation and Occupancy Models. *J. Comput. Civ. Eng.* **2014**, *28*, 50–62. [CrossRef]
22. Khahro, S.H.; Subramaniam, U.; Ghosh, A. Efficiency for Energy, Economy and Environment (3E) Benefit. *Energies* **2021**, *14*, 4487. [CrossRef]
23. Maimon, O.; Rokach, L. (Eds.) *Data Mining and Knowledge Discovery Handbook*; Springer: Cham, Switzerland, 2008; ISBN 9780387699349.
24. Szekely, G.J.; Rizzo, M.L. Hierarchical Clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method. *J. Classif.* **2005**, *22*, 151–183. [CrossRef]
25. Murtagh, F.; Legendre, P. Ward's Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm. *arXiv* **2011**, arXiv:1111.6285.
26. Kruskal, J.B.; Black, P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *J. Classif.* **2012**, *6*, 4–6. [CrossRef]
27. Distances between Clustering, Hierarchical Clustering. Available online: https://www.stat.cmu.edu/~{}cshalizi/350/lectures/08/lecture-08.pdf (accessed on 8 July 2020).
28. López-Rodríguez, M.A.; Santiago, I.; Trillo-Montero, D.; Torriti, J.; Moreno-Munoz, A. Analysis and Modeling of Active Occupancy of the Residential Sector in Spain: An Indicator of Residential Electricity Consumption. *Energy Policy* **2013**, *62*, 742–751. [CrossRef]

29. Blight, T.S.; Coley, D.A. Sensitivity Analysis of the Effect of Occupant Behaviour on the Energy Consumption of Passive House Dwellings. *Energy Build.* **2013**, *66*, 183–192. [CrossRef]
30. Pérez-Fargallo, A.; Pulido-Arcas, J.A.; Rubio-Bellido, C.; Trebilcock, M.; Piderit, B.; Attia, S. Development of a New Adaptive Comfort Model for Low Income Housing in the Central-South of Chile. *Energy Build.* **2018**, *178*, 94–106. [CrossRef]
31. Feng, X.; Yan, D.; Wang, C.; Sun, H. A Preliminary Research on the Derivation of Typical Occupant Behavior Based on Large-Scale Questionnaire Surveys. *Energy Build.* **2016**, *117*, 332–340. [CrossRef]
32. Guerra Santin, O.; Itard, L.; Visscher, H. The Effect of Occupancy and Building Characteristics on Energy Use for Space and Water Heating in Dutch Residential Stock. *Energy Build.* **2009**, *41*, 1223–1232. [CrossRef]
33. Yao, M.; Zhao, B. Window Opening Behavior of Occupants in Residential Buildings in Beijing. *Build. Environ.* **2017**, *124*, 441–449. [CrossRef]
34. Andersen, R.; Fabi, V.; Toftum, J.; Corgnati, S.P.; Olesen, B.W. Window Opening Behaviour Modelled from Measurements in Danish Dwellings. *Build. Environ.* **2013**, *69*, 101–113. [CrossRef]
35. Wang, W.; Chen, J.; Song, X. Modeling and Predicting Occupancy Profile in Office Space with a Wi-Fi Probe-Based Dynamic Markov Time-Window Inference Approach. *Build. Environ.* **2017**, *124*, 130–142. [CrossRef]
36. Shen, M.; Cui, Q.; Fu, L. Personality Traits and Energy Conservation. *Energy Policy* **2015**, *85*, 322–334. [CrossRef]
37. Chen, C.; Xu, X.; Day, J.K. Thermal Comfort or Money Saving? Exploring Intentions to Conserve Energy among Low-Income Households in the United States. *Energy Res. Soc. Sci.* **2017**, *26*, 61–71. [CrossRef]
38. Abrahamse, W.; Steg, L. How Do Socio-Demographic and Psychological Factors Relate to Households' Direct and Indirect Energy Use and Savings? *J. Econ. Psychol.* **2009**, *30*, 711–720. [CrossRef]
39. Schweiker, M.; Shukuya, M. Comparison of Theoretical and Statistical Models of Air-Conditioning-Unit Usage Behaviour in a Residential Setting under Japanese Climatic Conditions. *Build. Environ.* **2009**, *44*, 2137–2149. [CrossRef]
