Advances in Applied Energy 3 (2021) 100038

Contents lists available at ScienceDirect

Advances in Applied Energy

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

Exploring the influence of urban context on building energy retrofit performance: A hybrid simulation and data-driven approach

Alex Nutkiewicz*, Benjamin Choi, Rishee K. Jain

Urban Informatics Lab, Department of Civil & Environmental Engineering, Stanford University, 473 Via Ortega, Stanford, CA 94305, United States

A R T I C L E I N F O

Keywords:
Urban building energy model
Deep learning
Retrofit analysis
Inter-building effect
Urban context
Energy optimization

A B S T R A C T

Cities are an integral part to meeting the world's sustainable energy goals. Specifically, retrofits have been implemented to improve energy efficiency and reduce carbon emissions in the buildings sector. Recent simulation, reduced-order, and data-driven approaches have been used to predict the current energy consumption of urban buildings. However, these efforts are limited in their ability to evaluate potential impacts of future retrofits as they are unable to account for inter-building energy interactions that can influence urban building energy performance. To overcome these limitations, we extend a previously developed hybrid data-driven urban energy simulation (DUE-S) model that leverages building energy simulations and deep learning models by now predicting the impact of various building energy retrofits on multiple spatiotemporal scales across a city. We evaluate this approach on a case study of 29 densely co-located buildings in downtown Sacramento, California, USA. Our results indicate that accounting for urban context can compound the impact of retrofits on individual buildings by up to 7.4% as they also influence the electricity use of their surroundings. Finally, we show how DUE-S can provide insights on how to select buildings for retrofit that captures a potential compounding energy savings effect. We develop a greedy optimization algorithm that minimizes the number of required retrofits needed to achieve maximal energy savings across an urban study area. As a result, this work underscores how a flexible urban energy prediction model such as DUE-S can help inform energy-related decisions for a variety of urban-minded stakeholders including architects, engineers, planners, and policymakers.

1. Introduction

Cities are epicenters of social, cultural, and economic activity. Over half of the world's population currently resides in cities, and that number is expected to increase to over 68% by 2050 – requiring the creation and expansion of hundreds of urban areas globally [1]. As a result of rapid urbanization, cities now account for over 75% of all primary energy use and 80% of all greenhouse gas emissions [2]. To address the challenges of energy security, climate change, and economic growth, 77 countries and over 100 cities have begun instituting aggressive emissions reductions targets in conjunction with the COP 21 Paris Climate Accord [3]. A prime opportunity to reduce urban greenhouse gas emissions is in the buildings sector, which currently accounts for over 40% of all primary energy consumption [4]. Furthermore, a significant amount of the building stock that will exist in 2050 has already been constructed [5]; thus, to improve the energy efficiency of the existing building stock, retrofits are key to achieving ambitious climate targets set by urban sustainability stakeholders.

Existing approaches to assess the energy impacts of building retrofit programs are primarily done on an individual building scale using either simulation-based or data-driven models. It is well understood that building energy use is significantly influenced by inter-building effects (e.g., reflection can cause glare and heat gain in surrounding buildings, urban shade from buildings can cool their surroundings) [6] and microclimatic effects (e.g., urban heat island can trap heat within cities and cause increased cooling demand) [7,8]. While simulation-based models can reproduce the underlying thermodynamic effects of various retrofits on individual building energy use, it remains a challenge for these models to account for inter-building energy dynamics and urban context. And although emerging machine learning-based approaches can uncover hidden temporal patterns within building energy data, they struggle to forecast the effects of proposed retrofits due to the lack of available observations needed to train a prediction model [9]. Without accurate characterization for how large-scale retrofits may perform in an urban area, the decisions made by relevant stakeholders may have unintended energy, emissions, and economic consequences [10] that shape the life cycle of both retrofitted buildings and their surrounding neighbors.

Ongoing developments in sensing technologies and open data initiatives have led to a windfall of data streams describing the urban built environment. Existing approaches to characterize the influence of large-

* Corresponding author: Alex Nutkiewicz, 473 Via Ortega Room 271B, Stanford, CA, 94305, United States.
E-mail address: alexer@stanford.edu (A. Nutkiewicz).

https://doi.org/10.1016/j.adapen.2021.100038
Received 18 May 2021; Accepted 20 May 2021
Available online 25 May 2021

scale urban building retrofits on energy consumption are typically done through traditional physics-based modeling or emerging algorithms in machine learning. However, while often treated as two separate fields, our work demonstrates how these approaches can in fact be complementary to one another. Physical representations of buildings through simulation are interpretable and allow for extrapolating what may happen under non-observed conditions, like future building retrofits. Machine learning approaches can extract hidden temporal patterns and insights from large urban data streams. The synergy of these modeling approaches has already been demonstrated in the Earth systems science domain [11], where the outputs of physical models are used as features for deep learning algorithms to detect extreme weather events [12], predict lake temperatures [13], and monitor regional vegetation systems [14]. However, this modeling approach has yet to take hold in the buildings domain, where integrating these two methods could allow us to thermodynamically understand and more accurately predict the complex spatiotemporal nature of building energy use.

The primary objective of this paper is to expand upon a previously developed scalable data-driven urban energy simulation (DUE-S) model. To do so, we modify the deep learning architecture to better account for the time-series patterns present in building energy consumption data and add the ability for DUE-S to estimate the effects of large-scale retrofit programs on multiple spatiotemporal energy scales across a city. To meet their aggressive climate goals, cities must have interpretable, quantitative tools to evaluate the existing energy performance of their buildings and be able to evaluate various strategies to reduce urban energy use and subsequent environmental impacts. Characterizations of urban building energy use at multiple temporal (e.g., daily, monthly, annually) and spatial (e.g., single building, urban) scales through a DUE-S model can help inform the assessment of existing and planned urban energy scenarios. The rest of this paper is organized as follows: Section 2 presents and overview of existing work on urban building energy modeling for retrofit assessment and discusses the primary literature gaps; Section 3 discusses the integrated DUE-S modeling approach and introduces how it can be extended for retrofit analysis; Section 4 provides an overview of our case study of 29 commercial and mixed-use buildings in downtown Sacramento, California, USA; Section 5 discusses the results of our case study; Section 6 outlines the limitations and future work, and Section 7 concludes the paper.

2. Background

To support large-scale urban energy and greenhouse reduction measures, policymakers and other urban energy stakeholders rely on various tools to evaluate and select various retrofit measures for either targeted or widespread adoption. As a result, significant academic and industrial efforts have been made through simulation-based and data-driven approaches to improve prediction accuracy and functionality for answering a broad suite of urban energy sustainability problems. Here, we highlight some of the cross-domain challenges that impede the widespread adoption of both simulation-based and data-driven urban modeling tools.

2.1. Urban context

Buildings rely on energy to balance thermal loads (e.g., add or remove sensible or latent heat to achieve desired thermal comfort conditions within a building) and electrical loads (e.g., lighting systems, HVAC equipment, plug loads used by occupants) to support the functioning of the activities inside and around them. While these loads are heavily dependent on occupant dynamics [15,16], they are also significantly influenced by a building's urban context: factors that may include microclimatic effects [17], neighboring buildings [6] and the wider urban form [18], vegetation [19], or other urban systems [20]. For example, it is well understood that phenomena such as the urban heat island effect, typically caused by the combination of generated anthropogenic heat and increased sensible heat storage from urban structures, can increase the outdoor temperature of urban areas when compared to nearby rural areas [7] – resulting in increased demand for energy-intensive active cooling [21]. Throughout the day, as the sun moves around a city, buildings and other tall structures can cast shadows that can affect lighting and cooling loads in neighboring buildings [22]. Similar demands for lighting and cooling solutions can be required of buildings that are hit by so-called “death rays” of concentrated light that bounce off reflective windows or building facades [23,24]. Research in fluid dynamics has explored how air flows around and within cities as typical wind patterns can shift based on where trees and buildings are located relative to one another [25]. Finally, as cities shift to district energy systems such as heating and cooling networks [26], energy hubs [27], and the use of distributed energy resources [28], individual buildings should instead be treated as a broader, interconnected energy network.

Building energy models, physics-based tools often used to evaluate building energy consumption, are limited in the extent to which their surroundings are modeled – often only including surrounding shading structures from other buildings or wildlife. These models simulate the thermodynamic energy processes of buildings by abstracting building geometry to a network of connected nodes. Each node is then used to create and solve heat balance equations based on the assumed non-geometric building parameters inputted into the model. However, because of the large number of nodes and associated equations to solve, accurately simulating the energy consumption of a single building is time and resource intensive. Doing this for hundreds, if not thousands, of buildings across a city – each with extensive required information on how they are constructed and operated – is even more intensive. To combat this challenge, extensive work in the urban building energy modeling domain has focused on how to reduce the amount of required data to simplify data processing and computation time required of the model. Efforts to reduce the required number of inputs to simulation models have included addressing the absolute required levels of detail [29] for a “minimum viable urban building energy model (UBEM)” based on its final application [30].

The most common approach to simplify inputs is through modeling groups of buildings as archetypes – either true representations or virtual abstractions of buildings with similar characteristics in an urban area [31]. These archetypes are defined by *segmenting* buildings based on properties such as construction age, use type, and size, where the buildings in each bin of buildings identical to one another. Then non-geometric characteristics (e.g., materials, operating schedules, HVAC systems) are *characterized* for each archetype. So rather than each building requiring unique non-geometric inputs, only the archetypes require them – significantly reducing the total number of unique data points required for a full urban simulation model. Because urban context is inherently complex, it cannot be simplified in a similar manner and is not extensively modeled in energy simulations [32].

The most common approach to account for climate is through Typical Meteorological Files (TMY) that represent either the city or local region's meteorological conditions. While this data is freely available for thousands of cities globally, it does not characterize extreme weather events or future effects of climate change [33]. Additionally, the data sources for many of these files are also gathered from airports or other sparse regions and therefore cannot capture local microclimatic conditions or the effects of long-wave radiation between buildings in a city [34]. Several studies have improved upon this simplified approach by modeling with historical weather data [35] and using tools that can account for mutual shading between buildings [22]. Common modeling tools used to improve upon basic weather files include the Urban Weather Generator [36], which accounts for the effects of urban geometry on urban heat island, and the Canopy Interface Model (CIM) that produces high-resolution profiles of urban microclimates [37]. Other complex tools such as ENVI-met [38] leverage complex Computational Fluid Dynamics (CFD) models to quantify the effects radiation exchange has on building energy consumption. Overall, methods to account for urban context remain ad hoc as, depending on the region and the level of detail to which urban microclimates are modeled, the effects of urban context can significantly vary their impact on building energy consumption [7].

As the deployment of open data initiatives and sensing technology expands globally, a wealth of spatial and temporal data describing the urban built environment has allowed us to uncover the energy use patterns that describe buildings. As a result, many data-driven methods have been able to achieve high degrees of accuracy while allowing us to understand the drivers that most commonly influence building energy consumption, such as building materials and use type [39]. These models to predict energy use include simpler multiple linear regressions [40] and decision trees [41] as well as more complex artificial neural networks or multi-layered deep learning models [42]. This work has since expanded to larger spatial scales, where neural networks [43], random forests [44], and support vector regression [45] are most commonly used to predict building energy use. Neural networks have also been used in conjunction with socioeconomic data to predict building and transportation energy use intensity on multiple spatial scales [46]. Neighborhoods of urban buildings that have been modeled as a network of nodes interacting with one another, coupled with neural networks, have been shown to influence multi-building energy usage [47]. Aside from using local historical weather data, data-driven models typically do not account for urban context when forecasting building energy consumption. There are machine learning-based models that have predicted future urban weather conditions [48], estimated the effects of urban context on the urban heat island [49], and relied on neural networks to evaluate the degree to which urban context influences urban building energy use [50], but these findings have yet to be translated into inputs to model and forecast building energy consumption. But by coupling neural networks and a physics-based simulation model, the authors in [20] were able to quantify the effects of cool pavements on building energy demand. Purely data-driven approaches to predicting energy consumption and estimating the influence of urban context on buildings have been largely considered separate tasks and therefore require more integration to better understand urban building energy dynamics.

2.2. Retrofit analysis

Retrofitting existing buildings can help policymakers achieve their climate targets. However, finding an effective method to identify an optimal strategy for selecting retrofits and, more specifically, which buildings to retrofit, remains a challenge across both simulation and data-driven domains.

Energy simulation models are a widely used approach in retrofit analysis because they rely on engineering algorithms grounded in building physics theory to produce interpretable, and occasionally disaggregated, predictions of how different retrofits influence energy consumption for various end uses. The Combined Energy Simulation and Retrofitting (CESAR) tool uses simulation to assess the current energy demand and emissions reduction potentials of various building stock retrofit scenarios [51]. The Urban Modeling Interface (umi) has been used to simulate the energy demand of existing neighborhoods and subsequently propose district-level interventions to improve urban energy efficiency [52,53]. Similarly, URBANopt has been developed to understand the effect of various energy technology upgrades and balance energy loads with renewable energy resources at the district level [54]. In the case of some simulation tools, integrated web-based interfaces can quickly provide energy, emissions, or cost savings estimates on a diverse library of possible energy conservation measures (ECMs). For example, CityBES relies on a web-based interface to help users identify the energy savings potential and cost effectiveness of various individual building ECMs and larger-scale ECM packages for commercial buildings [22]. The ability for these tools to graphically display the effects of retrofit scenarios make them interpretable for energy planning and decision making regardless of the user's technical knowledge of building physics.

To be considered reliable tools for retrofit decision making, building simulation models rely on calibration – the process of iteratively or statistically updating building energy model parameters to match simulation outputs to measured energy data [55]. Because this process is required when any changes are made to a model (i.e., a retrofit scenario), retraining and recalibrating an urban-scale model for each proposed design change can make simulation-based methods prohibitively time and resource intensive [56]. For example, when spreadsheet (simplified statistical approaches to calculating building EUIs) and simulation-based methods for building-level retrofit analysis were applied to a university campus in Boston, the simulation approach was shown to have taken three times as long (260 h for spreadsheet approach vs. 600 for urban simulation approach) – primarily due to the calibration time required for the simulation model [57]. Additionally, because simulation models are typically only calibrated on a single spatiotemporal scale [58,59], simulation-based approaches have limited ability to estimate the influence retrofitting one building may have on another. Computational complexity can be diminished by relying on reduced-order models – where such methods to calculate urban energy use are analogous to resistor-capacitor (RC) electrical circuit networks. City Energy Analyst, for example, relies on RC networks to quickly assess the energy, emissions, and financial implications of various urban retrofit scenarios and determine optimal schemes for distributed energy generation [60]. While they are capable of simulating large-scale models with reduced computational requirements, these simplified methods often require strong modeling assumptions or over-simplifications (e.g., modeling a building as a single thermal zone, maintaining a constant heating setpoint year-round) [26,61]. These tools demonstrate the ability for simulation-based models to help in the planning of large-scale retrofit programs. Because of the tradeoff between modeling with a higher level of detail and increased computational complexity, additional work remains to efficiently produce reliable insights on the effects of widespread urban retrofits.

Unlike physics-driven energy simulation models, data-driven applications for retrofit analysis rely on leveraging massive amounts of data, and the majority of work done in this space primarily consists of benchmarking and load shape analysis [62]. Benchmarking, typically done at a city scale, is used to quantify and understand building energy use patterns, identify the most inefficient buildings, and target the worst of them for retrofits through policymaking or other competition-based incentives [41]. Large building portfolio managers (e.g., universities, large tech campuses) also employ benchmarking techniques on whole-building energy use data to target specific buildings for energy retrofits [63], especially due to the often-limited capital available for this type of investment. Similar to some simulation-based approaches, web-based tools have also been developed to better visualize energy benchmarking performance [64]. Although benchmarking methods are able to compare the relative energy efficiency of buildings across an urban area, they still require decisionmakers to select which buildings to retrofit and do not specify the specific ECM, or combination of ECMs, that would improve their efficiency. While load shape analysis, which identifies operational efficiencies at different times during the day, month, or year [65], often provides a more temporally granular understanding of building energy efficiency, they also do not propose specific retrofits that could improve upon existing energy performance.

Data-driven methods are able to accurately model the statistical patterns of building energy use and quantify the impacts of various energy covariates but rely exclusively on the mathematical patterns in the data and not the underlying physics of the thermal and energy systems that influence it. While it is difficult to provide a data-driven model with access to all relevant observations in its training set to make an informed prediction of a building's future energy use, research in this domain has instead dealt with this challenge by relying on available data to identify important building features that could dictate future retrofit decision making. For example, recurrent neural networks have been used to evaluate smart thermostat data of residential buildings to identify their thermal characteristics and target the worst performing ones for retrofit [66]. Feature selection has also been used to identify building characteristics that are most influential in retrofit decision making [67,68]. While these models can provide retrofit recommendations for large building stocks, these approaches would still require some physics-based model to determine the influence of retrofit packages on energy consumption. Thus, without a comprehensive library of observed retrofit implications or the use of a physical modeling tool, data-driven methods for retrofit decision making are limited in their use for widespread urban energy planning.

Overall, while both physical and data-driven models have been used to assess the impacts of retrofits in urban areas, the limitations of each approach could be alleviated by leveraging the benefits of the other. Simulation-based retrofit analysis tools largely suffer from overparameterization, making the process of calibration computationally expensive and uncertain due to there being so many possible model inputs to update. Finding a computationally efficient method to quickly calibrate and produce results regarding a suite of possible retrofit design options would make them much more useful to policymakers. Data-driven tools, while able to interpret large quantities of data, lack the underlying physical understanding of how retrofits influence building energy use, influence energy use, and without a training set that encompasses all possible retrofit design combinations, they cannot be exclusively used to make decisions about urban retrofits. However, by integrating the physical context of simulation with the forecasting ability of machine learning, we should be able to improve the predictive ability of a complex spatiotemporal problem such as urban building energy use. As a result, this paper aims to demonstrate how physics and data-driven models can be integrated to explore the energy implications of various ECMs while capturing the non-linear and complex interactions of the urban context.

3. Methods

In this section, we describe the procedure that builds upon previous iterations of the DUE-S model [29], as shown in Fig. 1. DUE-S is a two-part process that integrates baseline energy simulation models (Step 1) with a deep learning model (Step 2) to capture the spatiotemporal dynamics of urban building energy use. The first step in the DUE-S framework is to build baseline energy simulation models that produce periodic time series data to capture the underlying energy use dynamics of each building in the urban study area. Then, the output data from the simulation models is fed into a deep learning model, where the objective of this model is to learn the relationship between the simulated energy consumption and the actual, metered energy use for each building to help predict future energy consumption at multiple spatiotemporal scales.

The initial iteration of DUE-S relied on a deep convolutional neural network (i.e., residual network) with the intent of capturing the spatial relationships between buildings in an urban area. However, this paper looked better capture the temporal patterns of electricity use through a long, short-term memory network (LSTM) to improve prediction accuracy. Furthermore, while the initial project focused on the accurate, multi-scale prediction of urban building energy consumption, we extended this work by introducing the capability of conducting a retrofit analysis that takes advantage of the learned relationships between the simulated and metered energy consumption of each building in the urban study area. Specifically, after training the LSTM to understand the relationship between simulated and metered energy consumption, the model can be repurposed for urban retrofit assessment by holding the learned parameters constant and feeding in new data from modified simulation models. These modified energy simulation models are changed based on proposed retrofit policies and can subsequently capture the physical dynamics of how retrofits influence building energy use. With this new input dataset from the modified simulation models, the same LSTM can then predict what future energy consumption would look like under that retrofit scenario at varying spatiotemporal scales.

3.1. Step 1: capturing energy use dynamics through energy simulation models

Physics-based energy modeling is often overparameterized [69], and, when expanded to the urban scale, often requires an extensive amount of time and resources to produce accurate results. Sourcing and cleaning the required data needed for an urban energy model is often the most difficult aspect of the modeling process as the available data needed to create inputs for each individual building will vary between them. While there is a tradeoff between increasing levels of detail, the time and resources needed to collect it, and improved model accuracy [29], the objective of the DUE-S simulation models is to establish the underlying first-order physical dynamics of building energy use. Rather than relying on the simulation model to predict building energy use, the simulation outputs are instead used as inputs to a deep learning model that is used to predict energy consumption. Therefore, we use simplified geometries and inputs based on virtual archetypes to quickly develop our building energy simulations.

Energy simulation models require three primary types of inputs: weather, geometric characteristics, and non-geometric characteristics. Energy simulation models often rely on Typical Meteorological Year (TMY) datasets, which contain the typical weather characteristics (e.g., dry bulb temperature, relative humidity, solar insolation, wind speed) at the hourly time scale for a region. However, these measurements are more representative of rural landscapes rather than dense city centers [33]. Due to the limitation of TMY files neglecting weather differences caused by the urban context, we use historical weather data taken at hourly temporal granularity, as this type of data is available across more diverse geographies and better represents of observed urban weather patterns.

To construct the geometric representations of buildings in an energy simulation model, we draw from GIS shapefiles available from local municipal websites. These shapefiles contain geometric information on each building's footprint that are tied to additional characteristics (e.g., building height, age of construction, use type, elevation) through an identifying parcel number. The building footprints can be merged with accompanying geometric information in a modeling tool (e.g., Rhino, SketchUp) by "extruding" them with their building heights to create "massing models" – simplified geometric representations of a building – often referred to as "2.5-D" models. If additional information is available, these models can be further divided into floors either from the shapefile or using an assumed floor-to-floor height with the goal of understanding their shape and orientation relative to one another.

Finally, to describe the non-geometric characteristics (e.g., HVAC systems, operating schedules, materials and constructions), we draw generalized inputs from the U.S. Department of Energy's (DOE) Commercial Reference Buildings database [70]. Based on the 2003 Commercial Building Energy Consumption Survey (CBECS), the Commercial Reference Buildings define non-geometric inputs of virtual archetypical models for 16 commercial building types under 3 ages of construction in the United States. Overall, these models represent about 70% of the U.S. national building stock. While these models only apply to buildings located in the United States, other countries and international entities have developed their own methods to standardize non-geometric inputs [71]. While this is a highly generalized approach to define commercial building characteristics, specific inputs are often not available through open data initiatives. At a minimum, having information on each building's use type and age of construction can help identify which archetypical model from a national database like the Commercial Reference Buildings is necessary to develop these generalized building energy models. However, in the event that more detailed information on specific building characteristics is available, it can be substituted in place of the Commercial Reference Building inputs. Any input parameters not

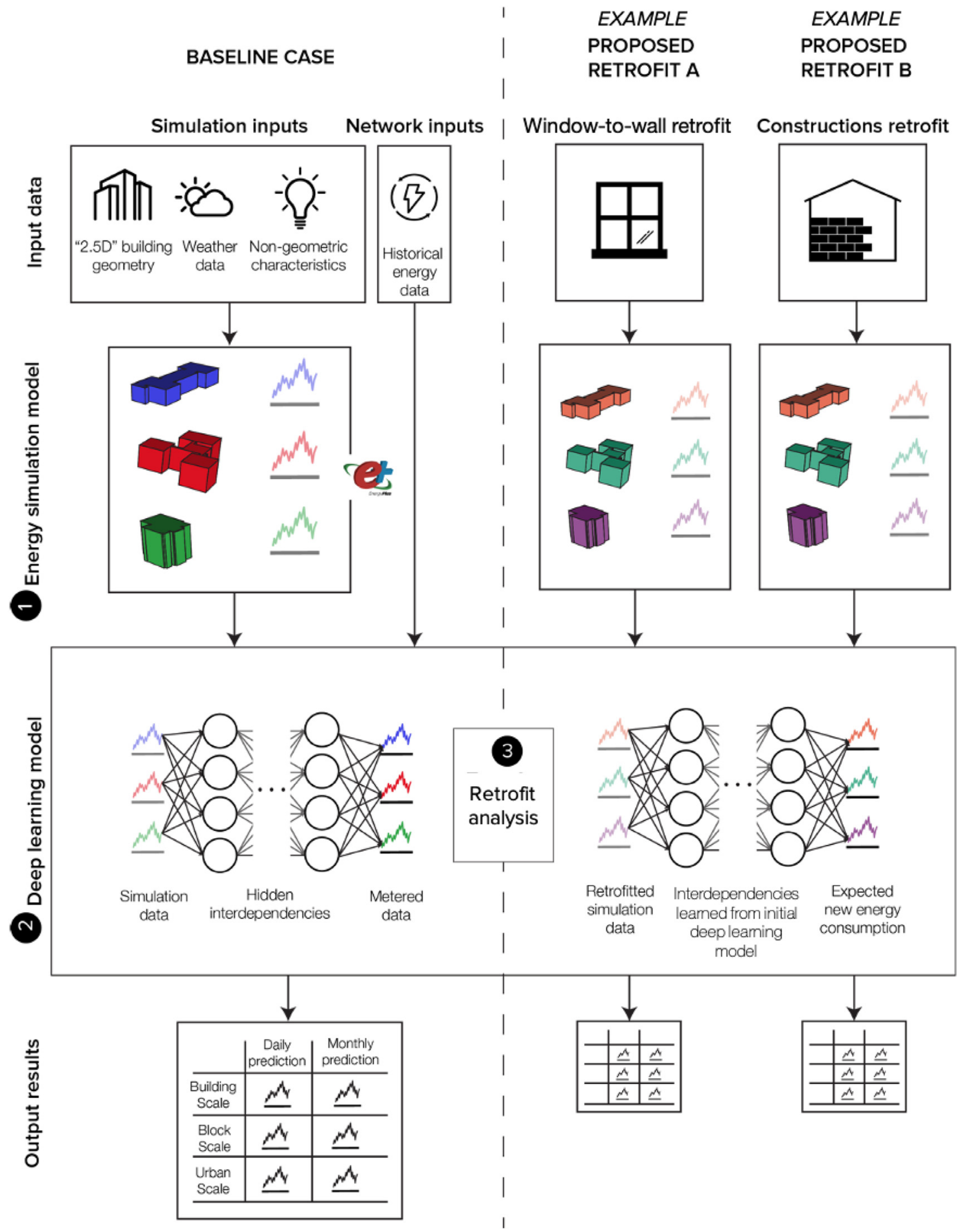

Fig. 1. DUE-S modeling framework. Step 1 consists of building a baseline urban building energy model to produce hourly predictions of first-order energy consumption of each building in the study area. This data, along with actual metered energy consumption data, is used as an input to Step 2 – a deep learning model that aims to map the relationship between simulated and metered energy use data and produce the final predictions of urban building energy use on multiple spatiotemporal scales. Finally, Step 3 shows the process of how DUE-S combines the underlying physics of simulation and prediction accuracy of deep learning to generate predictions of how large-scale retrofits would impact building energy use across an urban area.

specified by these sources are given default values from the engineering simulation tool (e.g., EnergyPlus, DOE-2, IES-VE) used in making the first part of the DUE-S model.

After preparing all primary inputs for weather, building geometry, and non-geometric characteristics, these models are prepared using an energy simulation engine (e.g., EnergyPlus). The 2.5-D massing models are first constructed and further divided into floors. Each building in the model is then assigned an archetype (one type of the Commercial Reference Buildings) that matches its age of construction and use type. For example, a recently constructed, large office building would be modeled using the non-geometric characteristics of a "New Construction, Large Office" Commercial Reference Building archetype. Finally, each building is simulated to produce an output time series dataset of hourly, whole-building energy consumption. We emphasize that the objective of this step in the DUE-S modeling process is to capture the underlying energy use dynamics of these buildings. At this point, a typical urban-scale

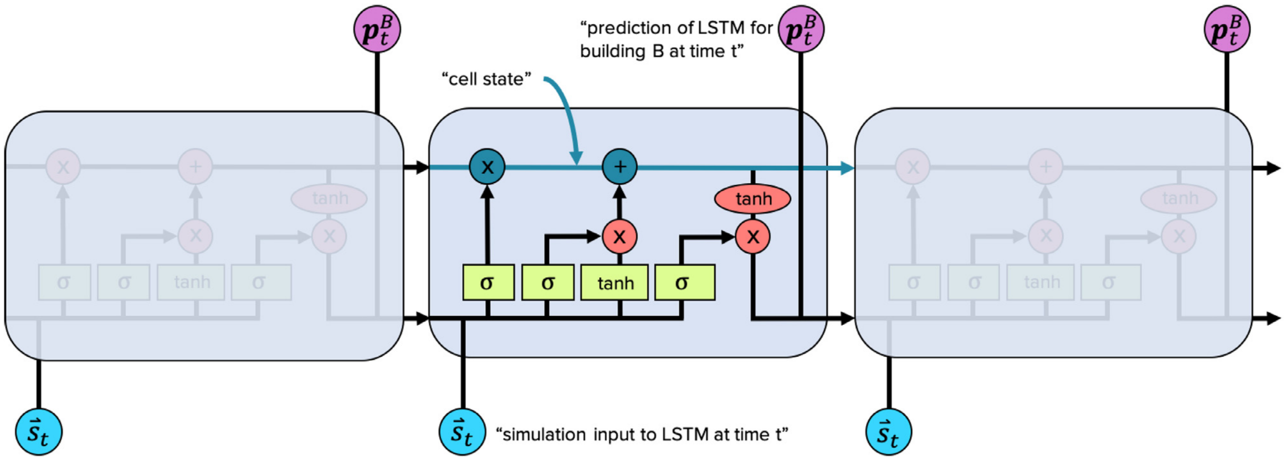

Fig. 2. A diagram of a chain of three long short-term memory (LSTM) cells that make up a larger LSTM neural network. The outputs from an energy simulation model at time t are fed into a cell to predict the metered energy consumption of building B_k at time t. The primary benefit of using an LSTM for a time series prediction task is the cell state, represented by the horizontal line at the top of each cell. The cell state can carry long-term information about previous predictions and determine if it is useful to predict the following ones.

energy model would undergo a detailed, time-intensive calibration step to improve its energy use prediction accuracy [55]. However, instead of relying on calibration, we use the outputs of our simulation model as an input for the second stage of the DUE-S modeling workflow: a long, short-term memory network (LSTM) – whose objective is to produce the final predictions of energy consumption for each building in the urban study area.

3.2. Step 2: predicting metered energy consumption through a deep learning model

Much of the uncertainty and error in energy simulation predictions arise from the assumptions made in the modeling process as well as the hidden urban context effects that are often not captured by individual building energy simulations [34]. Even when calibration is done to improve upon the initial prediction results, this process is often very time-intensive and still requires additional assumptions on the model's most uncertain or error-prone simulation inputs. To mitigate these issues in simulation-based methods, we use a deep learning algorithm to map simulated energy consumption to actual metered energy consumption. As the deep learning algorithm trains, it learns the differences between each building's simulated and actual energy consumption (i.e., uncertainties associated with input parameters and urban context). When the deep learning algorithm is used to make future energy use predictions, it can rely on these previously learned relationships to improve its accuracy. Finally, because inter-building effects and urban context are dynamic, these phenomena are often non-linear and thus can benefit from a neural network and its characteristic hidden layers to better capture this non-linearity. To implement this method, the model, at minimum, requires metered energy data for each building in the study area. While less granular data (e.g., monthly) could be used to implement a deep learning model, increasing the amount of data for each building (e.g., daily, hourly, interval) is likely to improve the final prediction accuracy of the final model's results.

When selecting a deep learning architecture for this task, we wanted to use a model suitable for a many-to-many prediction task – essentially where the past energy consumption of *many* buildings is used to predict the future energy consumption of *many* buildings. While a previous iteration of the DUE-S modeling framework relied on convolutional neural networks (CNN) in the form of ResNets to account for the spatial nature of the urban built environment [29], we aim to improve upon previous prediction results by instead focusing on the time series aspect of energy consumption data. Recurrent neural networks (RNNs) are a class of neural networks that use previous outputs as inputs for future predictions, and because they are often used in the application of predicting time series data [72,73], we utilize this architecture for the extension of DUE-S.

One of the challenges in using simple RNNs is that they often rely on the most recent inputs to predict the next output. When dealing with building energy data, where weekly, monthly, or even annual patterns of consumption are important in forecasting, the inability for simple RNNs to consider these long-term dependencies can result in higher error rates. To circumvent this issue, we use a variant of the RNN called the long short-term memory (LSTM) network [74], diagrammed in Fig. 2. As shown in Fig. 2, the cell state, represented by the horizontal line in the top of the illustration, runs through the entire chain of LSTM cells. The cell state carries previously learned information from one cell to the next, and through a series of gates (represented by the sigmoid operator and multiplication signs), the LSTM cell has the ability to add or remove information to the cell state in order to make predictions. While we also evaluated both simpler (k-nearest neighbors, support vector regression) and more complex models using convolutional neural networks (CNN-LSTM) and LSTM autoencoders [75], we in practice found that a simple LSTM has better prediction accuracy than the simpler models yet achieved comparable performance to more complex deep learning architectures with less memory. For our prediction task, we use 2 sequential LSTM 64-unit layers and 2 time-distributed fully connected layers configured for a many-to-many prediction task that can output energy consumption predictions on multiple spatial and temporal scales.

To evaluate the impacts of urban context on energy consumption, we compare two scenarios: *With Context* and *No Context*. To model urban context using the *With Context* scenario, all simulation outputs $\vec{s}_t$ are inputted to the deep learning model regardless of the target building B_k (Fig. 3). The model is then free to learn the time series characteristics of different retrofits as well as complex interdependencies between physics-based simulations of multiple buildings. Conversely, the *No Context* scenario instead uses the LSTM where no other buildings in the study area are used to predict a target building's electricity usage. Instead, only simulation output $\vec{s}_t^B$, representing the simulation output of only target building B_k is fed into the LSTM to produce a prediction of energy consumption. After the *With Context* and *No Context* models are trained, their outputs, w_t^B and n_t^B, corresponding to LSTM prediction outputs of the *With Context* and *No Context* scenarios, respectively, are compared to characterize the effects of urban context.

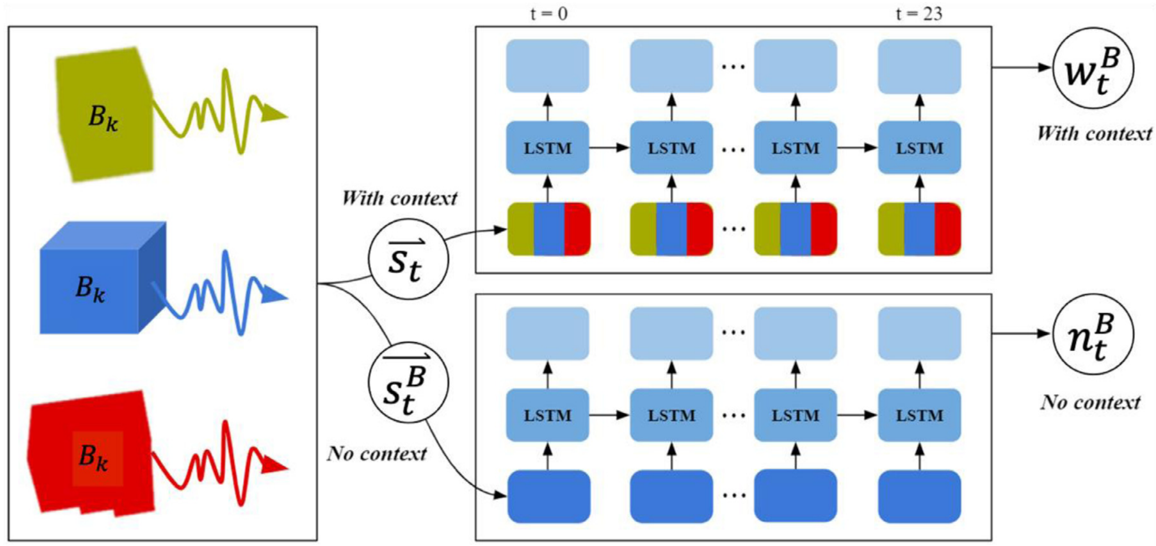

Fig. 3. Training the DUE-S model for *With Context* and *No Context* scenarios. *With Context* scenarios predict building energy consumption using all buildings' energy simulation outputs as inputs to the LSTM (i.e., all buildings are used to predict a single target building B_k energy consumption). *No Context* scenarios predict building energy consumption only using the target building's energy simulation output (i.e., only a single target building B_k is used to predict its own energy consumption).

3.3. Step 3: assessing retrofit scenarios through an integrated urban energy model

We expand on the original DUE-S framework by introducing the capability for the model to conduct large-scale retrofit analyses. One of the primary challenges in using a purely simulation-based approach to retrofit analysis is that because these models are calibrated to their original scenario, if any changes or retrofits are made to the building, its systems, or its surroundings, the model will need to be re-trained with additional ground truth data. Urban building energy models are often also calibrated to a single spatial and/or temporal scale and may be unable to model energy use at other scales. And while purely data-driven approaches model the statistical patterns of energy consumption, they are unable to consider the underlying building physics of thermal and energy systems. Because urban buildings may have hundreds of possible retrofit options to consider, it is impossible to have a robust training set that can consider the influence of all potential urban retrofit options.

To run the retrofit analysis (diagrammed in Fig. 1), the initial baseline energy simulations are modified with the proposed retrofits and re-simulated. The new outputs of these retrofitted simulations become the new inputs to the previously trained deep learning model discussed as part of Step 2 of the DUE-S modeling workflow (Section 3.2). Because the deep learning model has already learned the relationship between simulated and actual energy consumption, we hold all of its learned parameters constant and pass through the new simulation outputs to predict the new metered energy consumption on multiple spatiotemporal scales. As a result of this model architecture, DUE-S does not require additional re-training or re-calibration to be operationalized and therefore significantly reduces the amount of time needed to evaluate a single urban retrofit scenario.

Finally, in a true retrofit scenario, it is not often the case that all buildings will be retrofitted. Thus, we developed a greedy optimization algorithm in order to determine how many and which specific buildings could be targeted to maximize cumulative energy savings across buildings in an urban area. As shown in Fig. 4, for each building candidate, our model predicts the marginal change in energy savings achieved from its addition to the retrofitted subset. The candidate with the greatest marginal savings is permanently added to the subset (Retrofit Pool). The algorithm continues until all there are no building candidates left that result in marginal energy savings.

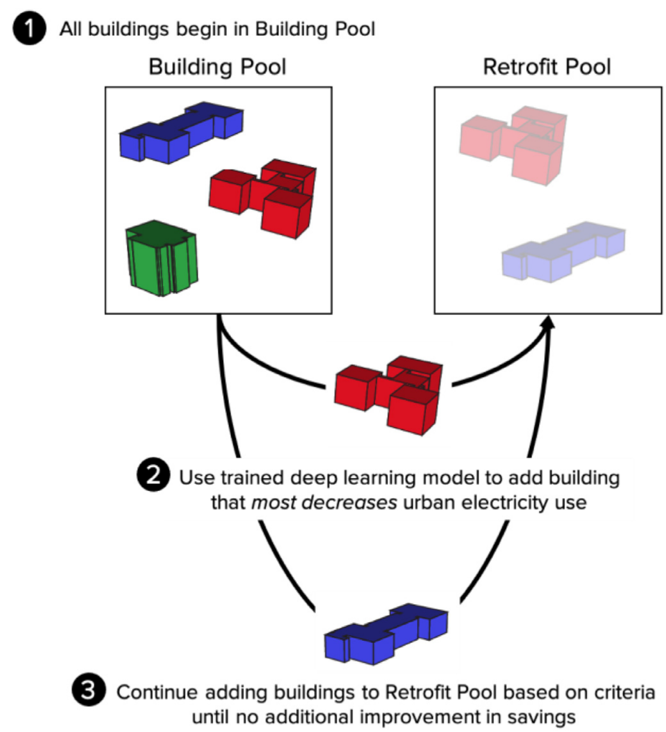

Fig. 4. Schematic representation of greedy optimization algorithm for selecting urban building retrofits. To determine the optimal buildings to retrofit in a given urban area, each building is first assessed to determine which of them decreases urban energy use the most; the building with the greatest impact is then added to the Retrofit Pool. This process is repeated with all remaining buildings in the Building Pool until there are no additional improvements in urban energy savings.

4. Case study

4.1. Study area and data inputs

We evaluated the performance of the expanded DUE-S framework on a dense cluster of 29 commercial and mixed-use buildings in down-

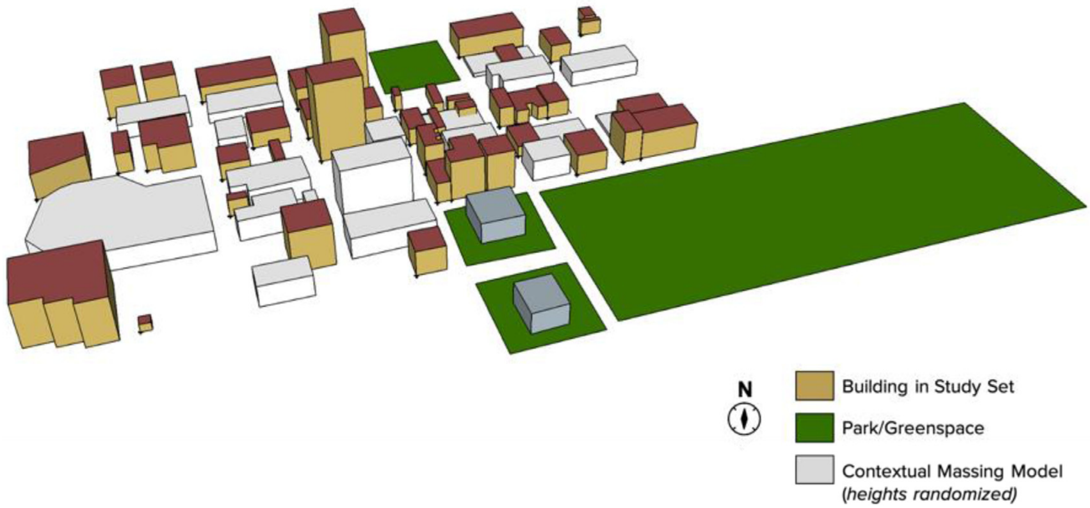

Fig. 5. Diagram of urban study area in Sacramento, California, USA.

Table 1
Data acquired for baseline energy simulation models and deep learning model.

Use	Data Source	Data Field
Weather data	Solcast [76]	Hourly historical weather info (inputs typical of an Energy Plus Weather (EPW) file
Building geometry	Sacramento GIS and tax assessor database [77]	Building area Building height Number of floors
Non-geometric inputs	US Department of Energy Commercial Reference Buildings [70] Sacramento tax assessor database [77]	Building use type Occupancy schedules Building constructions and materials HVAC type Heating and cooling loads
Metered electricity data	Sacramento Municipal Utility District (SMUD)/Local utility	Whole-building hourly electricity consumption

town Sacramento, California, USA (Fig. 5). Because one of the underlying goals of this work is to understand the inter-building dynamics that influence energy consumption, this site was selected based on the availability of whole-building energy use data describing a closely co-located group of buildings. The study area is located in the central business district with other surrounding buildings and a large greenspace on the southern border. While DUE-S has the capacity to work for multiple energy sources, this case study specifically analyzes electricity consumption as it is most relevant to the warm and temperate climate of Sacramento, California (natural gas is primarily used for heating, which is not often necessary in Sacramento). Table 1 shows the data used for both the baseline energy simulations and the deep learning model where, with the exception of the electricity consumption, all other data is publicly available. To accompany the 3 years of electricity data, 3 years of hourly historical weather data was collected from Solcast [76]. The building geometries (i.e., building areas extruded by its building heights) were constructed using shapefiles from Sacramento's GIS and tax assessor databases. Using that same database, we used its information on each building's primary use type and age of construction to match each building to a corresponding Commercial Reference Building. That corresponding Commercial Reference Building then served as the template of inputs that would be fed into each baseline energy model in EnergyPlus.

After simulating each building for 3 years with the corresponding historical weather files, the simulation output – hourly whole-building electricity consumption – was then fed as an input to the long short-term memory network (LSTM). Ground truth data for the LSTM consisted of hourly electricity consumption data for 2016–2018 for each of the 29 buildings in the urban study area. To avoid data leakage that may arise when randomly splitting sequential time series data [78], we split the first two years into consecutive sequences of training (18 months) and validation (6 months) sets. The last year (2018) was reserved for the testing set. We trained and validated our model by using historical simulations as inputs to predict historical observed energy usage. Specifically, we inputted the 24 most recent simulation timesteps $S = [s_{t-23} \cdots s_t]$ (i.e., one full day) to generate predictions P. As we describe in Section 3.2, the LSTM is modified to evaluate both *With Context* and *No Context* scenarios: *With Context* scenarios rely on all simulation outputs to predict each building's electricity consumption whereas the *No Context* scenarios only rely on the target's building electricity use to predict its future consumption.

Although we used only one model to predict energy consumption for all 29 buildings, we optimized our model from a single loss curve to avoid difficulties in model training diagnosis stemming from multiple optimization objectives and loss curves. We accomplished this by one-hot encoding the target building for prediction as a model input. Essentially, we created a binary categorical variable for each building, where only the target building is coded as "1″ to tell the model this is the only building for which it should predict electricity use. Therefore, for each target building b_k, the model output at each timestep was a scalar prediction p_t^k of electricity use for target building b_k instead of the length N output vector $\vec{p_t}$. Model optimization was performed using an Adam optimizer with a learning rate of 0.001 – chosen for its adaptive learning rate and invariance to scaling of the objective function.

Finally, to conduct the retrofit analysis, we re-simulated the initial baseline energy models under three commonly utilized [79] retrofits described in Table 2. In addition to selecting common retrofit types, we also wanted to explore the differences between retrofits that were in-

Table 2
Proposed retrofit scenarios for downtown Sacramento study area.

Retrofit Type	Baseline Scenario	Retrofit Scenario	
		All buildings retrofitted	One block of buildings retrofitted
“Lighting”	DOE Commercial Reference Buildings (dependant on specific building)	Switch to LED bulbs (~27% decrease in lighting power density)	
“Windows”		Updated to ASHRAE 90.1–2010 materials (dependant on building type, but generally lower U-value and solar heat gain coefficient)	
“Full"		Include both “Lighting” and “Window” retrofits	

ternally (i.e., indoor lighting systems) or externally (building fenestration) focused. The new simulation outputs were then fed into the trained LSTM to produce new predictions for how each type of retrofit would impact the electricity consumption of each building in the urban study area. These retrofit scenarios are also evaluated under both *With Context* and *No Context* scenarios described in Section 3.2.

4.2. Validation metrics

Validation is a critical step in the energy modeling process that evaluates the accuracy to which a model can predict building or urban-scale energy use. The most common metrics used in simulation-based approaches are mean bias error (MBE), which reflects the level of overestimation versus underestimation of energy consumption and the model's long-term performance, and the coefficient of variation (CV) of root mean square error (RMSE) – a measure of the variability of accuracy overcover a period of time. While many governing bodies have set limits on acceptable error rates for MBE and CV(RMSE) [80–82], these benchmarks were established for use on *individual* building energy models. Currently there is no universal standard or measurement system to evaluate the performance of urban-scale energy models, and while many previous works have adopted the scale set for individual building performance [29,58], urban-scale models should be placed under stricter requirements because of the higher granularity to which these models are estimating building energy use.

Data-driven models used in forecasting applications are often evaluated using other metrics, such as mean average percentage error (MAPE), mean squared error (MSE), and mean average error (MAE). For each building, we sought out an absolute error value roughly proportional to its average energy use, so we strategically used MAE loss in our training loop for its intermediary between MAPE and MSE. MAPE optimizes for low-energy buildings with proportionally larger noise, while MSE penalizes the relatively larger errors in high-energy buildings disproportionately. Thus, we validated the hybrid data-driven DUE-S model using MAE:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|$$

Where n is the number of timesteps (either hourly, daily, or monthly depending on the model's output), Y_i is the actual measurement of electricity use for the selected timestep and $\hat{Y}_i$ is the predicted value from the LSTM. These estimations were made for both each individual building in the study area as well as the full urban-scale result.

5. Results and discussion

After creating the initial DUE-S model describing the existing buildings in the case study area, we calculated the mean average error (MAE) per the equation described in Section 4.2. The baseline prediction error results were calculated for both the individual and urban spatial scales as well as on the hourly, daily, and monthly timesteps and are shown in Fig. 6. Overall, DUE-S improves its prediction accuracy as the spatiotemporal scales increase (i.e., monthly accuracy is higher than daily; urban-scale is more accurate than individual-scale). This pattern

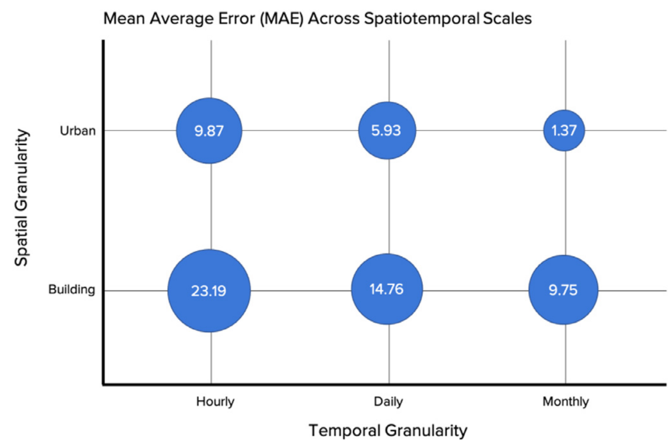

Fig. 6. DUE-S electricity prediction error on multiple spatiotemporal scales, measured in mean average error (MAE). Prediction accuracy improves as the spatial or temporal granularity of estimation decreases.

is consistent with both previous results in the DUE-S model [29] as well as other urban-scale energy modeling [52] and data-driven forecasting [83] studies. Models tend to perform better on larger granular scales because aggregation of errors will often reduce the variability in the overall prediction of electricity consumption. In other words, when looking at urban-scale accuracy, the errors of buildings that perform the worst are partially compensated by the better performing buildings. And similarly, when looking at daily or monthly performance, the hours that are the least accurate are partially compensated by the ones that are more accurate.

After establishing the baseline DUE-S model that describes the energy performance of existing buildings, we ran the retrofit analysis on the 12 scenarios described in Fig. 7. We simulated the effects of both when all buildings underwent the same retrofit and when only a select block of buildings underwent the retrofits. We also compared the differences between how the hybrid DUE-S model predicts building energy consumption when accounting for urban context (*With Context* scenario) and when it does not (*No Context* scenario).

The results of our case study indicate that the greatest reduction in whole-building electricity use amongst retrofitted buildings occurs when they undergo window retrofits (Fig. 8). Overall, as expected, the majority of the buildings in our study area decrease electricity use after undergoing retrofits. For those buildings that did slightly increase use, we note that these buildings had lower energy usage (≤ 200 kWh hourly electricity use) and likely due to errors introduced by poor parametrization of the simulations (i.e., buildings with increases tend to have similar simulation results to their baseline ones). The “window” and “full” retrofit scenarios, both of which receive window retrofits, see greater changes in electricity consumption than the “lighting” retrofit. We further investigate these trends by analysing the percentage change in electricity use resulting from the “full” retrofit scenario on a monthly (Fig. 9A, 9C) and hour-of-the-day (Fig. 9B, 9D)

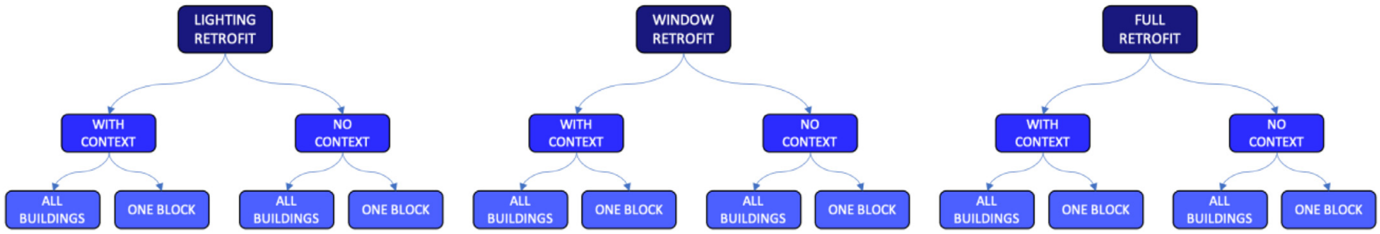

Fig. 7. Retrofit analysis scenarios performed as part of Sacramento case study. Scenarios involve those evaluating urban context, the total number of buildings retrofitted, and the type of retrofit performed on each building.

Fig. 8. Median hourly percentage reduction of buildings under various retrofit options in a *With Context* scenario. Lower energy buildings show the greatest reduction in electricity use, especially under "full" and "window" retrofit scenarios.

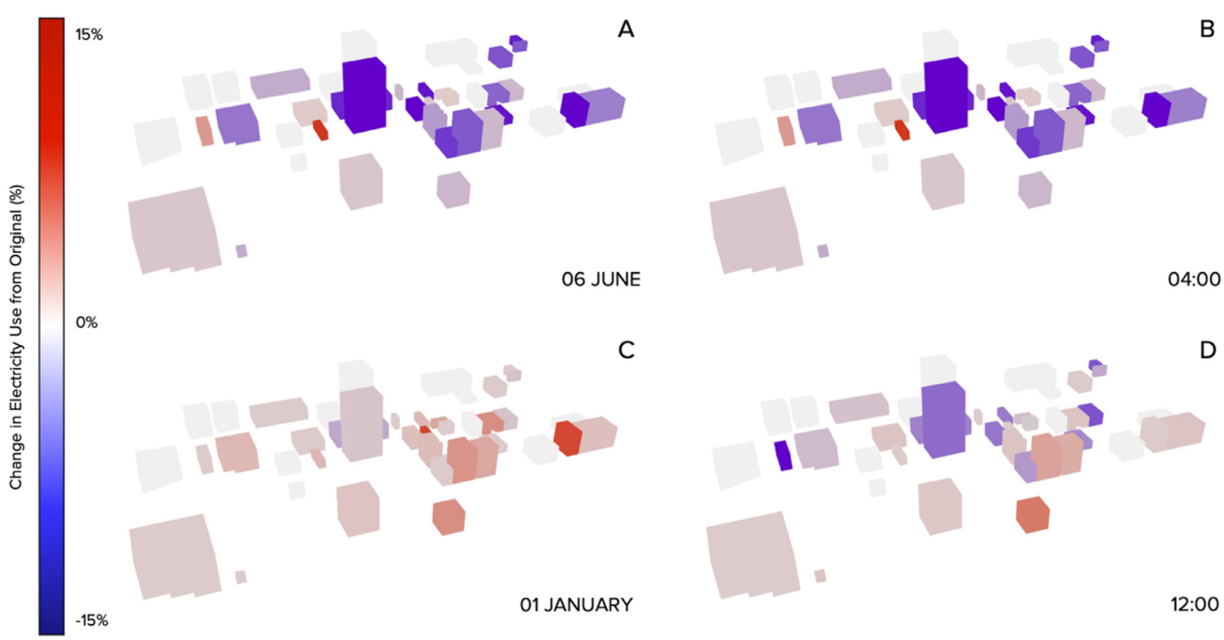

Fig. 9. Reduction in building-level electricity consumption from a "full" retrofit. These specific examples demonstrate percentage reduction for June (A), 05:00 (B), January (C), and 12:00 (D). Overall, buildings that reduce their electricity consumption after retrofitting see most improvement in summer months and shoulder hours of the day, which is when most stress is likely placed on the grid. Full monthly and hourly results can be found in Appendix Figs. A.1 and A.2.

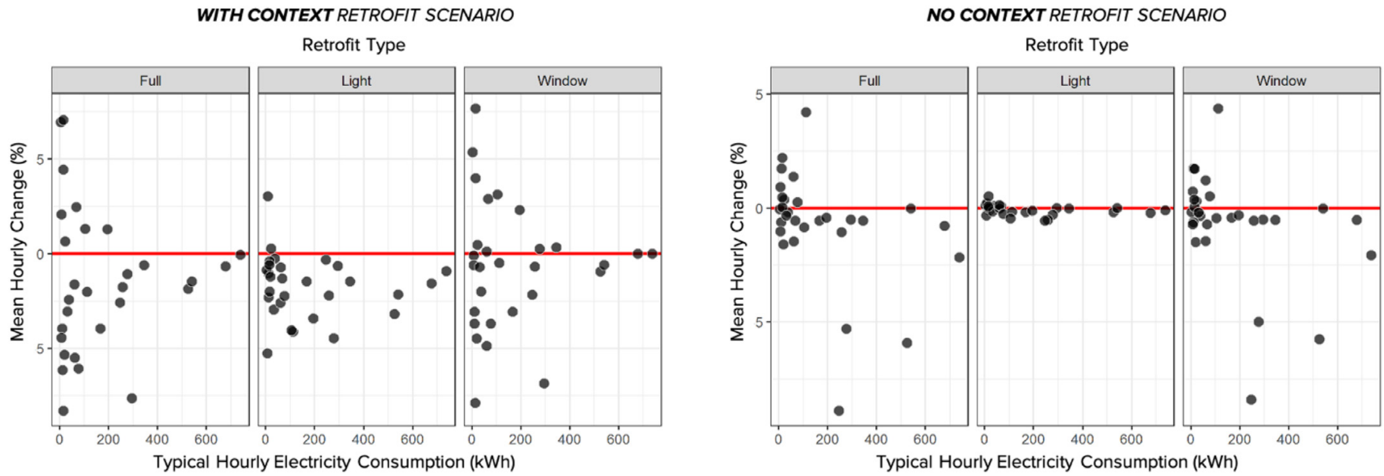

Fig. 10. Comparing mean hourly change in electricity use under a *With Context* and *No Context* scenario.

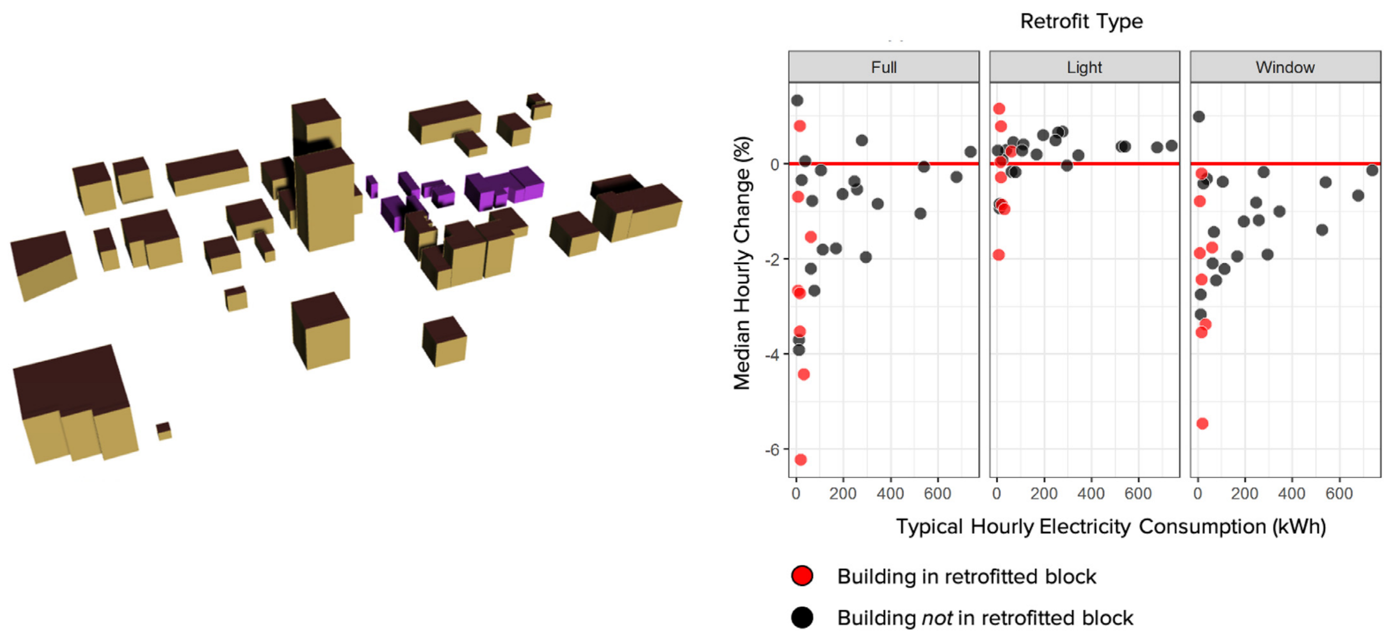

Fig. 11. Median hourly change in electricity use under a single block retrofit scenario (buildings in purple indicate they were retrofitted). Window retrofits show a greater impact on reduced electricity consumption in surrounding buildings.

scale. In general, buildings that reduce their electricity consumption post-retrofit see the largest improvements in the summer months (June – August) and in the early morning hours of the day (04:00 – 07:00) when high stress is placed on a heavily renewables-dependent grid to "start up" commercial buildings or cool them throughout the day [84]. Lowering electricity demand through such retrofits can reduce strain on the grid, creating a more flexible and resilient grid that benefits utility companies while also lowering energy costs for consumers. The results for all months and hours of the day can be viewed in Appendix Figs. A.1 and A.2, respectively.

As discussed in Section 3.2, the output of the DUE-S deep learning model can be modified to understand the differences in energy performance when and when not considering urban context. Comparing the shifts in hourly electricity use under a *With Context* and *No Context* model (Fig. 10), we find buildings generally experience greater decreases in electricity consumption when the model accounts for urban context. In a scenario in which only a single block of buildings undergoes retrofits (red dots in Fig. 11), we see that there is some level of influence of retrofits on decreased electricity use in surrounding buildings (black dots in Fig. 11). Here, we also note that "lighting" – an exclusively indoor retrofit – has a negligible effect on surrounding buildings, but the "window" and "full" retrofits that affect the building envelope do play a role in decreasing the electricity use of surrounding buildings. Under this block retrofit scenario, when plotting every building's median hourly change in electricity use by the proximity to the nearest retrofitted building (Fig. 12), we see a slight tendency for buildings closer to retrofitted buildings to have greater decreases in electricity consumption. Based on this analysis, if urban context were to play a significant role in determining changes in electricity use, we would expect to see the bubbles towards the bottom left of the plot. Overall, the results from our retrofit analysis imply that the buildings with "externally facing" window retrofits, which primarily decrease each window's U-factor and solar heat gain coefficient, have a greater effect in decreasing the energy consumption of their surroundings.

Finally, we want to show how the DUE-S model can be operationalized for urban energy stakeholders to inform their decision making about selecting buildings to retrofit. Specifically, we want to compare how DUE-S-based and simulation-based approaches determine the num-

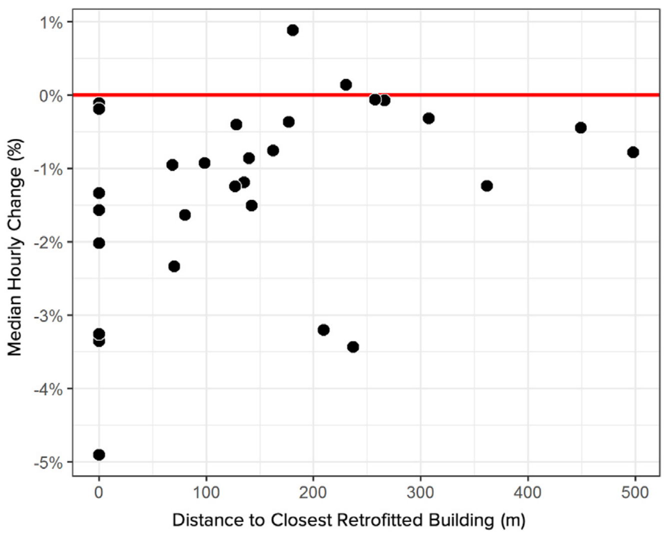

Fig. 12. Median hourly change in electricity use versus proximity to retrofitted buildings under a single block retrofit. Here, there is a slight tendency for smaller, nearby buildings to decrease electricity use.

ber of buildings needed to achieve the majority (~80%) of possible urban electricity savings. Under the circumstances in which all buildings receive "full" retrofits, the maximum projected electricity savings across the entire urban study area is 14.4% in a simulation-based approach and 13.8% in the DUE-S approach. While comparable in magnitude, these savings are achieved through different means. The DUE-S approach relies on its greedy optimization algorithm, described in Section 3.3, where buildings are selected for retrofit one at a time based on the marginal savings it will contribute. A simulation-only method, however, would select for buildings in order of the highest building-level electricity savings as it cannot account for the compounding effects of the inter-building influence that result from urban context.

The DUE-S approach is able to consider the "snowball" effect of compounding retrofits, we show that this case study area can achieve ~80% of maximum projected electricity savings with fewer building retrofits (6 retrofitted buildings) than as predicted by the simulation-based approach (11 retrofitted buildings) (Fig. 13). Furthermore, we compare the number of required buildings needed to achieve 80% electricity savings under the DUE-S *With Context* and *No Context* models (Fig. 14) and find that when accounting for urban context, the model prefers retrofitting more buildings, but less overall square footage (120,701 square feet in DUE-S *With Context* scenario vs. 427,628 square feet in DUE-S *No Context* scenario) to be retrofitted to achieve the same amount of savings. Because DUE-S is able to target a smaller area of buildings to maximize energy savings, this can be especially helpful for policymakers or building portfolio owners that have limited capital available to finance energy conservation measures. Finally, our optimization algorithm also reveals certain buildings that are repeatedly selected as most suitable for retrofitting (Fig. 15). Three buildings in particular (Buildings 35, 37, and 40) are found most often in each run of the optimization algorithm, and, like most of the selected buildings, are office buildings classified as having a construction date before 2004. While a greedy optimization algorithm may not necessarily find the global optimum, our results show that this approach can significantly reduce the amount of square footage and associated costs of retrofitting a large urban area, especially when accounting for urban context. In our specific example, based on the results of the optimization, we would expect a retrofit program targeting older office buildings to yield the greatest reduction in the city's overall energy consumption. This finding should not replace the decision-making process for urban sustainability policymakers but instead serve as a helpful starting place to determine which urban buildings are best to retrofit.

Overall, the results of this case study highlight the merit of a hybrid approach to urban energy retrofit analysis, but it also demonstrates why consideration for inter-building effects and urban context is so critical in developing urban building energy models. Through our analysis, we learn that accounting for urban context can compound the impact of retrofits on individual buildings as they also influence the electricity use of their surroundings. Finally, the results of our optimization approach show how policymakers can greatly benefit from using hybrid physical models to inform retrofit strategies that maximize electricity savings while reducing the costs and logistical challenges associated with retrofitting a large number of buildings. This work represents a key step in progressing modeling efforts at the intersection of building simulation and machine learning and furthers our understanding of the influence of urban context on building energy use.

6. Limitations and future work

Our model shows how accounting for urban context can cause greater whole-building energy savings effects across an urban area when compared to examining individual buildings in isolation. Because our model is focused on predicting building energy use at larger spatial scales, we are currently unable to disaggregate our results by end uses (e.g., chillers, fans, lighting). While this is likely possible using the outputs from EnergyPlus, this capability is currently out of scope for this work. Furthermore, additional work needs to be done to fully validate the findings of this study. While this study required limited detail on

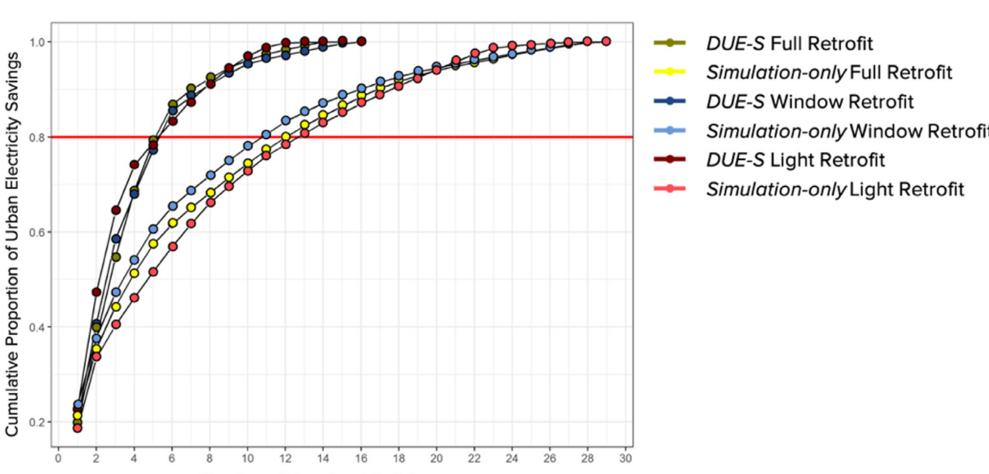

Fig. 13. Comparing the number of buildings needed to achieve the majority of urban energy savings. DUE-S requires fewer buildings to achieve ~80% of urban energy savings compared to a simulation-only approach.

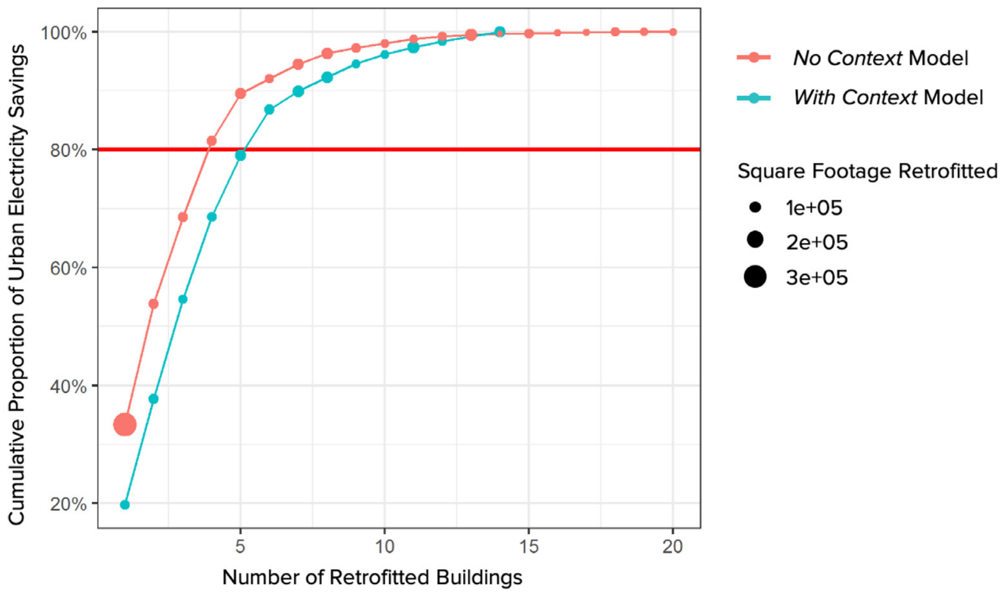

Fig. 14. Comparing the *With Context* and *No Context* results for retrofit optimization. While the *With* Context model requires more buildings to be retrofitted to achieve ~80% of urban electricity savings, it requires a lower amount of overall square footage to undergo those retrofits.

	Full Retrofit		Window Retrofit		Light Retrofit	
With Context	Building 10	Pre-1980, Office	Building 17	1980-2004, Office	Building 30	Pre-1980, Office
	Building 40	1980-2004, Office	Building 10	Pre-1980, Office	Building 10	Pre-1980, Office
	Building 17	1980-2004, Office	**Building 40**	1980-2004, Office	**Building 40**	1980-2004, Office
	Building 30	Pre-1980, Office	Building 30	Pre-1980, Office	**Building 35**	Pre-1980, Office
	Building 35	Pre-1980, Office	**Building 35**	Pre-1980, Office	Building 28	1980-2004, Retail
	Building 37	1980-2004, Office	**Building 37**	1980-2004, Office	Building 17	1980-2004, Office
No Context	Building 15	1980-2004, Retail	Building 15	1980-2004, Retail	**Building 40**	1980-2004, Office
	Building 40	1980-2004, Office	**Building 40**	1980-2004, Office	**Building 35**	Pre-1980, Office
	Building 25	1980-2004, Office	Building 25	1980-2004, Office	Building 25	1980-2004, Office
	Building 37	1980-2004, Office	**Building 37**	1980-2004, Office	Building 33	Pre-1980, Office
	Building 35	Pre-1980, Office	**Building 35**	Pre-1980, Office	Building 15	1980-2004, Retail
	Building 26	Pre-1980, Office	Building 26	Pre-1980, Office		
	Building 33	Pre-1980, Office	Building 33	Pre-1980, Office		

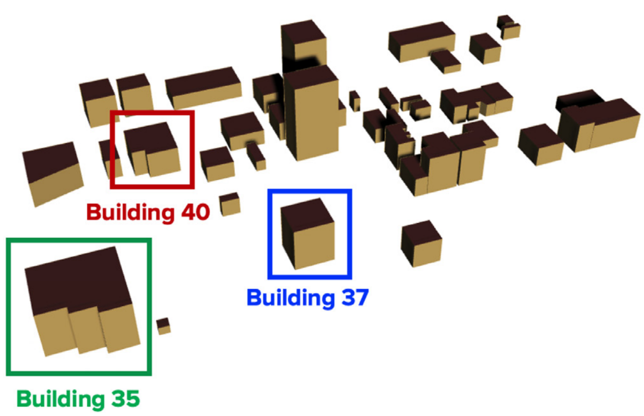

Fig. 15. Buildings selected through DUE-S optimization model to retrofit, listed in order of selection by greedy optimization algorithm. Building 40, Building 35, and Building 37 emerge most frequently as buildings that should be selected for retrofitting.

each building's specific geometric and non-geometric input, a future study aims to look at how increasing their levels of detail may improve the prediction accuracy of DUE-S. This additional analysis can help us better understand how DUE-S performs on larger urban study areas with different electricity and natural gas usage patterns (i.e., how does DUE-S perform on a city with greater natural gas-based heating demand) and explore the impacts of urban context on other building types (e.g., residential, industrial). While we simulated the effects of "virtual" retrofits on real buildings, future case studies using the DUE-S modeling framework aim to use observed energy use data from a neighborhood that has previously gone through a large-scale retrofit program. By having access to data on electricity consumption both pre- and post-retrofit, we can better validate our model's retrofit analysis approach and conduct a sensitivity analysis to better quantify the role that urban context plays in how retrofits influence the energy use of surrounding buildings. And while the ownership structure of a building (e.g., owner/operator vs.

leasing tenant) is an important criterion when selecting energy conservation measures for specific buildings, our scope of work instead focused on how our model could generally predict future energy performance under various retrofits.

Finally, this project, as is the case with any urban energy modeling study, was limited in the number of buildings and length of analysis based on the availability of energy data for a dense cluster of urban buildings. Especially for deep learning models, where an extremely large number of observations are required to effectively train, validate, and test models, the accuracy of DUE-S is heavily reliant on the availability of a large dataset. While our model was able to produce significant results predicting existing electricity consumption and possible changes under virtual retrofits, a larger dataset – both in the number of observations and the number of buildings – may help broaden confidence in the usefulness of a hybrid modeling approach to a wide variety of urban environments worldwide.

7. Conclusion

This paper aimed to expand upon a previously introduced hybrid data-driven urban energy simulation (DUE-S) model that leverages building energy simulations and deep learning models. We modified the deep learning architecture to better capture the temporal nature of time-series building energy data and expanded the framework to predict the impact of various building energy retrofits on multiple spatiotemporal scales across a city. While simulation and deep learning have individually made significant progress in characterizing and highlighting buildings suitable for energy efficiency retrofits, we demonstrated the potential of a hybrid approach in a case study of 29 densely co-located commercial and mixed-use buildings in downtown Sacramento, California, USA. The primary objective of this paper was to demonstrate how an integrated data-driven and physics-based model could characterize the performance of various retrofit scenarios. The bulk of existing research in the building retrofit domain is segmented between simulation and data-driven approaches; however, the results of our case study show that by integrating these two types of modeling methods, we can unlock hidden insights how urban context can influence urban building retrofits. Our results showed that in accounting for inter-building effects, our energy prediction model can uncover the hidden relationships of how buildings influence each other's energy consumption. Additionally, DUE-S can provide insights on how to select buildings for retrofit that captures the potential "snowball" effect of urban context.

The DUE-S model can scale to any size building portfolio or urban area to yield insights on building energy consumption on multiple spatial and temporal scales. With the exception of whole-building energy use data, all of the data required to implement DUE-S is widely available and open-source, meaning that a variety of building and urban-level stakeholders could use this sort of model to make energy or sustainability-related decisions. For example, building designers and facility managers could use this integrated model to understand the compounding influence design and retrofit decisions have on not only their building, but the surrounding ones as well (e.g., window retrofits modifying the energy consumption of surrounding buildings). This is especially important for owners/operators of large building portfolios, such as universities or large technology campuses, where these effects can multiply their own financial, energy, or emissions consequences. This tool can also be used by policymakers to predict and visualize the effects of various retrofit scenarios to provide better awareness for how widespread retrofit measures may influence both building-level and broader urban energy efficiency. For example, our retrofit optimization algorithm showed how cities can maximize their overall energy savings while minimizing the number of required buildings needed to do so. But instead, if the algorithm were to maximize savings in carbon emissions, DUE-S could be used to help select retrofits for buildings that would reduce energy consumption during times of the day when the grid relied on more carbon-emitting fuel types (e.g., late afternoon/early evening when natural gas peaker plants are needed to meet energy demand). Finally, as the world's cities continue to expand, data is being increasingly used to plan their growth to meet the demand of an urbanizing population. DUE-S can leverage data from existing cities to help inform key decisions related to the energy efficiency of an expanding urban area (e.g., developing planning policies related to urban morphology and land use, predict future demand to size and plan city energy systems). In the end, tools capable of accurately predicting and characterizing urban building energy usage under various retrofit scenarios will be key in transitioning our cities to a more sustainable energy future.

Declaration of Competing Interest

None.

Acknowledgements

We would like to thank Sacramento Municipal Utility District (SMUD) for providing input data for our case study analysis. The research presented in this paper was supported in part by the Center for Integrated Facility Engineering and the U.S. National Science Foundation (NSF) under Grant No. 1941695. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of U.S. NSF.

Appendix

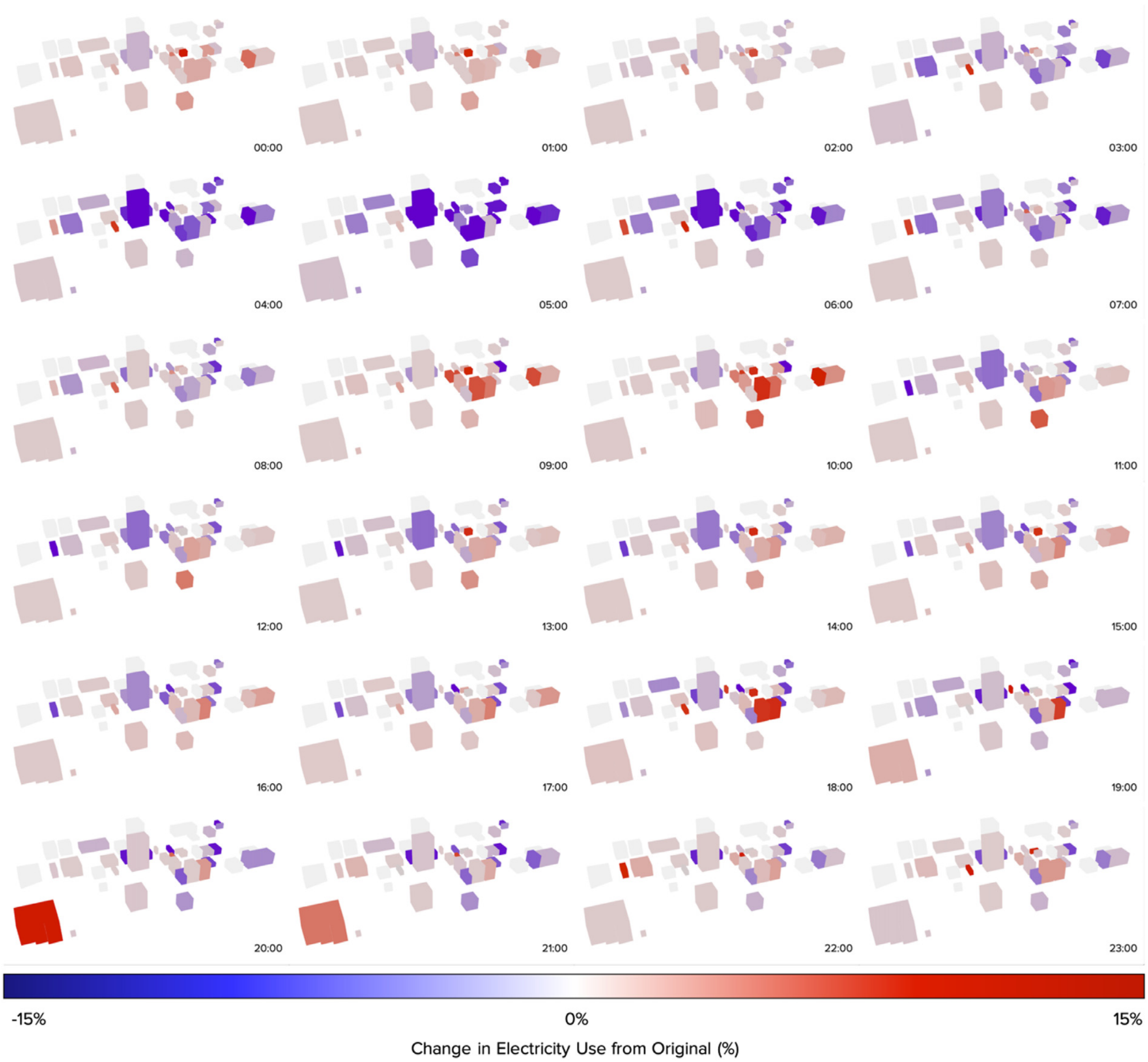

Fig. A1. Monthly building-level electricity consumption change for a *With Context*, full retrofit scenario in which all buildings receive the full retrofit.

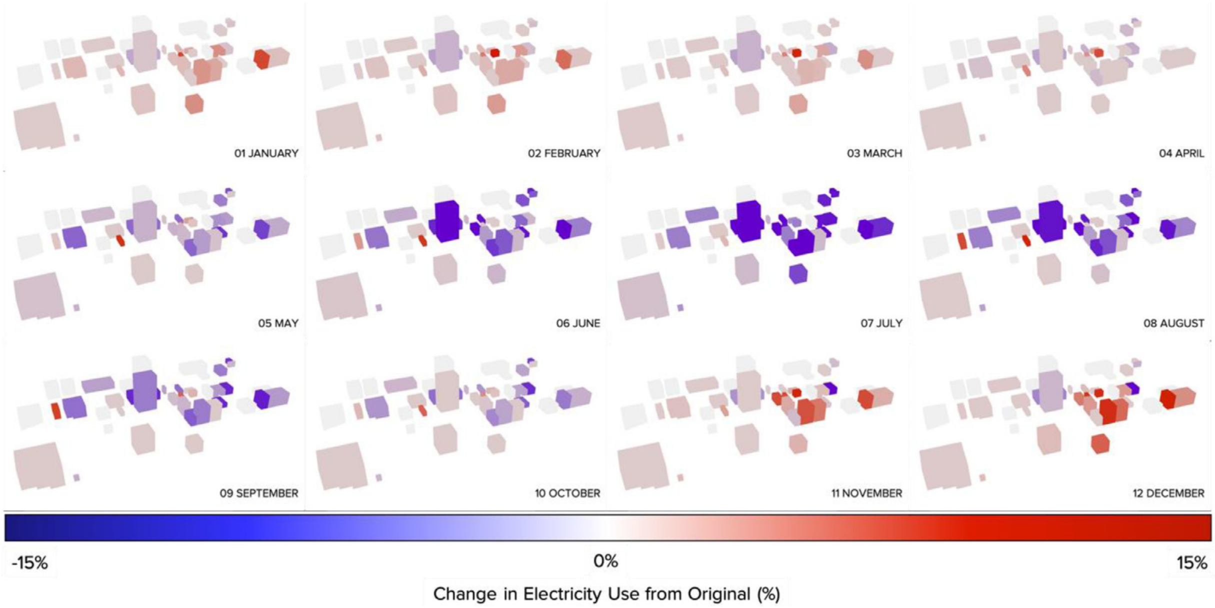

Fig. A2. Average hour-of-the-day building-level electricity consumption changes for a *With Context*, full retrofit scenario in which all buildings receive the full retrofit.

	Pre-1980 Construction	1980-2004 Construction	ASHRAE 90.1-2010 Construction
U-factor	5.84	5.84	1.82
Solar heat gain coefficient (SHGC)	0.54	0.44	0.25
Visible light transmittance (VLT)	0.38	0.27	0.2

Fig. A3. Updated Office window specifications as part of “Window” retrofit.

References

[1] World Urbanization Prospects: The 2018 Revision. 2019. doi:10.18356/b9e995fe-en.
[2] US Energy Information Administration (EIA). International Energy Outlook 2017 Overview. 2017.
[3] International Institute for Sustainable Development. 77 Countries, 100+ Cities Commit to Net Zero Carbon Emissions by 2050 at Climate Summit 2019. https://sdg.iisd.org/news/77-countries-100-cities-commit-to-net-zero-carbon-emissions-by-2050-at-climate-summit/ (accessed November 17, 2020).
[4] Buildings. Int Energy Agency 2020. https://www.iea.org/topics/buildings (accessed November 17, 2020).
[5] Architecture 2030. Why the Buildings Sector? 2020. https://architecture2030.org/buildings_problem_why/ (accessed November 17, 2020).
[6] Pisello AL, Taylor JE, Xu X, Cotana F. Inter-building effect: simulating the impact of a network of buildings on the accuracy of building energy performance predictions. Build Environ 2012;58:37–45. doi:10.1016/j.buildenv.2012.06.017.
[7] 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;174:407–19. doi:10.1016/j.energy.2019.02.183.
[8] Santos LGR, Afshari A, Norford LK, Mao J. Evaluating approaches for district-wide energy model calibration considering the Urban Heat Island effect. Appl Energy 2018. doi:10.1016/j.apenergy.2018.01.089.
[9] Seyedzadeh S, Pour Rahimian F, Rastogi P, Glesk I. Tuning machine learning models for prediction of building energy loads. Sustain Cities Soc 2019;47:101484. doi:10.1016/j.scs.2019.101484.
[10] U.S. Department of Energy (DOE) Energy Efficiency & Renewable Energy Office. Commercial Buildings Integration Program 2020.
[11] Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, et al. Deep learning and process understanding for data-driven Earth system science. Nature 2019;566:195–204. doi:10.1038/s41586-019-0912-1.
[12] Liu Y, Racah E, Correa J, Khosrowshahi A, Lavers D, Kunkel K, et al. Application of deep convolutional neural networks for detecting extreme weather in climate datasets. ArXiv Prepr ArXiv160501156 2016.
[13] Karpatne A, Watkins W, Read J, Kumar V. Physics-guided neural networks (pgnn): an application in lake temperature modeling. arXiv 2017 ArXiv Prepr ArXiv171011431 n.d.
[14] Fer I, Kelly R, Moorcroft PR, Richardson AD, Cowdery EM, Dietze MC. Linking big models to big data: efficient ecosystem model calibration through Bayesian model emulation. Biogeosciences 2018. doi:10.5194/bg-15-5801-2018.
[15] Salim FD, Dong B, Ouf M, Wang Q, Pigliautile I, Kang X, et al. Modelling urban-scale occupant behaviour, mobility, and energy in buildings: a survey. Build Environ 2020. doi:10.1016/j.buildenv.2020.106964.
[16] Sonta AJ, Simmons PE, Jain RK. Understanding building occupant activities at scale: an integrated knowledge-based and data-driven approach. Adv Eng Informatics 2018. doi:10.1016/j.aei.2018.04.009.
[17] Gros A, Bozonnet E, Inard C, Musy M. Simulation tools to assess microclimate and building energy – A case study on the design of a new district. Energy Build 2016;114:112–22. doi:10.1016/j.enbuild.2015.06.032.
[18] Brian C, Wolfgang L, Alexander E. Operating energy demand of various residential building typologies in different European climates. Smart Sustain Built Environ 2018;7:226–50. doi:10.1108/SASBE-08-2017-0035.
[19] Dimoudi A, Nikolopoulou M. Vegetation in the urban environment: microclimatic analysis and benefits. Energy Build 2003;35:69–76. doi:10.1016/S0378-7788(02)00081-6.
[20] Xu X, AzariJafari H, Gregory J, Norford L, Kirchain R. An integrated model for quantifying the impacts of pavement albedo and urban morphology on building energy demand. Energy Build 2020. doi:10.1016/j.enbuild.2020.109759.
[21] Davis LW, Gertler PJ. Contribution of air conditioning adoption to future energy use under global warming. Proc Natl Acad Sci 2015;112:5962–7.
[22] 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;205:323–35. doi:10.1016/j.apenergy.2017.07.128.
[23] Han Y, Taylor JE, Pisello AL. Exploring mutual shading and mutual reflection inter-building effects on building energy performance. Appl Energy 2017;185:1556–64. doi:10.1016/j.apenergy.2015.10.170.
[24] Xie J. Can Reflective Buildings That Produce “Death Rays” Be Prevented? Bloom CityLab 2013 https://www.bloomberg.com/news/articles/2013-09-04/can-reflective-buildings-that-produce-death-rays-be-prevented (accessed November 17, 2020).
[25] Gorlé C. Improving Predictions of the Urban Wind Environment Using Data. Technol + Des 2019;3:137–41. doi:10.1080/24751448.2019.1640522.
[26] Strzalka A, Bogdahn J, Coors V, Eicker U. 3D City modeling for urban scale heating energy demand forecasting. HVAC&R Res 2011;17:526–39. doi:10.1080/10789669.2011.582920.
[27] Orehounig K, Evins R, Dorer V. Integration of decentralized energy systems in neighbourhoods using the energy hub approach. Appl Energy 2015;154:277–89. doi:10.1016/j.apenergy.2015.04.114.

[28] Jain RK, Qin J, Rajagopal R. Data-driven planning of distributed energy resources amidst socio-technical complexities. Nat Energy 2017;2:17112. doi:10.1038/nenergy.2017.112.
[29] Nutkiewicz A, Yang Z, Jain RK. Data-driven Urban Energy Simulation (DUE-S): a framework for integrating engineering simulation and machine learning methods in a multi-scale urban energy modeling workflow. Appl Energy 2018;225:1176–89. doi:10.1016/j.apenergy.2018.05.023.
[30] Ang YQ, Berzolla ZM, Reinhart CF. From concept to application: a review of use cases in urban building energy modeling. Appl Energy 2020. doi:10.1016/j.apenergy.2020.115738.
[31] Reinhart CF, Cerezo Davila C. Urban building energy modeling – A review of a nascent field. Build Environ 2016;97:196–202. doi:10.1016/j.buildenv.2015.12.001.
[32] Lauzet N, Rodler A, Musy M, Azam M-H, Guernouti S, Mauree D, et al. How building energy models take the local climate into account in an urban context – A review. Renew Sustain Energy Rev 2019;116:109390. doi:10.1016/j.rser.2019.109390.
[33] Crawley DB, Huang YJ. Does it matter which weather data you use in energy simulations. User News 1997;18:2–12.
[34] Hong T, Chen Y, Luo X, Luo N, Lee SH. Ten questions on urban building energy modeling. Build Environ 2020. doi:10.1016/j.buildenv.2019.106508.
[35] Li W. Quantifying the Building Energy Dynamics of Manhattan, New York City, Using an Urban Building Energy Model and Localized Weather Data. Energies 2020;13:3244.
[36] Bueno B, Norford L, Hidalgo J, Pigeon G. The urban weather generator. J Build Perform Simul 2013;6:269–81. doi:10.1080/19401493.2012.718797.
[37] Mauree D, Coccolo S, Kaempf J, Scartezzini J-L. Multi-scale modelling to evaluate building energy consumption at the neighbourhood scale. PLoS ONE 2017. doi:10.5281/zenodo.847291.Funding.
[38] Yang X, Zhao L, Bruse M, Meng Q. An integrated simulation method for building energy performance assessment in urban environments. Energy Build 2012;54:243–51. doi:10.1016/j.enbuild.2012.07.042.
[39] Zhao H, Magoulès F. A review on the prediction of building energy consumption. Renew Sustain Energy Rev 2012;16:3586–92. doi:10.1016/j.rser.2012.02.049.
[40] Howard B, Parshall L, Thompson J, Hammer S, Dickinson J, Modi V. Spatial distribution of urban building energy consumption by end use. Energy Build 2012;45:141–51. doi:10.1016/j.enbuild.2011.10.061.
[41] 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. doi:10.1016/j.enbuild.2017.12.040.
[42] Amarasinghe K, Marino DL, Manic M. Deep neural networks for energy load forecasting. In: 2017 IEEE 26th Int. