Review

A Review of Multi-Domain Urban Energy Modelling Data

Daniel Bishop [1], Patricio Gallardo [2] and Baxter L. M. Williams [3,*]

[1] Department of Civil and Natural Resources Engineering, University of Canterbury, 8041 Christchurch, New Zealand; daniel.bishop@canterbury.ac.nz (D.B.)

[2] EPE Centre, University of Canterbury, 8041 Christchurch, New Zealand; patricio.gallardoocampo@epecentre.ac.nz (P.G.)

[3] Department of Mechanical Engineering, University of Canterbury, 8041 Christchurch, New Zealand

* Corresponding author. E-mail: baxter.williams@pg.canterbury.ac.nz (B.L.M.W.)

ABSTRACT: Urban energy models (UEMs) simulate energy use at the urban scale and are used to inform urban planning, policy development, infrastructure development, and digital twin monitoring and forecasting. Recent technological improvements have spurred interest in large, multi-domain UEMs, which analyse multiple interconnected parts of these energy systems, such as geography, transport, and buildings. Reviews have focussed on single domains or aspects of UEM data. However, multi-domain UEMs require detailed multi-domain data inputs to provide accurate results. This paper provides a comprehensive review of data requirements and a repository of data-specific information for researchers, including data formats, sources, acquisition methods, bridging methods, and challenges. The review was conducted using academic search engines and the authors' direct research experience. Domains are characterised by Climate, Geographic, Building, Transportation, Demographics, Energy Networks and Consumption, and Distributed Energy Resources. Additionally, challenges common to multiple sectors are identified, and methods for addressing these are proposed. The paper concludes with a series of recommendations drawing from the general and sector-specific challenges. Overall, a large amount of data exists, but their use by urban energy modellers is limited due to lack of coordination and standardisation, and concerns over privacy and commercial interests. Coordinated public effort is required to overcome these limitations and improve the results of UEMs in the future.

Keywords: Urban energy model; Accessibility; Availability; Building energy model; Urban energy system; Data challenges

1. Introduction

Urban energy modelling (UEM) involves simulating energy use at the urban scale, from the neighbourhood to the city level. For example, one type of UEM is urban building energy models (UBEMs), which represent buildings at the district scale [1] and are used to simulate heating and air-conditioning loads to inform the design of urban power and energy networks, such as district heating schemes, as well as to investigate energy efficiency scenarios, such as the aggregate impact of retrofitting insulation [2]. Other UEMs include domains such as local climate, buildings, transportation, and energy networks and resources. UEMs, in general have applications including urban planning [3], policy development [4], infrastructure development [5], and digital twin monitoring and forecasting [6]. In each of these applications, UEM can inform the improvement of energy efficiency, resilience, and sustainability by simulating various scenarios, such as integrating solar PV at the district scale [7].

Modelling approaches can be categorised as bottom-up or top-down [8]. Bottom-up, or 'physics-based,' models capture the physical characteristics of urban elements. This modelling approach can simulate changes in the urban environment and thus provide more insight than top-down, or 'data-driven', approaches [9], as changes to the physical system can be tested by modifying values, such as increasing the share of rooftop area to solar PV. However, bottom-up and top-down models require detailed data inputs for accurate modelling.

Given the wide range of UEM domains and applications and the detailed data requirements of both bottom-up and top-down models, it is important for urban energy modellers to understand the available data formats, sources, bridging methods, acquisition methods, and limitations, such as granularity and accessibility.

Several reviews of UEM data have been conducted. Herrera et al. [10] review the methods of weather data creation for building simulation, including a critical analysis of each methodology and a discussion of challenges for weather data, such as climate change and the urban heat island (UHI) effect. Software tools for the generation of future weather files are reviewed and critically analysed by Moazami et al. [11], who conclude care should be taken when selecting and using tools for the generation of future weather data, as weather predictions can vary depending on the prediction method. The effects of green and blue infrastructure (GBI) on temperature moderation in urban environments are reviewed by Bartesaghi-Koc et al. [12], and the modelling techniques and data requirements for simulating the effects of urban GBI are systematically reviewed by Liu et al. [13], who highlights the complexity and diversity of input data for modelling GBI and suggest improvements in data availability and accessibility. Advances in light detection and ranging (LiDAR) systems and the use of LiDAR data for generation of digital elevation models are reviewed by Liu [14], with a specific focus on data filters, interpolation methods, resolution, and data reduction.

Gunduz et al. [15] summarise research in indoor modelling and mapping, a key component of UEMs requiring accurate thermal models or tracking of people and objects within buildings. The results of in-situ insulation tests are reviewed by O'Hegarty et al. [16], including a discussion of over- or under- performing insulation relative to standards and reported values. Modelling approaches for building occupant behaviour are reviewed and categorised by Happle et al. [17], and the application of these modelling approaches to UEMs is discussed. Fuentes et al. [18] review research in domestic hot water (DHW) consumption and the impacts of factors such as climate and seasonality on DHW use, focusing on tools for generating DHW consumption profiles for UEM applications.

Understanding distributed energy resources (DERs) is important for energy researchers and modellers, as discussed by Rathore et al. [19] in their review of several types of solar photovoltaic (PV) cells and their applications. Small-scale wind turbines, an increasingly common type of DER, are reviewed by Tummala et al. [20], who compare different kinds of turbines and highlight the effects of factors such as positioning and aero-acoustic performance. An analytical review of wind turbines and wind resource evaluations in urban environments is conducted by Tasneem et al. [21], which discusses the importance of precise wind mapping and accurate wind data. Energy storage is another common type of DER, as discussed by Rahman et al. [22] in their review of energy storage technologies, which provides a summary of costs and emissions data of available storage technologies.

Additionally, a number of reviews of UEM in general have been conducted. Oraiopoulos and Howard [23] present the results of a systematic analysis of UEMs with their results validated against measured data highlighting the importance of accurate, relevant input data for model accuracy. Dahlström et al. [24] review advancements in and challenges of urban energy modelling. Goy et al. [25] review the impacts of different types of input data in UBEMs and present the results of a case study to rank the impact of input parameters on space heating demand. A comprehensive review of approaches, methods, and tools for UEM is conducted by Ali et al. [26], which identifies challenges and promising techniques. Johari et al. [27] provide a critical review of the field of UEM, including a discussion of possibilities, challenges, and potential future improvements. While they highlight the importance of increasing the accessibility and availability of data for UEMs, their review does not specifically review data requirements and challenges.

To date, data reviews focus on a narrow set of modelling data, often limited to a single domain. Research interests in specific domains of UEM often drive this specificity. For example, researchers quantifying the UHI effect are most concerned with local weather and external building geometries [28]. In contrast, researchers interested in accurately modelling building energy consumption and Indoor Environmental Quality (IEQ) are most concerned with interior building layouts and occupant behaviour [29].

However, practical UEM applications are naturally multi-domain, given the functional requirements for urban planning and infrastructure development are multi-domain. For example, power network planners are interested in the impact of weather pattern changes on heating and cooling loads and the impact of distributed energy resources (DERs) like solar PV. This practical need for multi-domain UEM is reflected in the growing body of recently developed multi-domain tools for UEM [30]. Additionally, linking domains in UEMs offers the potential for greater accuracy, given that the physical systems they represent are linked by themselves. For example, building heating and cooling causes heat-flux into and out of the surrounding local climate and, at the same time ,the local climate affects building heating and cooling loads. Thus, coupling building energy models and local climate models can increase the accuracy of both.

Overall, given the practical needs and opportunities for increased accuracy with multi-domain UEMs, a comprehensive multi-domain review of data for UEMs is required. However, to date, no such review exists.

This work provides a comprehensive review of multi-domain UEM data, with domains categorized as Climate, Geographic, Building, Transportation, Demographics, Energy Networks and Consumption, and Distributed Energy Resources, as shown in Figure 1. Data formats, sources, acquisition methods, bridging methods, and challenges for each domain are identified. Additionally, key overall challenges for UEM data are explored, key considerations and practical implications are summarised, and recommendations are provided for multi-domain UEM data.

Climate	Weather	Climate projections	Local climate	
Geographic	Terrains & Elevations	Natural & Manmade		
Buildings	Geometry	Construction	Occupant behaviour	Building Systems
Transportation	Networks	Travel		
Demographics				
Energy Networks	Networks	Usage		
Distributed Energy Resources	Technological Specifications			

Figure 1. Overview and categorisation of data requirements for urban energy models.

The literature review for this work was conducted using several academic search engines, including *Google Scholar*, *Scopus*, and *Consensus*, an AI-powered academic search platform. These tools were chosen to ensure comprehensive coverage of relevant studies across multiple domains. Some references were selected according to the authors' direct research experience, providing further depth and context to the data sources used to develop multi-domain urban energy models.

This work is structured as follows. Section 2 prefaces the work with a review of general data challenges across all domains of UEM. Section 3 provides a data review of each UEM domain: Climate, Geographic, Building, Transportation, Demographic, Energy Network, and Distributed Energy Resources. Section 4 identifies key considerations, practical implications, and recommendations for UEM data and data research. Section 5 concludes the review.

2. Data Challenges

Several general challenges exist for researchers seeking to obtain sufficient, and sufficiently accurate data for UEM. Many of these challenges stem from inconsistencies in data availability, and variations in data quality between regions, which pose particular challenges for constructing large-scale UEMs [1]. While different modelling applications have different data requirements, challenges for urban energy modelling data typically fall into four key areas [25]: (i) data availability; (ii) data accessibility; (iii) data quality; and (iv) data impact.

2.1. Data Availability

Data required for some urban energy models can be difficult to collect [23]. For example, many models require information about building energy systems, ranging from simple residential open fires to complex commercial heating, ventilation, and air conditioning (HVAC) systems. These data are difficult to collect, particularly in the residential sector, so information about building energy systems is typically estimated based on factors such as the year in which the building was constructed [31]. Other examples include building occupancy and energy consumption data, which can also be difficult to collect. Building occupancy and, to a lesser extent, energy consumption can be assumed to follow standard profiles [32,33] or can arise as the result of simple behavioural rules [34]. However, while these methods offer means of estimating the required data, they may lead to inaccuracies, particularly if the estimates are based on information from other regions.

Additionally, data availability varies between different regions [26]. Thus, even easily collected data can be unavailable in some regions [23]. Data collection methods can also vary. For example, different surveying techniques can lead to different measurements of building footprint [35]. Thus, datasets collected using different methods can be incompatible due to differing assumptions. Furthermore, even where data are available, data granularity can be insufficient for constructing accurate energy models [25], meaning low-granularity data are often interpolated for sufficient resolution, which can limit model accuracy.

2.2. Data Accessibility

Limited accessibility means that even if data are collected, they may not be publicly available. Cities in most economically developed countries have large databases with energy-related information, but these are often unavailable for use by modellers [36]. Accessibility can be restricted because data are proprietary, in unsuitable formats, dispersed across multiple databases, and/or due to confidentiality/privacy concerns [25].

2.3. Data Quality

Even if they are collected and publicly available, data can be of variable quality. Data quality can be limited by variability in collection methods, as standardised collection techniques do not exist for many kinds of data required for UEMs [37,38]. Thus, obtaining data from different databases, as is often required due to limited availability/accessibility, can limit data quality due to inconsistencies in the collection methods [36]. Additionally, data quality can be limited due to infrequent collection, as dynamic sectors, such as the building sector, can change rapidly, limiting temporal validity. For example, the lack of consideration for building refurbishments, which are often not reflected in publicly available building databases, can lead to overestimations in heating demand by up to 180% [39].

2.4. Data Impact

The impact on modelling outcome varies between different types of input data [23]. Consequently, the importance of availability/accessibility can also vary between different types of data. For example, building construction characteristics typically have a larger impact on model outcomes than building geometry. Still, data on the former tend to be less accessible than data on the latter [25].

2.5. Addressing Data Challenges

Because of these challenges, improving data availability, accessibility, and quality is considered a priority [25,40,41]. Specifically, researchers agree that the collection of energy-related data should be increased and made publicly available [42–44], which could be best achieved by making these data available via an online public platform [25].

However, while such changes would increase data availability and accessibility, data quality and impact should not be overlooked. Thus, priority should be given to collecting high-impact data, such as building construction characteristics, first, and to ensuring that data collection methods are accurately reported. Standardised data collection methods should also be adopted where possible to avoid inconsistency between different databases.

3. Data in Urban Energy Modelling

3.1. Climate Data

Weather drives building energy consumption and renewable energy generation. Weather datasets include external air temperature; humidity; wind speed and direction; pressure; precipitation; surface albedo; solar radiation and its components, including direct and diffuse radiation; and sky illuminance and its components [45]. Location and time are recorded with latitude, longitude, elevation, and local time stamps. Data are compiled at sub-hourly to hourly time-steps, and statistical information. Such as maximums, minimums, averages, and frequency distributions, are also often reported [10]. Weather data are collected at weather stations, which are typically in remote locations away from the influence of micro-climatic effects [46].

Weather datasets may consist of the historical weather conditions for a specific year, as with actual meteorological year (AMY) data, or typical data, as with typical meteorological year (TMY) data [10]. TMY weather conditions are annual weather datasets representative of historical years, derived from historical data, and constructed by statistical methods [47].

Recorded weather data's temporal and spatial resolution limits their applicability and accuracy for UEMs. Typically, data from the closest weather station are used, which may not accurately represent the weather at the location of interest. However, weather generators producing "synthetic" weather data have been developed to overcome these limitations [10]. "Meteonorm" [48] generates stochastic TMY at time intervals as low as one minute and has been used for building simulation [49]. "LARS-WG" [50] is a stochastic weather generator that can downscale the spatial resolution of weather data [51] and generate data for locations between weather stations. In some cases where only low-resolution data are available, the highest-resolution dataset is used [52].

The urban heat island (UHI) effect is a phenomenon where towns and cities experience higher temperatures than surrounding areas due to differences in vegetation, surface albedo, increased heat retention, reduced airflow due to built-up intensity, and/or anthropogenic heat emissions [53]. Additionally, interactions between buildings in the urban environment include shading between adjacent buildings, longwave radiant heat exchange, and solar reflection [54].

Consequently, the urban form has an impact on building conditioning loads. Tools have been developed to capture the urban microclimate due to the UHI and the interactions between buildings, including ENVI-met [13,55] and Urban Weather Generator (UWG) [56,57].

Anthropogenic climate change means weather changes can occur within the lifetimes of current buildings and infrastructure, so climate change must be considered by urban energy modellers. Climate models are used to predict future weather patterns, which in turn are used to simulate buildings and predict future energy performance [58]. The Intergovernmental Panel on Climate Change (IPCC) produces climate models for multiple scenarios. The models include Global Circulation Models (GCM) and finer-resolution Regional Climate Models (RCM) [59]. However, their finest resolution is a 25km-by-25km surface grid [10]. Tools are available to generate representative localised weather data from these climate models [11], including "CCWorldWeatherGen" [60] and "WeatherShift" [61].

3.2. Geographic Data

The urban environment contains natural and human-made features that can impact the energy performance of buildings and the effectiveness of renewable resources. Natural features include terrain elevation, parks, trees, and bodies of water. Human-made features include transportation infrastructure, such as roads and buildings. Many natural and human-made elements participate in microclimatic effects, affect wind-flow, and produce shading, thus affecting energy performance. Failure to consider the effects of geographic elements, such as terrain, can limit model accuracy. For example, some UBEMs assume lang surfaces are flat rather than modelling the effects of changing elevation, leading to inaccurate results [62,63]. Thus, the effects of terrain and other geographic data are important elements for consideration in UEMs.

3.2.1. Terrain and Elevation

Two types of Digital Elevation Models (DEM) exist: Digital Surface Models (DSM), which include natural and human-made features; and Digital Terrain Models (DTM), which exclude human-made features and provide the elevation of bare land. Elevation models are created from contour lines, topographic maps, global positioning system measurements, photogrammetry techniques, radar interferometry, stereo satellite images, and laser scanning, with vertical measurement accuracies varying according to the method used [64]. Satellite techniques can produce DEM spanning the globe, with vertical accuracies of 1.5-6 metre root-mean-square error (RMSE). In contrast, airborne laser scanning techniques, also known as light detection and ranging (LiDAR), require localised measurement and have lower vertical accuracies of approximately 0.15 metre RMSE [14,64].

A number of satellite-derived DEMs are freely available, which have been used to map urban structures [62,65,66], including the Shuttle Radar Topography Mission (SRTM) elevation data, captured in the year 2000 with a 30-metre surface mesh [67]; Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), which has a 30-metre surface mesh and was released in 2009 [68]; Advanced Land Observing Satellite (ALOS) World 3D surface model, which has a 30-metre surface mesh and is based on observations between 2006 and 2011 [69]; and the Global Multi-resolution Terrain Elevation Dataset of 2010 (GMTED2010) [70].

In addition to lower-accuracy satellite-derived DEM, higher-resolution LiDAR-derived DEM have also been produced. However, the availability of these higher-resolution data is limited. Furthermore, given its higher resolution of measurement, LiDAR has been used to determine building ages [71], facades [72], and building geometries [73]. "OpenTopography" is an initiative to support data sharing and access to high-resolution topography [74]. Highly accurate LiDAR datasets have been shared on this platform, including for large tracts of New Zealand, collected between 2010 and 2022 [75].

Dataset age presents a barrier to the applicability of elevation datasets, particularly for extracting building features. Where building features are extracted, the building must have been built before the dataset was collected, which limits the use of such techniques for recently constructed buildings. Additionally, dataset accuracy presents another barrier. Although satellite-derived datasets are freely available and provide global coverage, they have low accuracy for building purposes and can produce significant errors when used to extract building dimensions. [62].

3.2.2. Natural & Human-Made Features

Built infrastructure, roading and buildings, and Green and Blue Infrastructure (GBI), which describes vegetation and bodies of water, affect the local microclimate due to heat retention, changes to albedo, evapotranspiration, solar shading, and wind flow modification. GBI has been used to modulate local climate to increase comfort and decrease energy consumption [13] and have been incorporated into UEM simulations as energy efficiency interventions [76,77].

Local climate zone (LCZ) is a standard for characterising urban areas for local climate analysis, which is divided into 10 urban zones and 7 natural zones. Zones include categories for compact and open high-rise, mid-rise, and low-

rise buildings, pathed areas, bush cover, and water [78]. Subclassification, "levels", quantifies further details, such as tree morphology and soil type [78]. These zones have specific attributes relating to their impact on the local climate.

Various means have been used to measure climate relevant land qualities, including remotely sensed spectral data; airborne LiDAR and spectral airborne-based LiDAR; low- and high- resolution satellite imagery; terrestrial laser scanning; aerial and terrestrial photography; and in-situ inspections [12]. The World Urban Database and Access Portal Tools (WUDAPT) project is a global initiative to collate and disseminate relevant climate data [79], including LCZ maps [80], which can be generated automatically from satellite data [81]. LCZ classifications, such as ENVI-met, are often used as a model input for urban climate modelling packages [82].

3.3. Building Data

Building energy modelling (BEM) involves detailed physics-based modelling of buildings to understand energy consumption and evaluate the effects of energy interventions [83]. With advances in computing power, the related field of urban building energy modeling (UBEM) has emerged, involving the simulation of numerous buildings at an urban scale.While the two fields overlap, UBEM typically represents many buildings at a lower level of detail (LoD), while BEM represents fewer buildings (or one building) at a higher LoD [84]. Thus, while the two fields require similar data types, the impact and importance of different data can vary between BEM and UBEM.

Both BEM and UBEM can be implemented in various software environments, including open-source applications. For example, EnergyPlus is an open-source building energy simulation package developed by the United States Department of Energy, which can model a range of building energy functions [85].

3.3.1. Data Formats

Several data formats can be used to store UBEM data. Common formats include IFC and gbXML, which are standard for BEM; and GeoJSON, Shapefiles, and CityGML, which are more common for UBEM [63]. The Industry Foundation Class (IFC) [86] is used to describe building and construction data and "green building." XML (gbXML) [87] is designed for sharing data between building design and simulation tools. GeoJSON [88] is a JSON-based file format for geospatial data, which is commonly used in mapping applications and UBEM platforms [63] and is easily implementable in a range of applications. The OpenStudio City Modelling Framework extends the use of the JSON format for UrbanOpt [89]. Shapefiles are a type of geospatial vector data format used to store and display geographic information, consisting of a set of files representing geographic features such as points, lines, and polygons [90]. Shapefiles are commonly used in GIS systems [63] and are thus well-suited to urban modelling applications, as GIS systems are specifically designed to handle and analyse spatial and geographic data and have well-developed resources for those tasks.

CityGML (City Geographic Markup Language) is an open and standardised data format specifically for cities and landscapes, which was developed by the Open Geospatial Consortium [91]. CityGML has several features [92] suiting the requirements of UEMs: (i) The format has native support for city elements such as city furniture, buildings, transportation, vegetation, and water bodies; (ii) The format supports multi-scale modelling. For example, geometries may be represented in four increasing levels of detail (LOD0-3); in LOD0, buildings are represented by surface footprints, while in LOD3, buildings have highly detailed representations described by sets of surfaces, internal and external geometries, apertures, and shading devices; and (iii) the format supports the development of application domain extensions (ADE), sets of additional classes, attributes, and relations.

EnergyADE is an ADE developed to allow for detailed single-building energy simulations and city-wide bottom-up energy assessments, including data for building physics, occupant behaviour, materials, construction, and building energy systems [93]. EnergyADE is not intended for use with centralised energy infrastructures, such as district heating networks, as these are the focus of other extensions, such as the Utility Network ADE. However, the Utility Network ADE project appears to be stagnant, as the project is still in draft status and was last updated in 2012 [94].

3.3.2. Building Geometry

Building geometry data requirements increase with increasing LoD [95]. UBEM often uses models with lower complexity than those in single-building BEM, reflected by the lower LoD typically used for models [63], which typically consists of external and assumed geometries, and single thermal zoning instead of multi-zoning [44]. For example, UBEMs often include approximations such as the use of single, centrally located windows on building faces, which are sized to meet assumed or known window-to-wall ratios (WWRs) [96]. Thus, UBEM can have lower per-building data requirements than high-LoD BEM.

Diverse types of building geometries exist. Building external geometry data include building footprints, height, orientation, floor number and height, and roof shape. Façade geometry data include the size and positioning of doors, windows, and shading devices. Data may be simplified, such as simplifying the size and distribution of windows to a

single WWR. Internal geometry data include an internal layout of partitions and the zone type. For example, zones can be classified as occupied/unoccupied, or according to use, such as office, bedroom, living area, and circulation area.

While building plans and building information models (BIM) often contain all the required building information for UBEM, access to these data can be limited, and is typically restricted to the municipalities. Additionally, as it can be cumbersome and time-intensive to extract the relevant data when available, it is typically infeasible to utilise BIM at the urban scale. Furthermore, in the case of aged building stock, BIM may not be available, and building plans have a variety of formats, such as scanned drawings. Overall, while building data are often available, extracting the required information for urban-scale analysis is typically impractical. Conversely, automated feature extraction from BIM or drawings presents an opportunity to enhance the availability of building geometric data for UEM.

Building footprints provide building orientation and overall shape and can be vertically extruded to produce a 3D building shell [27], a common workflow for UBEM production. Footprints can be generated from direct surveys, satellite imagery, aerial imagery, oblique photogrammetry, and LiDAR [63], where image recognition and extraction can automate footprint generation [97].

Google's Open Buildings footprints [98], generated from satellite imagery using machine learning [99], and Open Street Maps footprints, generated by users from various sources [100,101], provide freely available building footprints, which have been used for UBEM [62,102,103]. Another footprint source is local municipality cadastral records, which may have greater accuracy than footprints derived from satellite imaging and have also been used for UBEM [104]. Some cadastral records are openly available [105–110]. Accurate commercial mapping services are also available [62]. In addition to footprint, mapping services provide urban layout data, such as the distribution and orientation of buildings, parks, and roading.

Building heights can be obtained from building plans and records [111]; obtained from Energy Performance Certificates [24,112], which are required for certain buildings in the European Union (EU); obtained from tax records [113]; or calculated by assuming floor-to-floor heights from the number of stories [114,115]. Heights can also be extracted from normalized Digital Surface Models (nDSM), which are created by subtracting the DSM from the DTM [62]. Satellite-derived DEMs typically have low accuracy [63,116], but due to higher precision, LiDAR-derived DEMs can calculate building heights with greater accuracy. Similarly, LiDAR data can produce detailed roof geometries [73]. Additional methods to obtain building heights include oblique photogrammetry from aerial photography [117] and shadow measurement [118]. Where building heights have already been characterised, these data may be available as attributes in Google and OSM datasets [98,100] or city GIS datasets [119].

Fenestration information is often not included in building datasets [44,63], likely due to the difficulty of obtaining these data. Thus, several means to determine simple WWRs include assuming from building archetypes [96], professional judgement [44], aerial infrared thermography [120], and manual extraction from photographs [111,121]. However, the simplified WWRs lack accuracy for muti-zoned models, so detailed fenestration data, such as the size and positioning of individual windows, are required. Methods to extract detailed fenestration data include low-resolution aerial photographs with analytical extraction techniques [122,123], terrestrial and satellite photography and manual extraction [124], street-based photography and artificial intelligence extraction techniques [115], oblique aerial images and automated [115,122,125], LiDAR scanning of façades, and automated extraction [72]. However, many of these techniques are cumbersome, and thus not readily implementable in UBEMs.

External shading devices like louvres and overhangs, can affect building energy use by reducing solar gains and cooling loads. No research has been found for the automatic or manual characterisation of external shading devices, either by remote sensing or other means.

Modelling building interior partitions is common in traditional BEM but not in UBEM, due to the difficulty of collecting interior layout data at scale. However, model LoD and thermal zoning strategies do affect energy performance [112], so including building interior modelling will increase the accuracy of BEMs and UBEMs. Several techniques exist for indoor mapping, including the use of one or a mixture of laser sensors and scanners, stereo-imaging, and sonar [15,126]. However, these techniques are often manual time-intensive, and expensive. Crowdsourcing the task of harvesting internal boundaries to smart phones has been explored to overcome these limitations, and several platforms exist to convert smartphone data into indoor floorplans [126]. No largescale open datasets for indoor maps exist, however, Google Maps has added support for viewing multi-story indoor maps [127]. The dataset is small and limited to commercial and public spaces, which have an incentive publicise building layouts. Current datasets have limited applicability to the residential building stock and UBEM. Additionally, no research has been found integrating generated indoor maps, or derived thermal zoning, into UBEM. No large-scale open datasets for indoor maps exist. However, Google Maps has added support for viewing multi-story indoor maps [127]. This Google dataset is small, and limited to commercial and public spaces, which have an incentive to publicise building layouts. Current datasets have limited applicability to the residential building stock and UBEM. Additionally, no research has been found for the integration of indoor maps, or derived thermal zoning, into UBEMs.

3.3.3. Building Constructions

Building material thermal performance data are required for accurate thermal modelling in BEM and UBEM. These data include material type, airtightness, and thermodynamic characteristics such as U-values. In BEM, physical constructions can be extracted from building plans and surveys, then U-values and other characteristics obtained from standard values based on the obtained data. However, this procedure can be inaccurate, as measured values can differ substantially from idealised assumptions [16]. Thus, where available, directly measured data are superior to those calculated from standard values.

Construction data for UBEMs can be collected from energy performance certificates [104,128]; professional judgement based on historical standards, construction type, and building age [44,129]; and national building archetype databases, such as Tabula in the European Union [58,130–132]. These methods have a low accuracy, as they are typically not based on building measurements or direct observations, and thus can fail to account for key features, such as retrofitted insulation.

Although not yet implemented into UBEM, thermal performance data may be measured directly, which has the potential to overcome data accuracy and accessibility limitations. U-values can be calculated from on-site data, such as heat flux measurements and infrared thermography [133,134]. Large-scale acquisition of thermal property data via aerial and UAV infrared thermography has been proposed [120,135], although the technique has low accuracy. Additionally, image recognition may be used to classify exterior construction material [136], which can be used to allocate buildings to the appropriate archetype. For this task, source images can be extracted from public repositories, such as Google Street View [124]. Overall, the limited accuracy and availability of building construction data are ongoing limitations for UBEM.

3.3.4. Occupancy and Occupant Behavior

Occupant behaviour includes occupancy schedules, electrical plug loads, lighting schedules, temperature set points, window operation, solar shading operation, and hot water usage [137]. Each of these behaviours is important for UEM [25], as they drive building energy demand and can affect the timing of energy consumption, so they are key to understanding peak energy demands in energy networks [138,139] and can contribute to the increased potential for energy demand response [140]. Four general approaches are used to represent occupancy and occupant behaviour in UEMs, including deterministic, stochastic, and agent-based methods [17,137,141]. Deterministic occupancy schedules are often assumed based on expert judgement and are most often used in UBEM [23]. However, where available, measured data from individual buildings are the most accurate [142].

Time use surveys (TUS) are statistical surveys collecting data on how people spend their time. These surveys have been conducted in multiple countries and continents. By determining where occupants are likely to be at different times of day, TUS and similar datasets can be used to derive occupancy and behaviour schedules for residential and commercial buildings [63,137]. Data from smart electricity meters have also been used to determine occupancy and appliance use [143].

Deterministic occupant behaviour profiles are provided by guidance documents from institutions such as CIBSE and ASHRAE [144,145] and from building standards [146]. Stochastic profiles account for variability in human activity, which is critical for accurately predicting peak loads in large energy networks due to load diversification. Stochastic profiles can be based on people, such as creating a location profile for each occupant, or spaces, such as creating an occupancy profile for each room [17]. Several stochastic occupancy and behaviour profiles have been developed, including stochastic occupancy schedules based on the American Time-Use Survey [147]; and StROBe and newStROBe, derived from Belgian household data [148,149]. Agent-based models (ABMs), which simulate human behaviour and their interactions with other agents and the environment, have been used to generate occupancy and behaviour profiles [150]. The use of ABMs to generate behavioural profiles is promising at the urban scale, as ABMs can couple transportation and building utilisation [17] and lead to more accurate urban-scale occupancy distributions. Additional models have been developed to generate specific occupant behaviour, such as electrical plug loads [151], hot water consumption [18,152], and both [153].

3.3.5. Building Systems

The correct characterisation of building systems, such as the rated capacity and efficiency of Heating Ventilation and Air Conditioning (HVAC) and Domestic Hot Water (DHW) systems, is important as they are a significant driver of building energy consumption [23]. Building energy system characteristics have been extracted from existing databases, such as census data, the European EPC database, and the US Building Performance database [26,96,154]. Most commonly, archetypes are used, where systems are assigned to buildings based on the year of construction, the existence of district heating systems, and professional judgement [44,119,128,155]. Building systems have also been

ignored where the type of analysis is focused on other aspects of the energy system [104]. While not yet deployed in BEM or UBEM, non-intrusive monitoring from smart power meters, is becoming increasingly ubiquitous, and signal processing may be used to determine installed equipment, appliance loads, ratings, use, and efficiencies [143,156–159].

3.4. Transportation Data

Transport has been studied at the urban level for classical 'four-step' transport demand models, which include trip generation, trip distribution, modal split, and traffic assignments [160]; land use-transport interaction modelling; accessibility analysis [161,162]; transport poverty studies [163]; and transport energy requirements [164]. Including transportation in UEMs provides a broader and more useful energy system model [3] and can improve overall model accuracy by accounting for the interconnection between vehicle use and building occupancy [24].

A range of data sources exists for modelling urban transportation. Vehicle licensing records provide annual travel demand based on odometer readings and vehicle demographics, such as year, make, and model. Travel diaries and surveys, which record individual or household travel patterns for several days, have been used for multivariate statistical mode choice modelling [165]. Traffic counts, which record the vehicles passing through sampled streets, have been used for congestion monitoring, support analysis of travel patterns, and validation of other transport models [166,167]. Census data provide demographic data, including car ownership, and can include travel-specific data, such as travel demand, and cognitive factors, such as travel attitudes and preferences [165]. Census data benefit from large sample size and the inclusion of multiple factors such as region, household, and income, which improve their usefulness to modellers. Geographic data include land use and the spatial layout of road and transit networks, and departure and destination locations, such as residential areas, food, health, education, and retail facilities [168]. Geographic data are available from open sources, such as OpenStreetMap. The General Transit Feed Specification (GTFS) is an open standard for public transportation schedules and associated geographic information, such as rail and bus stops, routes, and timetables [169], which can support urban transport accessibility studies [170].

In addition to classical sources, technological developments such as the internet of things (IoT) and the advent of Big Data have provided several new sources of transportation data [171,172]. Global positioning system (GPS) loggers are now present in smart phones, wearable technologies, and many other devices, including vehicles. GPS data are used by services such as Google Maps and ride-sharing apps to provide information on congestion and adjust route generation. Similarly, social media posts are often 'geo-tagged', marking their position using the GPS capabilities of smart phones. Thus, social media data-harvesting can extract population-level travel behaviours and locations of interest but lacks accuracy at an individual level due to privacy restrictions [173,174]. Similarly, IoT devices, such as traffic cameras and smart traffic lights, can generate real-time traffic data [175,176]. Smart ticketing systems in public transport record travel times and entry and exit locations so they can provide detailed information on public transit use. However, these data are limited to the transit network and thus do not reveal passengers' actual origins or destinations. Records of network connections, such as cellular and wireless networks, contain time and approximate location. They can be used to infer travel routes but are limited to passengers connected to these networks during their transit.

Overall, while data-rich resources are becoming increasingly available, there are still considerable barriers to overcome before large-scale integration into UEMs. Specific challenges include integrating and sharing data, which can be challenging due to data volumes and private ownership, data ownership and privacy, and data quality and standards [171].

3.5. Demographic Data

Demographic data describes the characteristics of a population and includes socio-economic data, which can include age, gender, ethnicity, income, employment status, household composition, and location, all of which can affect energy consumption [177]. For example, space heating use typically increases with income level and with age [178]. Thus, regions with higher household income levels and/or older populations may have higher energy consumption, so socio-economic data are important inputs for accurate UEMs.

Demographic data are commonly employed in top-down UEMs [179], where regression is performed to correlate demographic characteristics with energy consumption patterns, and projections can be performed, such as forecasts of population growth. In bottom-up UEMs, demographic data are used indirectly to assign other inputs, such as space heating methods.

Sources for demographic data include surveys, billing data, tax records, and census records [26,180–182]. Census data are typically the most common demographic data source, due to the high detail, breadth, quality, and accessibility of data collected. Limitations of census data are the frequency of collection, which is commonly several years, and the lack of specific energy-related data. Due to privacy concerns, census data are anonymised, which limits connections between energy and census datasets at an individual or household level, thus preventing correlations between demographic data and energy consumption patterns at this level. Instead, aggregate correlations are used, which can limit the efficacy of census data for UEM.

3.6. Energy Network and Consumption Data

Energy networks can include electricity networks, distribution of gas and other fuels, and district heating schemes. In these networks, the placement and sizing of elements, such as distribution pipes, electricity lines, pumps, and electricity transformers, can vary. Thus, network-specific data are required to identify limitations, such as the capacity of electricity distribution transformers, and for planning and system design.

Consumption data in these energy networks are required to tune and validate bottom-up energy models, and are the primary input for top-down, data-driven models [26]. Energy consumption data can be collected at any level: dwelling, building, district, region, or nation. These data are typically collected by meters, such as electricity meters at the building or district level. Lower-level energy consumption data are the least aggregated and thus the most useful for high-granularity UEMs considering individual buildings or districts. However, access to dwelling- and building- level data is limited due to privacy concerns, so access typically requires the consent of each consumer [183]. Additionally, because meters are typically owned by energy retail companies, the data are usually proprietary, requiring the owner's cooperation for access. For example, electricity consumption data are plentiful in New Zealand due to the nationwide roll-out of smart electricity meters [184], but these data are owned by the electricity retail companies and are typically unavailable for energy modelling. To address this limited accessibility, some studies in New Zealand have collected non-proprietary, anonymised electricity consumption data from small numbers of consumers [184]. However, while these studies can help to address data privacy and ownership concerns, their small sample sizes raise concerns about representativeness. Additionally, small-scale studies of energy consumption can be particularly prone to issues of incompleteness [185], and issues can arise when combining datasets because of differences between them [186].

District-level consumption data can be sourced from energy distributors and are aggregated at sufficient resolution to overcome privacy concerns [104]. However, these data are proprietary, so access requires the distributor's participation. As distributors are not typically incentivised to participate in data-sharing, access to these data is often limited [44]. For similar reasons, access to energy network system design, such as network layouts and system capacities, is often limited [34]. However, high-level data are sometimes available directly from distributors, such as PowerCo in New Zealand, who provide an interactive map of their medium-voltage electricity network for planning purposes [187].

High-level energy consumption data are usually more openly accessible than lower-level data. At the state or national level, data reporting is often mandated, making energy network and consumption data frequently publicly accessible. For example, in New Zealand, the Electricity Authority reports electricity power flow and location-specific electricity costs [188]. However, the high-level aggregation in these data means they may be poorly suited to models with higher granularity.

Overall, high-level consumption and network data are typically more widely accessible than lower-level data, but their aggregation means they may be less useful for energy modellers. Conversely, low-level data are typically more useful but are often inaccessible because of privacy and/or commerciality concerns. Thus, improving the accessibility of high-granularity data, such as from individual buildings' energy meters, would increase the quality of UEMs requiring these data. However, the privacy of energy consumers must be protected, so datasets should be appropriately anonymised and retain useful, non-identifiable information, such as the region from which the data were collected.

3.7. Distributed Energy Resources

Distributed energy resources, such as solar PV, micro-wind turbines, and battery technologies, are increasingly being integrated into electricity systems at the building and district levels. Thus, variables affecting distributed energy resources, such as resource potentials, weather data, technological specifications, and costs, are increasingly important data inputs for UEMs. In addition to the weather data sources described in Section 3.1, the World Bank's Global Wind Atlas and Global Solar Atlas projects provide verified estimates of wind and solar resource potential for locations around the world [189,190]. Detailed resource potentials can also be estimated at the national level. OpenEI (Open Energy Info) provides the Open Energy Data Initiative data portal, where high-resolution resource potential data are available for North America [191], and Data Europa provides similar data for Europe [192].

Energy technology costs, related emissions, and specifications, such as PV panel efficiencies and temperature coefficients, power curves for micro-wind turbines, inverter conversion efficiencies, and battery efficiencies, are important considerations for the effective design of urban energy systems. The most relevant energy technologies at the urban scale are solar PV, small wind turbines, and battery technologies, as these are the most common [193]. Accurate technology specifications are available from manufacturers' product datasheets. However, at early design stages, or when making technology- rather than product-level decisions, general figures are most useful. Generalised energy technology data have been produced by institutional scientific review and are generally included in the internal libraries of energy modelling software platforms, such as SAM and HOMER [194,195]. Solar PV technologies have various established and emerging cell architectures. Important solar cell parameters include cell efficiency and power performance with temperature. Solar PV cell types are reviewed in [19], where cell type efficiencies and temperature

coefficients are quantified, and qualitative data is provided on size, cost, and high-temperature performance. NREL provides technical reports on the module costs of commercially available and emerging Solar PV technologies [196].

Various terms are used for small-scale wind generation, which is the most relevant at the urban scale, including micro-wind, small-wind, and building-integrated wind generation. Wind electricity generation at this scale is still an emerging technology, and less mature than solar PV in its development and ubiquity. Due to these factors, research reviews are limited to reviewing the technology types [20] and the connection between the local wind resources and small-scale wind turbines [21]. There is an absence of data on the cost of installation and operation, and the overall energy costs, such as LCOE calculations, for small-scale wind, which is likely due to the limited number of installations and the strong dependence of these costs on installation location and local topology, making costs and energy yields far more variable for micro-wind than for solar PV.

Energy storage technological options are reviewed in [22], where factors such as rated power, specific energy, energy efficiency, discharge time, response time, lifetime cycles, self-discharge, and relative costs are compared. The review highlights the high variability within technologies and provides ranges of these factors for different technologies. Additionally, at an institutional level, the European Union department of Energy has published a database of European energy storage technologies and facilities [197], which provides simplified costs and specifications. Increasingly, appliances with flexible electricity demand, such as hot water cylinders [198,199] and electric vehicles [34], are being used as distributed energy resources to provide ancillary services to the power system, such as reducing peak electricity demand and increasing utilisation of intermittent renewable generation. Thus, understanding this process, known as demand response (DR), is becoming increasingly important for urban energy modellers, as DR programs can influence electricity time-of-use and total energy use. The constraints and considerations of DR, including technical, economic, and behavioural factors affecting DR programs, are reviewed in [183].

In addition to the technology-specific sources described above, the United States Energy Information Administration produces an "Annual Energy Outlook" representing their assessment of the total system costs to develop and install various electricity generation technologies [200]. Additionally, financial consultancy Lazard produces annual reports on the levelized cost of energy (LCOE) for different generation and storage technologies [201].

Finally, electricity generation-related emissions are reviewed and reported by the Intergovernmental Panel on Climate Change (IPCC) in the "IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation," which reviews Life-Cycle Analyses (LCAs) of electricity generation technologies, calculates the lifetime emissions attributed to each generation technology, and compares emissions with lifetime generation [202]. For combustion sources, such as wood or gas boilers, the IPCC publishes an "Emissions Factor Database" to catalogue emissions factors for a combination of fuels and combustion techniques [203].

In general, energy resource data have several limitations. Due to economies of scale, energy technology data are often available for utility-scale data but are unsuitable for smaller urban-scale applications. Energy technology data can also fail to account for variability, as data are typically presented with a single representative figure and, where they are provided, data ranges are often inordinately large and are unaccompanied by information on when these ranges apply, so are impractical for use by energy modellers. Thus, a large-scale repository of DER products, with manufacturers' specifications and, indicative costs and typical energy yields, would be a useful and timesaving addition to urban energy systems design. Such a repository would allow direct comparisons between technologies and products with a high level of accuracy. However, no such repository currently exists.

4. Key Considerations, Practical Implications, and Recommendations

This section contains general summaries, takeaways from each domain, and general recommendations for improving multi-domain urban energy modelling data.

Climate data

- Climate data are well-developed and are typically available from government meteorological agencies and/or weather-forecasting companies.
- Established methods exist for downscaling spatially, such as for a particular location from large-scale data, and temporally, and for accounting for the effects of urban heat islands.
- Climatic changes may be apparent within the lifetimes of existing infrastructure, so future weather patterns are a key factor in design. Thus, urban energy models should incorporate these changes.

Geographic data

- A range of measurement techniques exist to determine terrain data from satellite imagery and aerial photography.
- The most ubiquitous open-source data are satellite data, which have limited accuracy, while the most accurate and useful data are from LiDAR. However, the availability of LiDAR data is limited.
- LiDAR has proven to be useful across a broad range of urban energy data requirements. Thus, increased collection and availability of LiDAR data represents a worthwhile investment.

Building data

- Building external geometries are the most well-developed building data input, with raw data from building footprints, LiDAR-extracted building heights, and fenestrations achieved through oblique aerial photography. Floor number and height, accurate window placement and size, and solar shading geometries are important for building energy modelling. However, methods of accurate determination of these data are limited.
- Conversely, interior building layouts, which are required to determine thermal zoning and accurate energy and comfort assessments, cannot be determined at scale using currently available methods.
- Building construction data are typically obtained from energy performance certificates and historical standards. However, real values can differ from the hypothetical values obtained from these performance certificates. Thermography and in-situ provide real measurements of thermal properties but are limited in accuracy or scalability.
- Occupant behaviour is highly variable but highly impactful for energy modelling, particularly for time-of-use considerations. Various overlapping methodologies exist to simulate building occupant behaviour, including deterministic, stochastic, and agent-based models. Stochastic methods are not necessarily accurate for any individual building but can produce accurate results in aggregate and are thus well-suited for urban energy models.
- Building systems are a significant driver of building energy consumption, which can only be determined accurately for any individual building through performance certificates or other records. However, surveys can provide accurate distributions of technology utilisation. Signal processing of data from smart electricity meters can provide a method for remote determination of building system characteristics at scale without in-situ measurements. Still, these data are typically proprietary and thus not openly accessible.

Transportation data

- Transportation network data and public transport routes and scheduling, are well-developed, digitised, and widely available.
- Vehicle ownership, travel demand, and travel distribution data are highly impactful for urban planning purposes, particularly with the increased adoption of electric vehicles. However, access to these data is limited.

Demographic data

- Demographic data are widely available at aggregate scales, such as city, state, or nation, and can thus provide accurate distributions for assignment in urban energy models.
- While high-granularity demographic data are collected, privacy concerns limit the accessibility of these data.

Energy network and consumption data

- Aggregate energy consumption data are widely accessible in many places, as national- or state- level data are openly available in many countries.
- Individual building energy metering data exist, but access is limited due to privacy and commerciality concerns. Increased access to these data could provide considerable benefits for the tuning and validating various energy models.

Distributed energy resources

- Many DER technologies are available, the most common of which are solar PV, micro-wind turbines, and residential batteries. However, technical characteristics, and thus specific performance, can vary, and integration of DER in energy models requires access to manufacturer datasheets.
- To the best of the authors' knowledge, no open data repository exists for the technical performance characteristics of DER equipment.

Recommendations

- The establishment of robust methods for predicting regional climate patterns, including their uncertainty, would facilitate assessment of the effects of climatic changes in urban energy models. Thus, these methods should be investigated.
- LiDAR data provide the most accurate and useful terrain data for geographic terrain and determination of building heights, making it important for energy modelling. Thus, availability and accessibility of LiDAR data should be increased.
- The interior layout of buildings contains important information for energy modelling, but no established methods for the determination of interior layouts currently exist. Thus, suitable methods should be investigated.
- Overcoming privacy and commerciality concerns to increase access to anonymised energy consumption data would provide considerable benefit at little to no additional cost, as smart metering technology is already deployed and these data already exist. Thus, increasing accessibility to these data without compromising consumer privacy should be actively pursued.

5. Conclusions

Recent technological advancements have driven rapid growth in the field of urban energy modeling (UEM), enabling the analysis of interconnected, multi-domain urban energy systems and requiring data inputs from various domains.This work provides a comprehensive review of multi-domain UEM data requirements, including data formats, sources, acquisition methods, bridging methods, and challenges. Data inputs are categorized into climate, geography, building, transportation, demographics, energy networks and consumption, and distributed energy resources. Additionally, several key challenges are identified, which are common to multiple domains. Although specific challenges can vary depending on the requirements of a given model, in general, improving the availability, accessibility, and quality of high-impact data should be considered a priority. Key implications and recommendations for multi-domain UEM data are provided. Overall, substantial amounts of data exist, but their use is encumbered by a lack of coordination and standardisation of formats and due to privacy and commerciality concerns. Consequently, coordinated effort by researchers, data owners, and data collectors is required to increase access to these data, which will improve the results of multi-domain UEMs.

Author Contributions

Conceptualization, D.B.; Methodology, D.B., B.L.M.W.; Investigation, D.B., B.L.M.W.; Writing—Original Draft Preparation, D.B.; Writing—Review & Editing, D.B., P.G., B.L.M.W.

Ethics Statement

Not applicable.

Informed Consent Statement

Not applicable.

Funding

Declaration of Competing Interest

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

References

1. Reinhart CF, Cerezo Davila C. Urban Building Energy Modeling—A Review of a Nascent Field. *Build. Environ.* **2016**, *97*, 196–202.
2. Lin Z, Hong T, Xu X, Chen J, Wang W. Evaluating Energy Retrofits of Historic Buildings in a University Campus Using an Urban Building Energy Model That Considers Uncertainties. *Sustain. Cities Soc.* **2023**, *95*, 104602.
3. Yazdanie M, Orehounig K. Advancing urban energy system planning and modeling approaches: Gaps and solutions in perspective. *Renew. Sustain. Energy Rev.* **2021**, *137*, 110607.
4. Keirstead J, Jennings M, Sivakumar A. A Review of Urban Energy System Models: Approaches, Challenges and Opportunities. *Renew. Sustain. Energy Rev.* **2012**, *16*, 3847–3866.
5. Alhamwi A, Medjroubi W, Vogt T, Agert C. Development of a GIS-based platform for the allocation and optimisation of distributed storage in urban energy systems. *Appl. Energy* **2019**, *251*, 113360.
6. Xu W, Liu S. Novel economic models for advancing urban energy management and transition: Simulation of urban energy system in digital twin. *Sustain. Cities Soc.* **2024**, *101*, 105154.
7. Amado M, Poggi F. Solar energy integration in urban planning: GUUD model. In *Proceedings of the Energy Procedia*; Elsevier Ltd.: Amsterdam, The Netherlands, 2014; pp. 277–284.
8. Li W, Zhou Y, Cetin K, Eom J, Wang Y, Chen G, et al. Modeling urban building energy use: A review of modeling approaches and procedures. *Energy* **2017**, *141*, 2445–2457.
9. Ferrando M, Causone F, Hong T, Chen Y. Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches. *Sustain. Cities Soc.* **2020**, *62*, 102408.
10. Herrera M, Natarajan S, Coley DA, Kershaw T, Ramallo-González AP, Eames M, et al. A review of current and future weather data for building simulation. *Build. Serv. Eng. Res. Technol.* **2017**, *38*, 602–627.

11. Moazami A, Carlucci S, Geving S. Critical analysis of software tools aimed at generating future weather files with a view to their use in building performance simulation. In *Proceedings of the Energy Procedia*; Elsevier Ltd.: Amsterdam, The Netherlands, 2017; pp. 640–645.

12. Bartesaghi Koc C, Osmond P, Peters A. Evaluating the cooling effects of green infrastructure: A systematic review of methods, indicators and data sources. *Sol. Energy* **2018**, *166*, 486–508.

13. Liu Z, Cheng W, Jim CY, Morakinyo TE, Shi Y, Ng E. Heat mitigation benefits of urban green and blue infrastructures: A systematic review of modeling techniques, validation and scenario simulation in ENVI-Met V4. *Build. Environ.* **2021**, *200*, 107939.

14. Liu X. Airborne LiDAR for DEM generation: Some critical issues. *Prog. Phys. Geogr.* **2008**, *32*, 31–49.

15. Gunduz M, Isikdag U, Basaraner M. A review of recent research in indoor modelling & mapping. In *Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*; ISPRS Archives, Leibniz Universität Hannover: Hannover, Germany, 2016; pp. 289–294.

16. O'Hegarty R, Kinnane O, Lennon D, Colclough S. In-situ U-value monitoring of highly insulated building envelopes: Review and experimental investigation. *Energy Build.* **2021**, *252*, 111447.

17. Happle G, Fonseca JA, Schlueter A. A review on occupant behavior in urban building energy models. *Energy Build.* **2018**, *174*, 276–292.

18. Fuentes E, Arce L, Salom J. A review of domestic hot water consumption profiles for application in systems and buildings energy performance analysis. *Renew. Sustain. Energy Rev.* **2018**, *81*, 1530–1547.

19. Rathore N, Panwar NL, Yettou F, Gama A. A comprehensive review of different types of solar photovoltaic cells and their applications. *Int. J. Ambient Energy* **2021**, *42*, 1200–1217.

20. Tummala A, Velamati RK, Sinha DK, Indraja V, Krishna VH. A review on small scale wind turbines. *Renew. Sustain. Energy Rev.* **2016**, *56*, 1351–1371.

21. Tasneem Z, Al Noman A, Das SK, Saha DK, Islam MR, Ali MF, et al. An analytical review on the evaluation of wind resource and wind turbine for urban application: Prospect and challenges. *Dev. Built Environ.* **2020**, *4*, 100033.

22. Rahman MM, Oni AO, Gemechu E, Kumar A. Assessment of energy storage technologies: A review. *Energy Convers. Manag.* **2020**, *223*, 113295.

23. Oraiopoulos A, Howard B. On the accuracy of urban building energy modelling. *Renew. Sustain. Energy Rev.* **2022**, *158*, 111976.

24. Dahlström L, Broström T, Widén J. Advancing urban building energy modelling through new model components and applications: A review. *Energy Build.* **2022**, *266*, 112099.

25. Goy S, Maréchal F, Finn D. Data for urban scale building energy modelling: Assessing impacts and overcoming availability challenges. *Energies* **2020**, *13*, 4244.

26. Ali U, Shamsi MH, Hoare C, Mangina E, O'Donnell J. Review of urban building energy modeling (UBEM) approaches, methods and tools using qualitative and quantitative analysis. *Energy Build.* **2021**, *246*, 111073.

27. Johari F, Peronato G, Sadeghian P, Zhao X, Widén J. Urban building energy modeling: State of the art and future prospects. *Renew. Sustain. Energy Rev.* **2020**, *128*, 109902.

28. Li X, Zhou Y, Yu S, Jia G, Li H, Li W. Urban Heat Island Impacts on Building Energy Consumption: A Review of Approaches and Findings. *Energy* **2019**, *174*, 407–419.

29. Torabi M, Mahdavinejad M. Past and Future Trends on the Effects of Occupant Behaviour on Building Energy Consumption. *J. Sustain Archit. Civ. Eng.* **2021**, *29*, 83–101.

30. Sola A, Corchero C, Salom J, Sanmarti M. Multi-Domain Urban-Scale Energy Modelling Tools: A Review. *Sustain. Cities Soc.* **2020**, *54*, 101872.

31. Nageler P, Koch A, Mauthner F, Leusbrock I, Mach T, Hochenauer C, et al. Comparison of Dynamic Urban Building Energy Models (UBEM): Sigmoid Energy Signature and Physical Modelling Approach. *Energy Build.* **2018**, *179*, 333–343.

32. Hedegaard RE, Kristensen MH, Pedersen TH, Brun A, Petersen S. Bottom-up Modelling Methodology for Urban-Scale Analysis of Residential Space Heating Demand Response. *Appl. Energy* **2019**, *242*, 181–204.

33. Koschwitz D, Frisch J, van Treeck C. Data-Driven Heating and Cooling Load Predictions for Non-Residential Buildings Based on Support Vector Machine Regression and NARX Recurrent Neural Network: A Comparative Study on District Scale. *Energy* **2018**, *165*, 134–142.

34. Williams B, Bishop D, Hooper G, Chase JG. Driving Change: Electric Vehicle Charging Behavior and Peak Loading. *Renew. Sustain. Energy Rev.* **2024**, *189*, 113953.

35. Chen Y, Hong T, Luo X, Hooper B. Development of City Buildings Dataset for Urban Building Energy Modeling. *Energy Build* **2019**, *183*, 252–265.

36. Pasichnyi O, Wallin J, Levihn F, Shahrokni H, Kordas O. Energy Performance Certificates—New Opportunities for Data-Enabled Urban Energy Policy Instruments? *Energy Policy* **2019**, *127*, 486–499.

37. Rodriguez C, Coronado M, D'Alessandro M, Medina J. The Importance of Standardised Data-Collection Methods in the Improvement of Thermal Comfort Assessment Models for Developing Countries in the Tropics. *Sustainability* **2019**, *11*, 4180.

38. Coakley D, Raftery P, Keane M. A Review of Methods to Match Building Energy Simulation Models to Measured Data. *Renew. Sustain. Energy Rev.* **2014**, *37*, 123–141.

39. Nouvel R, Zirak M, Coors V, Eicker U. The Influence of Data Quality on Urban Heating Demand Modeling Using 3D City Models. *Comput. Environ. Urban Syst.* **2017**, *64*, 68–80.

40. Medjroubi W, Müller UP, Scharf M, Matke C, Kleinhans D. Open Data in Power Grid Modelling: New Approaches Towards Transparent Grid Models. *Energy Rep.* **2017**, *3*, 14–21.

41. Manfren M, Nastasi B, Groppi D, Astiaso Garcia D. Open Data and Energy Analytics—An Analysis of Essential Information for Energy System Planning, Design and Operation. *Energy* **2020**, *213*, 118803.

42. Do H, Cetin KS. Residential Building Energy Consumption: A Review of Energy Data Availability, Characteristics, and Energy Performance Prediction Methods. *Curr. Sustain. Renew. Energy Rep.* **2018**, *5*, 76–85.

43. Mathew PA, Dunn LN, Sohn MD, Mercado A, Custudio C, Walter T. Big-Data for Building Energy Performance: Lessons from Assembling a Very Large National Database of Building Energy Use. *Appl. Energy* **2015**, *140*, 85–93.

44. Cerezo DC, Reinhart CF, Bemis JL. Modeling Boston: A Workflow for the Efficient Generation and Maintenance of Urban Building Energy Models from Existing Geospatial Datasets. *Energy* **2016**, *117*, 237–250.

45. Wilcox S, Marion W. Users Manual for TMY3 Data Sets. *Nat. Renew. Energy Lab.* **2008**, *56*, 43156.

46. World Meteorological Organisation. Guide to Instruments and Methods of Observation Volume I—Measurement of Meteorological Variables. Available online: https://library.wmo.int/idurl/4/41650 (accessed on 2 May 2023).

47. Kalogirou SA. Generation of Typical Meteorological Year (TMY-2) for Nicosia, Cyprus. *Renew. Energy* **2003**, *28*, 2317–2334.

48. Remund J, Müller S, Schmutz M, Graf P. Meteonorm Version 8.0, 2020. Available online: www.meteotest.com (accessed on 2 May 2023).

49. Tsoka S, Tolika K, Theodosiou T, Tsikaloudaki K. Evaluation of Stochastically Generated Weather Datasets for Building Energy Simulation. In *Proceedings of the Energy Procedia*; Elsevier Ltd.: Amsterdam, The Netherlands, 2017; pp. 853–858.

50. Rothamsted Research LARS-WG Weather Generator. Available online: https://sites.google.com/view/lars-wg/ (accessed on 10 March 2023).

51. Semenov M, Brooks R. Spatial Interpolation of the LARS-WG Stochastic Weather Generator in Great Britain. *Clim. Res.* **1999**, *11*, 137–148.

52. Varne AR, Blouin S, Williams BLM, Denkenberger D. The Impact of Abrupt Sunlight Reduction Scenarios on Renewable Energy Production. *Earth ArXiv* **2024**, *1*, 7–26.

53. Zhou D, Zhao S, Liu S, Zhang L, Zhu C. Surface Urban Heat Island in China's 32 Major Cities: Spatial Patterns and Drivers. *Remote Sens. Environ.* **2014**, *152*, 51–61.

54. Hong T, Chen Y, Luo X, Luo N, Lee SH. Ten Questions on Urban Building Energy Modeling. *Build. Environ.* **2020**, *168*, 106508.

55. ENVI_MET ENVI-Met Software. Available online: https://www.envi-met.com/ (accessed on 10 March 2023).

56. Urban Microclimate Urban Weather Generator 4.1. Available online: https://urbanmicroclimate.scripts.mit.edu/uwg.php (accessed on 10 March 2023).

57. Nakano A, Bueno B, Norford L, Reinhart CF. Urban Weather Generator- A Novel Workflow for Integrating Urban Heat Island Effect within Urban Design Process. *Build. Simul.* **2015**, *1*, BS2015.

58. Buckley N, Mills G, Letellier-Duchesne S, Benis K. Designing an Energy-Resilient Neighbourhood Using an Urban Building Energy Model. *Energies* **2021**, *14*, 4445.

59. Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, et al. Climate Change 2021: The Physical Science Basis. *WGI Sixth Assess. Rep.* **2021**, *2*, 2391.

60. University of Southampton—Energy & Climate Climate Change World Weather File Generator for World-Wide Weather Data—CCWorldWeatherGen. Available online: https://energy.soton.ac.uk/climate-change-world-weather-file-generator-for-world-wide-weather-data-ccworldweathergen/ (accessed on 10 March 2023).

61. Arup WeatherShift v2.0. Available online: https://www.arup.com/services/digital-solutions-and-tools/weathershift/ (accessed on 2 May 2023).

62. Girindran R, Boyd DS, Rosser J, Vijayan D, Long G, Robinson D. On the Reliable Generation of 3D City Models from Open Data. *Urban Sci.* **2020**, *4*, 47.

63. Wang C, Ferrando M, Causone F, Jin X, Zhou X, Shi X. Data Acquisition for Urban Building Energy Modeling: A Review. *Build. Environ.* **2022**, *217*, 109056.

64. Alganci U, Besol B, Sertel E. Accuracy Assessment of Different Digital Surface Models. *Int. J. Geoinf.* **2018**, *7*, 114.

65. Partovi T, Fraundorfer F, Bahmanyar R, Huang H, Reinartz P. Automatic 3-D Building Model Reconstruction from Very High Resolution Stereo Satellite Imagery. *Remote Sens.* **2019**, *11*, 1660.

66. Wang P, Huang C, Tilton JC. Mapping Three-Dimensional Urban Structure by Fusing Landsat and Global Elevation Data. *ArXiv* **2018**, *1807*, 04368.

67. Earth Resources Observation and Science (EROS) Center EROS Archive—Digital Elevation—Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global. Available online: https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1 (accessed on 14 March 2023).

68. NASA Jet Propulsion Laboratory ASTER Global Digital Elevation Map. Available online: https://asterweb.jpl.nasa.gov/gdem.asp (accessed on 14 March 2023).

69. Advanced Land Observing Satellite ALOS Global Digital Surface Model "ALOS World 3D—30m (AW3D30)". Available online: https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/aw3d30_e.htm (accessed on 14 March 2023).

70. United Stated Geological Survey GMTED2010. Available online: https://www.usgs.gov/coastal-changes-and-impacts/gmted2010 (accessed on 14 March 2023).

71. Tooke TR, Coops NC, Webster J. Predicting Building Ages from LiDAR Data with Random Forests for Building Energy Modeling. *Energy Build.* **2014**, *68*, 603–610.

72. O'Donnell J, Truong-Hong L, Boyle N, Corry E, Cao J, Laefer DF. LiDAR Point-Cloud Mapping of Building Façades for Building Energy Performance Simulation. *Autom. Constr.* **2019**, *107*, 102905.

73. Sun S, Salvaggio C. Aerial 3D Building Detection and Modeling from Airborne LiDAR Point Clouds. *J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1440–1449.

74. OpenTopography OpenTopography. Available online: https://opentopography.org/ (accessed on 14 March 2023).

75. Land Information New Zealand Elevation Data. Available online: https://www.linz.govt.nz/products-services/data/types-linz-data/elevation-data (accessed on 14 March 2023).

76. Ayad A, Fahmy M, Kamel W. Urban Green and Blue Infrastructure Simulation in a Changing Climate from Microclimate to Energy Consumption: A Case Study in Alexandria, Egypt. In *Proceedings of the Building Simulation Conference Proceedings*; International Building Performance Simulation Association, Routledge: London, UK, 2019; pp. 3499–3507.

77. Žuvela-Aloise M, Koch R, Buchholz S, Früh B. Modelling the Potential of Green and Blue Infrastructure to Reduce Urban Heat Load in the City of Vienna. *Clim. Chang.* **2016**, *135*, 425–438.

78. Stewart ID, Oke TR. Local Climate Zones for Urban Temperature Studies. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 1879–1900.

79. Ching J, Mills G, Bechtel B, See L, Feddema J, Wang X, et al. WUDAPT: An Urban Weather, Climate, and Environmental Modeling Infrastructure for the Anthropocene. *Bull. Am. Meteorol. Soc.* **2018**, *99*, 1907–1924.

80. WUDAPT World Urban Database and Access Portal Tools. Available online: https://www.wudapt.org/ (accessed on 15 March 2023).

81. Demuzere M, Kittner J, Bechtel B. LCZ Generator: A Web Application to Create Local Climate Zone Maps. *Front. Environ. Sci.* **2021**, *9*, 637455.

82. Bande L, Manandhar P, Ghazal R, Marpu P. Characterization of Local Climate Zones Using ENVI-Met and Site Data in the City of Al-Ain, UAE. *Int. J. Sustain. Dev. Plan.* **2020**, *15*, 751–760.

83. Harish VSKV, Kumar A. A Review on Modeling and Simulation of Building Energy Systems. *Renew. Sustain. Energy Rev.* **2016**, *56*, 1272–1292.

84. Kamel E. A Systematic Literature Review of Physics-Based Urban Building Energy Modeling (UBEM) Tools, Data Sources, and Challenges for Energy Conservation. *Energies* **2022**, *15*, 8649.

85. EnergyPlus Documentation. Available online: https://energyplus.net/documentation (accessed on 15 April 2024).

86. BuildingSMARTInternational Industry Foundation Classes (IFC)—An Introduction. Available online: https://technical.buildingsmart.org/standards/ifc/ (accessed on 17 March 2023).

87. Green Building XML (gbXML) Schema GbXML. Available online: https://www.gbxml.org/ (accessed on 17 March 2023).

88. Internet Engineering Task Force (IETF) The GeoJSON Format. Available online: https://www.rfc-editor.org/rfc/rfc7946 (accessed on 17 March 2023).

89. Gruchalla K, Fleming K, Gruchalla K, Fleming K. *Building Simulation Conference: Building Performance Modeling*; SimBuild: Salt Lake City, UT, USA, 2016; pp. 133–140.

90. ESRI Reference—Shapefiles. Available online: https://doc.arcgis.com/en/arcgis-online/reference/shapefiles.htm (accessed on 17 March 2023).

91. Open Geospatial Consortium CityGML—Overview. Available online: https://www.ogc.org/standard/citygml/ (accessed on 17 March 2023).

92. Open Geospatial Consortium. OGC City Geography Markup Language (CityGML) Part 1 Conceptual Model Standard, 2023. Available online: https://www.ogc.org/standard/gml/ (accessed on 2 May 2023).

93. Agugiaro G, Benner J, Cipriano P, Nouvel R. The Energy Application Domain Extension for CityGML: Enhancing Interoperability for Urban Energy Simulations. *Open Geospat. Data Softw. Stand.* **2018**, *3*, 1–30.

94. CityGMLwiki CityGML UtilityNetworkADE. Available online: https://www.citygmlwiki.org/index.php?title=CityGML_UtilityNetworkADE (accessed on 17 March 2023).

95. Bishop D, Mohkam M, Williams BLM, Wu W, Bellamy L. The Impact of Building Level of Detail Modelling Strategies: Insights into Building and Urban Energy Modelling. *Energies* **2024**, *5*, 2280–2299.

96. Wang D, Landolt J, Mavromatidis G, Orehounig K, Carmeliet J. CESAR: A Bottom-up Building Stock Modelling Tool for Switzerland to Address Sustainable Energy Transformation Strategies. *Energy Build.* **2018**, *169*, 9–26.

97. Faisal K, Imam A, Majrashi A, Hegazy I. Machine Learning Approach to Extract Building Footprint from High-Resolution Images: The Case Study of Makkah, Saudi Arabia. *Int. J. Low Carbon Technol.* **2021**, *16*, 655–663.

98. Google Open Buildings. Available online: https://sites.research.google/open-buildings/#faq (accessed on 20 March 2023).

99. Sirko W, Kashubin S, Ritter M, Annkah A, Salah Y, Bouchareb E, et al. Continental-scale building detection from high resolution satellite imagery. *ArXiv* **2021**, *2*, 12286.

100. OpenStreetMap Foundation OpenStreetMap. Available online: https://www.openstreetmap.org/ (accessed on 20 March 2023).

101. Zhuo X, Fraundorfer F, Kurz F, Reinartz P. Optimization of OpenStreetMap Building Footprints Based on Semantic Information of Oblique UAV Images. *Remote Sens.* **2018**, *10*, 264.

102. Katal A, Mortezazadeh M, Wang L, Yu H. Urban Building Energy and Microclimate Modeling—From 3D City Generation to Dynamic Simulations. *Energy* **2022**, *251*, 123817.

103. Schiefelbein J, Rudnick J, Scholl A, Remmen P, Fuchs M, Müller D. Automated Urban Energy System Modeling and Thermal Building Simulation Based on OpenStreetMap Data Sets. *Build. Environ.* **2019**, *149*, 630–639.

104. Verellen E, Allacker K. Developing a Building Stock Model to Enable Clustered Renovation—The City of Leuven as Case Study. *Sustainability* **2022**, *14*, 5769.

105. Data San Francisco DataSF OpenData. Available online: Data San Francisco (accessed on 20 March 2023).

106. City of New York NYC OpenData. Available online: https://opendata.cityofnewyork.us/ (accessed on 20 March 2023).

107. City of Chicago Chicago Data Portal. Available online: https://data.cityofchicago.org/ (accessed on 20 March 2023).

108. Open Data Berlin Berlin3D. Available online: https://download-berlin3d.virtualcitymap.de/documents/manual.en.html#allgemein (accessed on 20 March 2023).

109. Comune di Milano Portale Open Data. Available online: https://dati.comune.milano.it/ (accessed on 20 March 2023).

110. Federal Office of Topography Swisstopo. Available online: https://www.swisstopo.admin.ch/ (accessed on 20 March 2023).

111. Caro-Martínez R, Sendra JJ. Implementation of Urban Building Energy Modeling in Historic Districts. Seville as Case-Study. *Int. J. Sustain. Dev. Plan.* **2018**, *13*, 528–540.

112. Faure X, Johansson T, Pasichnyi O. The Impact of Detail, Shadowing and Thermal Zoning Levels on Urban Building Energy Modelling (UBEM) on a District Scale. *Energies* **2022**, *15*, 1525.

113. Roth J, Martin A, Miller C, Jain RK. SynCity: Using Open Data to Create a Synthetic City of Hourly Building Energy Estimates by Integrating Data-Driven and Physics-Based Methods. *Appl. Energy* **2020**, *280*, 115981.

114. Li W, Zhou Y, Cetin KS, Yu S, Wang Y, Liang B. Developing a Landscape of Urban Building Energy Use with Improved Spatiotemporal Representations in a Cool-Humid Climate. *Build. Environ.* **2018**, *136*, 107–117.

115. Wang C, Wei S, Du S, Zhuang D, Li Y, Shi X, et al. A Systematic Method to Develop Three Dimensional Geometry Models of Buildings for Urban Building Energy Modeling. *Sustain. Cities. Soc.* **2021**, *71*, 102998.

116. Misra P, Avtar R, Takeuchi W. Comparison of Digital Building Height Models Extracted from AW3D, TanDEM-X, ASTER, and SRTM Digital Surface Models over Yangon City. *Remote Sens.* **2018**, *10*, 12.

117. Zhou J, Liu Y, Nie G, Cheng H, Yang X, Chen X, et al. Building Extraction and Floor Area Estimation at the Village Level in Rural China Via a Comprehensive Method Integrating UAV Photogrammetry and the Novel EDSANet. *Remote Sens.* **2022**, *14*, 1575.

118. Qi F, Zhai JZ, Dang G. Building Height Estimation Using Google Earth. *Energy Build.* **2016**, *118*, 123–132.

119. Chen Y, Hong T, Piette MA. Automatic Generation and Simulation of Urban Building Energy Models Based on City Datasets for City-Scale Building Retrofit Analysis. *Appl. Energy* **2017**, *205*, 323–335.

120. Dochev I, Gorzalka P, Weiler V, Estevam SJ, Linkiewicz M, Eicker U, et al. Calculating Urban Heat Demands: An Analysis of Two Modelling Approaches and Remote Sensing for Input Data and Validation. *Energy Build.* **2020**, *226*, 110378.

121. Cerezo C, Sokol J, AlKhaled S, Reinhart C, Al-Mumin A, Hajiah A. Comparison of Four Building Archetype Characterization Methods in Urban Building Energy Modeling (UBEM): A Residential Case Study in Kuwait City. *Energy Build.* **2017**, *154*, 321–334.

122. Cao J, Metzmacher H, O'Donnell J, Frisch J, Bazjanac V, Kobbelt L, et al. Facade Geometry Generation from Low-Resolution Aerial Photographs for Building Energy Modeling. *Build. Environ.* **2017**, *123*, 601–624.

123. Cao J, Metzmacher H, O'Donnell J, Kobbelt L, van Treeck C. BIM Geometry Generation from Low-Resolution Aerial Photographs for Building Energy Modeling. In Proceedings of BS2015: 14th Conference of International Building Performance Simulation Association, Hyderabad, India, 7–9 December 2015; pp. 512–519.

124. Ramallo-González AP, Vellei M, Brown M, Coley DA. Remote Facade Surveying of Windows Characteristics. In *Proceedings of the Energy Procedia*, Elsevier Ltd.: Amsterdam, The Netherlands, 2015; pp. 925–930.

125. Yang X, Qin X, Wang J, Wang J, Ye X, Qin Q. Building Façade Recognition Using Oblique Aerial Images. *Remote Sens.* **2015**, *7*, 10562–10588.

126. Zhou B, Ma W, Li Q, El-Sheimy N, Mao Q, Li Y, et al. Crowdsourcing-Based Indoor Mapping Using Smartphones: A Survey. *J. Photogramm. Remote Sens.* **2021**, *177*, 131–146.

127. Google Maps Google Indoor Maps. Available online: https://www.google.com/maps/about/partners/indoormaps/ (accessed on 23 March 2023).

128. Buckley N, Mills G, Reinhart C, Berzolla ZM. Using Urban Building Energy Modelling (UBEM) to Support the New European Union's Green Deal: Case Study of Dublin Ireland. *Energy Build.* **2021**, *247*, 111115.

129. Sokol J, Cerezo Davila C, Reinhart CF. Validation of a Bayesian-Based Method for Defining Residential Archetypes in Urban Building Energy Models. *Energy Build.* **2017**, *134*, 11–24.

130. Pasichnyi O, Wallin J, Kordas O. Data-Driven Building Archetypes for Urban Building Energy Modelling. *Energy* **2019**, *181*, 360–377.

131. Nageler P, Zahrer G, Heimrath R, Mach T, Mauthner F, Leusbrock I, et al. Novel Validated Method for GIS Based Automated Dynamic Urban Building Energy Simulations. *Energy* **2017**, *139*, 142–154.

132. Nouvel R, Schulte C, Eicker U, Pietruschka D, Coors V. CITYGML-Based 3D City Model for Energy Diagnostics and Urban Energy Policy Support. In Proceedings of the Proceedings of BS 2013: 13th Conference of the International Building Performance Simulation Association, Chambéry, France, 25–28 August 2013; pp. 218–225.

133. Lazaro M, Lazaro A, González B, Villarino R, Girbau D. Long-Range Wireless System for U-Value Assessment Using a Low-Cost Heat Flux Sensor. *Sensors* **2022**, *22*, 7259.

134. Mobaraki B, Castilla Pascual FJ, Lozano-Galant F, Lozano-Galant JA, Porras Soriano R. In Situ U-Value Measurement of Building Envelopes through Continuous Low-Cost Monitoring. *Case Stud. Therm. Eng.* **2023**, *43*, 102778.

135. Patel D, Schmiedt JE, Röger M, Hoffschmidt B. Approach for External Measurements of the Heat Transfer Coefficient (U-value) of Building Envelope Components Using UAV-Based Infrared Thermography. In Proceedings of the 14th Quantitative Infrared Thermography Conference (QIRT). Institute of Solar Research, Berlin, Germany, 25–29 June 2018.

136. Dimitrov A, Golparvar-Fard M. Vision-Based Material Recognition for Automated Monitoring of Construction Progress and Generating Building Information Modeling from Unordered Site Image Collections. *Adv. Eng. Inform.* **2014**, *28*, 37–49.

137. Dabirian S, Panchabikesan K, Eicker U. Occupant-Centric Urban Building Energy Modeling: Approaches, Inputs, and Data Sources—A Review. *Energy Build.* **2022**, *257*, 111809.

138. Khan I. Household Factors and Electrical Peak Demand: A Review for Further Assessment. *Adv. Build. Energy Res.* **2021**, *15*, 409–441.

139. Tanimoto J, Hagishima A, Sagara H. A Methodology for Peak Energy Requirement Considering Actual Variation of Occupants' Behavior Schedules. *Build. Environ.* **2008**, *43*, 610–619.

140. Williams B, Bishop D. Flexible Futures: The Potential for Electricity Demand Response in New Zealand. Available online: https://ssrn.com/abstract=4615974 (accessed on 5 March 2024).

141. Carlucci S, De Simone M, Firth SK, Kjærgaard MB, Markovic R, Rahaman MS, et al. Modeling Occupant Behavior in Buildings. *Build. Environ.* **2020**, *174*, 106768.

142. Kontar R, El Rakha T. Profiling Occupancy Patterns in Community-Scale Residential Buildings Using Measured Energy Use Data Clustering. In Proceedings of the 2018 Building Performance Analysis Conference and SimBuild. ASHRAE and IBPSA, Chicago, USA, 26–28 September 2018.

143. Liao J, Stankoviz L, Stankovic V. Detecting Household Activity Patterns from Smart Meter Data. In *Proceedings of the Proceedings—2014 International Conference on Intelligent Environments*; Institute of Electrical and Electronics Engineers Inc.: Venice, Italy, 2014; pp. 71–78.

144. Knebel DE. *Energy-Efficient Design of Low-Rise Residential Buildings*; American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. ASHRAE Standard; ASHRAE: Atlanta, GA, USA, 2007.

145. Hazim A. *AM11 Building Performance Modelling;* Chartered Institution of Building Services Engineers: London, UK, 2015.

146. Ministry of Business. H1 Energy Efficiency Verification Method H1/VM1, 2023. Available online: https://www.building.govt.nz/assets/Uploads/building-code-compliance/h1-energy-efficiency/asvm/h1-energy-efficiency-vm1-5th-edition-amendment-1.pdf (accessed on 5 March 2024).

147. Malekpour Koupaei D, Cetin KS, Passe U. Stochastic Residential Occupancy Schedules Based on the American Time-Use Survey. *Sci Technol. Built Environ.* **2022**, *28*, 776–790.

148. Baetens R, Saelens D. Modelling Uncertainty in District Energy Simulations by Stochastic Residential Occupant Behaviour. *J. Build Perform. Simul.* **2016**, *9*, 431–447.

149. Verbruggen S, Delghust M, Laverge J, Janssens A. Stochastic Occupant Behavior Model Based on Activity And Occupancy Patterns. In Proceedings of the Proceedings of Building Simulation 2019: 16th Conference of IBPSA. International Building Performance Simulation Association, Hazel lane, SD, USA, 2–4 September 2020; pp. 2310–2317.

150. Norouziasl S, Jafari A, Wang C. An Agent-Based Simulation of Occupancy Schedule in Office Buildings. *Build Environ.* **2020**, *186*, 107352.

151. Clemente S, Beauchêne S, Nefzaoui E. Generation of Aggregated Plug Load Profiles in Office Buildings. *Energy Build* **2021**, *252*, 111398.

152. Jordan U, Vajen K. DHWcalc: Program to Generate Domestic Hot Water Profiles with Statistical Means for User Defined Conditions. In Proceedings of the ISES Solar World Congress, Orlando, FL, USA, 6–12 August 2005.

153. Noah Pflugradt Load Profile Generator. Available online: https://www.loadprofilegenerator.de/ (accessed on 5 April 2023).

154. US DOE Building Performance Database. Available online: https://bpd.lbl.gov/ (accessed on 31 March 2023).

155. Mathur A, Fennell P, Rawal R, Korolija I. Assessing a Fit-for-Purpose Urban Building Energy Modelling Framework with Reference to Ahmedabad. *Sci. Technol. Built Environ.* **2021**, *27*, 1075–1103.

156. Kavousian A, Rajagopal R, Fischer M. Ranking Appliance Energy Efficiency in Households: Utilizing Smart Meter Data and Energy Efficiency Frontiers to Estimate and Identify the Determinants of Appliance Energy Efficiency in Residential Buildings. *Energy Build* **2015**, *99*, 220–230.

157. Guo Z, Wang ZJ, Kashani A. Home Appliance Load Modeling from Aggregated Smart Meter Data. *Trans. Power Syst.* **2015**, *30*, 254–262.

158. Dinesh C, Nettasinghe BW, Godaliyadda RI, Ekanayake MPB, Ekanayake J, Wijayakulasooriya JV. Residential Appliance Identification Based on Spectral Information of Low Frequency Smart Meter Measurements. *Trans. Smart Grid* **2016**, *7*, 2781–2792.

159. Liao J, Elafoudi G, Stankovic L, Stankovic V. Non-Intrusive Appliance Load Monitoring Using Low-Resolution Smart Meter Data. In *Proceedings of the 2014 IEEE International Conference on Smart Grid Communications*; Institute of Electrical and Electronics Engineers Inc.: Venice, Italy, 2015; pp. 535–540.

160. Zhong M, Shan R, Du D, Lu C. A Comparative Analysis of Traditional Four-Step and Activity-Based Travel Demand Modeling: A Case Study of Tampa, Florida. *Transp. Plan. Technol.* **2015**, *38*, 517–533.

161. Guida C, Caglioni M. Urban Accessibility: The Paradox, the Paradigms and the Measures. A Scientific Review. *J. Land Use Mobil. Environ.* **2020**, *13*, 149–168.

162. Pereira RHM, Herszenhut D. *Introduction to Urban Accessibility: A Practical Guide with R;* Instituto de Pesquisa Econômica Aplicada: Rio de Janeiro, Brazil, 2023.

163. Lowans C, Furszyfer Del Rio D, Sovacool BK, Rooney D, Foley AM. What Is the State of the Art in Energy and Transport Poverty Metrics? A Critical and Comprehensive Review. *Energy Econ.* **2021**, *101*, 105360.

164. Saunders MJ, Kuhnimhof T, Chlond B, da Silva ANR. Incorporating Transport Energy into Urban Planning. *Transp. Res. Part. A. Policy Pract.* **2008**, *42*, 874–882.

165. Buehler R. Determinants of Transport Mode Choice: A Comparison of Germany and the USA. *J. Transp. Geogr.* **2011**, *19*, 644–657.

166. Tamin OZ, Willumsen LG. Transport Demand Model Estimation from Traffic Counts. *Transportation* **1989**, *16*, 3–26.

167. Aletta F, Brinchi S, Carrese S, Gemma A, Guattari C, Mannini L, et al. Analysing Urban Traffic Volumes and Mapping Noise Emissions in Rome (Italy) in the Context of Containment Measures for the COVID-19 Disease. *Noise Mapp.* **2020**, *7*, 114–122.

168. Acheampong RA, Silva EA. Land Use–Transport Interaction Modeling: A Review of the Literature and Future Research Directions. *J. Transp. Land Use* **2015**, *8*, 11–38.

169. MobilityData General Transit Feed Specification. Available online: https://gtfs.org/ (accessed on 16 March 2024).

170. Jomehpour Chahar AJ, Smith-Colin J. Transit Deserts: Equity Analysis of Public Transit Accessibility. *J. Transp. Geogr.* **2020**, *89*, 102869.

171. Harrison G, Grant-Muller SM, Hodgson FC. New and Emerging Data Forms in Transportation Planning and Policy: Opportunities and Challenges for "Track and Trace" Data. *Transp. Res. Part C. Emerg. Technol.* **2020**, *117*, 102672.

172. Torre-Bastida AI, Del Ser J, Laña I, Ilardia M, Bilbao MN, Campos-Cordobés S. Big Data for Transportation and Mobility: Recent Advances, Trends and Challenges. *Intell. Transp. Syst.* **2018**, *12*, 742–755.

173. Maghrebi M, Abbasi A, Waller ST. Transportation Application of Social Media: Travel Mode Extraction. In *Proceedings of the IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*; Institute of Electrical and Electronics Engineers Inc.: Venice, Italy, 2016; pp. 1648–1653.

174. Rashidi TH, Abbasi A, Maghrebi M, Hasan S, Waller TS. Exploring the Capacity of Social Media Data for Modelling Travel Behaviour: Opportunities and Challenges. *Transp. Res. Part C Emerg. Technol.* **2017**, *75*, 197–211.

175. Feng X, Hu J. Research on the Identification and Management of Vehicle Behaviour Based on Internet of Things Technology. *Comput. Commun.* **2020**, *156*, 68–76.

176. Bandaragoda T, Adikari A, Nawaratne R, Nallaperuma D, Luhach AK, Kempitiya T, et al. Artificial Intelligence Based Commuter Behaviour Profiling Framework Using Internet of Things for Real-Time Decision-Making. *Neural Comput. Appl.* **2020**, *32*, 16057–16071.

177. Brounen D, Kok N, Quigley JM. Residential Energy Use and Conservation: Economics and Demographics. *Eur. Econ. Rev.* **2012**, *56*, 931–945.

178. Sardianou E. Estimating Space Heating Determinants: An Analysis of Greek Households. *Energy Build.* **2008**, *40*, 1084–1093.

179. Wang Y, Wu T, Li H, Skitmore M, Su B. A Statistics-Based Method to Quantify Residential Energy Consumption and Stock at the City Level in China: The Case of the Guangdong-Hong Kong-Macao Greater Bay Area Cities. *J. Clean. Prod.* **2020**, *251*, 119637.

180. Zhao F, Martinez-Moyano IJ, Augenbroe G. Agent-Based Modeling of Commercial Building Stocks for Policy Support. In Proceedings of Building Simulation 2011, 12th Conference of the International Building Performance Simulation Association, Sydney, Australia, 14–16 November 2011.

181. Yang Z, Roth J, Jain RK. DUE-B: Data-Driven Urban Energy Benchmarking of Buildings Using Recursive Partitioning and Stochastic Frontier Analysis. *Energy Build.* **2018**, *163*, 58–69.

182. Abbasabadi N, Ashayeri M, Azari R, Stephens B, Heidarinejad M. An Integrated Data-Driven Framework for Urban Energy Use Modeling (UEUM). *Appl. Energy* **2019**, *253*, 113550.

183. Williams B, Bishop D, Gallardo P, Chase JG. Demand Side Management in Industrial, Commercial, and Residential Sectors: A Review of Constraints and Considerations. *Energies* **2023**, *16*, 5155.

184. Stephenson J, Ford R, Nair NK, Watson N, Wood A, Miller A. Smart Grid Research in New Zealand—A Review from the GREEN Grid Research Programme. *Renew. Sustain. Energy. Rev.* **2018**, *82*, 1636–1645.

185. Chen W, Zhou K, Yang S, Wu C. Data Quality of Electricity Consumption Data in a Smart Grid Environment. *Renew. Sustain. Energy Rev.* **2017**, *75*, 98–105.

186. Klemenjak C, Reinhardt A, Pereira L, Makonin S, Bergés M, Elmenreich W. Electricity Consumption Data Sets: Pitfalls and Opportunities. In Proceedings of the BuildSys 2019—Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, Association for Computing Machinery, New York, NY, USA, 13–14 November 2019; pp. 159–162.

187. PowerCo Utility-Scale Generation. Available online: https://www.powerco.co.nz/get-connected/utility-scale-generation (accessed on 16 March 2024).

188. Electricity Authority Electricity Market Information. Available online: https://www.emi.ea.govt.nz/ (accessed on 18 April 2023).

189. World Bank Group Global. Wind Altas. Available online: https://globalwindatlas.info/en (accessed on 28 April 2023).

190. World Bank Group Global. Solar Atlas. Available online: https://globalsolaratlas.info/ (accessed on 28 April 2023).

191. NREL OpenEI. Available online: https://openei.org/wiki/Main_Page (accessed on 2 May 2023).

192. European Union Data Europa. Available online: https://data.europa.eu/en (accessed on 2 May 2023).

193. Ehsan A, Yang Q. Optimal Integration and Planning of Renewable Distributed Generation in the Power Distribution Networks: A Review of Analytical Techniques. *Appl. Energy* **2018**, *210*, 44–59.

194. HOMER Software HOMER Energy. Available online: https://www.homerenergy.com/ (accessed on 2 May 2023).

195. NREL System Advisor Model. Available online: https://sam.nrel.gov/ (accessed on 2 May 2023).

196. Fu R, Feldman D, Margolis R. U.S. Solar Photovoltaic System Cost Benchmark, 2018. Available online: https://www.nrel.gov/docs/fy21osti/77324.pdf (accessed on 3 May 2023).

197. Directorate-General for Energy Database of the European Energy Storage Technologies and Facilities. Directorate-General for Energy 2023. Available online: https://commission.europa.eu/select-language?destination=/node/13989 (accessed 5 March 2024).

198. Williams B, Bishop D, Docherty P. Assessing the Energy Storage Potential of Electric Hot Water Cylinders with Stochastic Model-Based Control. *J. R. Soc.* **2023**, *2*, 1–17.

199. Bishop D, Nankivell T, Williams B. Peak Loads vs. Cold Showers: The Impact of Existing and Emerging Hot Water Controllers on Load Management. *J. R. Soc.* **2023**, *1*, 1–26.

200. Energy Information Administration. U.S. Cost and Performance Characteristics of New Generating Technologies, Annual Energy Outlook, 2022. Available online: html?file=https%3A%2F%2Fwww.eia.gov%2Foutlooks%2Faeo%2Fassumptions%2Fpdf%2Ftable_8.2.pdf (accessed on 5 March 2023).

201. Lazard. Lazard's Levelized Cost of Energy Analysis, 2023. Available online: https://www.lazard.com/research-insights/levelized-cost-of-energyplus/ (accessed on 5 March 2024).

202. Moomaw W, Burgherr P, Heath G, Lenzen M, Nyboer J, Verbruggen A. Annex II: Methodology. In *IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation*; Edenhofer O, Pichs-Madruga R, Sokona Y, Seyboth K, Matschoss P, Kadner S, et al., Eds.; Cambridge University Press: Cambridge, UK, 2011.

203. IPCC Emissions Factor Database. Available online: https://www.ipcc-nggip.iges.or.jp/EFDB/main.php (accessed on 2 May 2023).