40. Jones, R.V.; Fuertes, A.; Gregori, E.; Giretti, A. Stochastic Behavioural Models of Occupants' Main Bedroom Window Operation for UK Residential Buildings. *Build. Environ.* **2017**, *118*, 144–158. [CrossRef]
41. Lenzen, M.; Wier, M.; Cohen, C.; Hayami, H.; Pachauri, S.; Schaeffer, R. A Comparative Multivariate Analysis of Household Energy Requirements in Australia, Brazil, Denmark, India and Japan. *Energy* **2006**, *31*, 181–207. [CrossRef]
42. Havas, L.; Ballweg, J.; Penna, C.; Race, D. Power to Change: Analysis of Household Participation in a Renewable Energy and Energy Efficiency Programme in Central Australia. *Energy Policy* **2015**, *87*, 325–333. [CrossRef]
43. Simshauser, P. From First Place to Last: The National Electricity Market's Policy-Induced "Energy Market Death Spiral". *Aust. Econ. Rev.* **2014**, *47*, 540–562. [CrossRef]
44. Szczurek, A.; Maciejewska, M.; Pietrucha, T. Occupancy Determination Based on Time Series of CO2concentration, Temperature and Relative Humidity. *Energy Build.* **2017**, *147*, 142–154. [CrossRef]
45. Wang, W.; Chen, J.; Hong, T. Occupancy Prediction through Machine Learning and Data Fusion of Environmental Sensing and Wi-Fi Sensing in Buildings. *Autom. Constr.* **2018**, *94*, 233–243. [CrossRef]
46. Wang, W.; Chen, J.; Hong, T. Modeling Occupancy Distribution in Large Spaces with Multi-Feature Classification Algorithm. *Build. Environ.* **2018**, *137*, 108–117. [CrossRef]
47. Zuraimi, M.S.; Pantazaras, A.; Chaturvedi, K.A.; Yang, J.J.; Tham, K.W.; Lee, S.E. Predicting Occupancy Counts Using Physical and Statistical Co2-Based Modeling Methodologies. *Build. Environ.* **2017**, *123*, 517–528. [CrossRef]
48. Richardson, I.; Thomson, M.; Infield, D. A High-Resolution Domestic Building Occupancy Model for Energy Demand Simulations. *Energy Build.* **2008**, *40*, 1560–1566. [CrossRef]
49. Li, Z.; Dong, B. Short Term Predictions of Occupancy in Commercial Buildings—Performance Analysis for Stochastic Models and Machine Learning Approaches. *Energy Build.* **2018**, *158*, 268–281. [CrossRef]
50. Wang, W.; Chen, J.; Hong, T.; Zhu, N. Occupancy Prediction through Markov Based Feedback Recurrent Neural Network (M-FRNN) Algorithm with WiFi Probe Technology. *Build. Environ.* **2018**, *138*, 160–170. [CrossRef]
51. Candanedo, L.M.; Feldheim, V.; Deramaix, D. A Methodology Based on Hidden Markov Models for Occupancy Detection and a Case Study in a Low Energy Residential Building. *Energy Build.* **2017**, *148*, 327–341. [CrossRef]
52. Wolf, S.; Møller, J.K.; Bitsch, M.A.; Krogstie, J.; Madsen, H. A Markov-Switching Model for Building Occupant Activity Estimation. *Energy Build.* **2019**, *183*, 672–683. [CrossRef]
53. Wei, Y.; Xia, L.; Pan, S.; Wu, J.; Zhang, X.; Han, M.; Zhang, W.; Xie, J.; Li, Q. Prediction of Occupancy Level and Energy Consumption in Office Building Using Blind System Identification and Neural Networks. *Appl. Energy* **2019**, *240*, 276–294. [CrossRef]
54. Rinaldi, A.; Schweiker, M.; Iannone, F. On Uses of Energy in Buildings: Extracting Influencing Factors of Occupant Behaviour by Means of a Questionnaire Survey. *Energy Build.* **2018**, *168*, 298–308. [CrossRef]
55. Pereira, P.F.; Ramos, N.M.M.; Simões, M.L. Data-Driven Occupant Actions Prediction to Achieve an Intelligent Building. *Build. Res. Inf.* **2020**, *48*, 485–500. [CrossRef]
56. Malik, J.; Bardhan, R.; Hong, T.; Piette, M.A. Contextualising Adaptive Comfort Behaviour within Low-Income Housing of Mumbai, India. *Build. Environ.* **2020**, *177*, 106877. [CrossRef]
57. Brewer, D. Do Preferred Thermostat Settings Differ by Sex? *Energy Build.* **2020**, *224*, 110140. [CrossRef]
58. Imagawa, H.; Rijal, H.B.; Shukuya, M. Development of Integrated Occupant-Behavioural Stochastic Model Including the Fan Use in Japanese Dwellings. *Energy Build.* **2020**, *226*, 110326. [CrossRef]

59. He, Y.; Li, N.; Zhang, H.; Han, Y.; Lu, J.; Zhou, L. Air-Conditioning Use Behaviors When Elevated Air Movement Is Available. *Energy Build.* **2020**, *225*, 110370. [CrossRef]
60. Candanedo, L.M.; Feldheim, V.; Deramaix, D. Data Driven Prediction Models of Energy Use of Appliances in a Low-Energy House. *Energy Build.* **2017**, *140*, 81–97. [CrossRef]
61. Calì, D.; Andersen, R.K.; Müller, D.; Olesen, B.W. Analysis of Occupants' Behavior Related to the Use of Windows in German Households. *Build. Environ.* **2016**, *103*, 54–69. [CrossRef]
62. Yue, T.; Long, R.; Chen, H. Factors Influencing Energy-Saving Behavior of Urban Households in Jiangsu Province. *Energy Policy* **2013**, *62*, 665–675. [CrossRef]
63. Jang, H.; Kang, J. A Stochastic Model of Integrating Occupant Behaviour into Energy Simulation with Respect to Actual Energy Consumption in High-Rise Apartment Buildings. *Energy Build.* **2016**, *121*, 205–216. [CrossRef]
64. Zhao, J.; Lasternas, B.; Lam, K.P.; Yun, R.; Loftness, V. Occupant Behavior and Schedule Modeling for Building Energy Simulation through Office Appliance Power Consumption Data Mining. *Energy Build.* **2014**, *82*, 341–355. [CrossRef]
65. Park, J.Y.; Dougherty, T.; Fritz, H.; Nagy, Z. LightLearn: An Adaptive and Occupant Centered Controller for Lighting Based on Reinforcement Learning. *Build. Environ.* **2018**, *147*, 397–414. [CrossRef]
66. Baptista, M.; Fang, A.; Prendinger, H.; Prada, R.; Yamaguchi, Y. Accurate Household Occupant Behavior Modeling Based on Data Mining Techniques. In Proceedings of the AAAI Conference on Artificial Intelligence, Quebec, QC, USA, 27–31 July 2014; pp. 1164–1170.
67. Matsumoto, S. How Do Household Characteristics Affect Appliance Usage? Application of Conditional Demand Analysis to Japanese Household Data. *Energy Policy* **2016**, *94*, 214–223. [CrossRef]
68. Cayla, J.M.; Maizi, N.; Marchand, C. The Role of Income in Energy Consumption Behaviour: Evidence from French Households Data. *Energy Policy* **2011**, *39*, 7874–7883. [CrossRef]
69. Reames, T.G. Targeting Energy Justice: Exploring Spatial, Racial/Ethnic and Socioeconomic Disparities in Urban Residential Heating Energy Efficiency. *Energy Policy* **2016**, *97*, 549–558. [CrossRef]
70. Dong, B.; Li, Z.; McFadden, G. An Investigation on Energy-Related Occupancy Behavior for Low-Income Residential Buildings. *Sci. Technol. Built Environ.* **2015**, *21*, 892–901. [CrossRef]
71. Reiss, P.C.; White, M.W. Household Electricity Demand, Revisited. *NBER Work. Pap. Ser. Househ.* **2001**, *43*, 1–56. Available online: https://www.nber.org/system/files/working_papers/w8687/w8687.pdf (accessed on 5 December 2021).
72. James, M.; Ambrose, M. Retrofit or Behaviour Change? Which Has the Greater Impact on Energy Consumption in Low Income Households? *Procedia Eng.* **2017**, *180*, 1558–1567. [CrossRef]
73. Langevin, J.; Wen, J.; Gurian, P.L. Simulating the Human-Building Interaction: Development and Validation of an Agent-Based Model of Office Occupant Behaviors. *Build. Environ.* **2015**, *88*, 27–45. [CrossRef]
74. Ryu, S.H.; Moon, H.J. Development of an Occupancy Prediction Model Using Indoor Environmental Data Based on Machine Learning Techniques. *Build. Environ.* **2016**, *107*, 1–9. [CrossRef]
75. Masood, M.K.; Jiang, C.; Soh, Y.C. A Novel Feature Selection Framework with Hybrid Feature-Scaled Extreme Learning Machine (HFS-ELM) for Indoor Occupancy Estimation. *Energy Build.* **2018**, *158*, 1139–1151. [CrossRef]
76. Pang, Z.; Xu, P.; O'Neill, Z.; Gu, J.; Qiu, S.; Lu, X.; Li, X. Application of Mobile Positioning Occupancy Data for Building Energy Simulation: An Engineering Case Study. *Build. Environ.* **2018**, *141*, 1–15. [CrossRef]
77. Zhu, Q.; Chen, Z.; Masood, M.K.; Soh, Y.C. Occupancy Estimation with Environmental Sensing via Non-Iterative LRF Feature Learning in Time and Frequency Domains. *Energy Build.* **2017**, *141*, 125–133. [CrossRef]
78. Wang, Y.; Shao, L. Understanding Occupancy Pattern and Improving Building Energy Efficiency through Wi-Fi Based Indoor Positioning. *Build. Environ.* **2017**, *114*, 106–117. [CrossRef]
79. Pedersen, T.H.; Nielsen, K.U.; Petersen, S. Method for Room Occupancy Detection Based on Trajectory of Indoor Climate Sensor Data. *Build. Environ.* **2017**, *115*, 147–156. [CrossRef]
80. Jiang, C.; Masood, M.K.; Soh, Y.C.; Li, H. Indoor Occupancy Estimation from Carbon Dioxide Concentration. *Energy Build.* **2016**, *131*, 132–141. [CrossRef]
81. Szczurek, A.; Maciejewska, M.; Wyłomańska, A.; Zimroz, R.; Żak, G.; Dolega, A. Detection of Occupancy Profile Based on Carbon Dioxide Concentration Pattern Matching. *Meas. J. Int. Meas. Confed.* **2016**, *93*, 265–271. [CrossRef]
82. Masood, M.K.; Soh, Y.C.; Jiang, C. Occupancy Estimation from Environmental Parameters Using Wrapper and Hybrid Feature Selection. *Appl. Soft Comput.* **2017**, *60*, 482–494. [CrossRef]
83. Jiefan, G.; Peng, X.; Zhihong, P.; Yongbao, C.; Ying, J.; Zhe, C. Extracting Typical Occupancy Data of Different Buildings from Mobile Positioning Data. *Energy Build.* **2018**, *180*, 135–145. [CrossRef]
84. Carmenate, T.; Rahman, M.M.; Leante, D.; Bobadilla, L.; Mostafavi, A. Modeling and Analyzing Occupant Behaviors in Building Energy Analysis Using an Information Space Approach. *IEEE Int. Conf. Autom. Sci. Eng.* **2015**, *2015*, 425–431. [CrossRef]
85. Saeidi, S.; Chokwitthaya, C.; Zhu, Y.; Sun, M. Spatial-Temporal Event-Driven Modeling for Occupant Behavior Studies Using Immersive Virtual Environments. *Autom. Constr.* **2018**, *94*, 371–382. [CrossRef]
86. Sun, K.; Hong, T. A Simulation Approach to Estimate Energy Savings Potential of Occupant Behavior Measures. *Energy Build.* **2017**, *136*, 43–62. [CrossRef]
87. Pereira, P.F.; Ramos, N.M.M.; Almeida, R.M.S.F.; Simões, M.L. Methodology for Detection of Occupant Actions in Residential Buildings Using Indoor Environment Monitoring Systems. *Build. Environ.* **2018**, *146*, 107–118. [CrossRef]

88. Ardakanian, O.; Bhattacharya, A.; Culler, D. Non-Intrusive Occupancy Monitoring for Energy Conservation in Commercial Buildings. *Energy Build.* **2018**, *179*, 311–323. [CrossRef]
89. Ramos, N.M.M.; Curado, A.; Almeida, R.M.S.F. Analysis of User Behavior Profiles and Impact on the Indoor Environment in Social Housing of Mild Climate Countries. *Energy Procedia* **2015**, *78*, 561–566. [CrossRef]
90. Lee, S.; Karava, P.; Tzempelikos, A.; Bilionis, I. Inference of Thermal Preference Profiles for Personalized Thermal Environments with Actual Building Occupants. *Build. Environ.* **2019**, *148*, 714–729. [CrossRef]
91. Gaetani, I.; Hoes, P.-J.; Hensen, J.L.M. Occupant Behavior in Building Energy Simulation: Towards a Fit-for-Purpose Modeling Strategy. *Energy Build.* **2016**, *121*, 188–204. [CrossRef]
92. Ren, X.; Yan, D.; Wang, C. Air-Conditioning Usage Conditional Probability Model for Residential Buildings. *Build. Environ.* **2014**, *81*, 172–182. [CrossRef]
93. Fabi, V.; Andersen, R.K.; Corgnati, S. Verification of Stochastic Behavioural Models of Occupants' Interactions with Windows in Residential Buildings. *Build. Environ.* **2015**, *94*, 371–383. [CrossRef]
94. Boudet, H.S.; Flora, J.A.; Armel, K.C. Clustering Household Energy-Saving Behaviours by Behavioural Attribute. *Energy Policy* **2016**, *92*, 444–454. [CrossRef]
95. Santin, O.G. Behavioural Patterns and User Profiles Related to Energy Consumption for Heating. *Energy Build.* **2011**, *43*, 2662–2672. [CrossRef]
96. Wang, W.; Hong, T.; Li, N.; Wang, R.Q.; Chen, J. Linking Energy-Cyber-Physical Systems with Occupancy Prediction and Interpretation through WiFi Probe-Based Ensemble Classification. *Appl. Energy* **2019**, *236*, 55–69. [CrossRef]
97. Wolf, S.; Calì, D.; Krogstie, J.; Madsen, H. Carbon Dioxide-Based Occupancy Estimation Using Stochastic Differential Equations. *Appl. Energy* **2019**, *236*, 32–41. [CrossRef]
98. Jiang, H.; Yao, R.; Han, S.; Du, C.; Yu, W.; Chen, S.; Li, B.; Yu, H.; Li, N.; Peng, J.; et al. How Do Urban Residents Use Energy for Winter Heating at Home? A Large-Scale Survey in the Hot Summer and Cold Winter Climate Zone in the Yangtze River Region. *Energy Build.* **2020**, *223*, 110131. [CrossRef]
99. Tekler, Z.D.; Low, R.; Gunay, B.; Andersen, R.K.; Blessing, L. A Scalable Bluetooth Low Energy Approach to Identify Occupancy Patterns and Profiles in Office Spaces. *Build. Environ.* **2020**, *171*, 106681. [CrossRef]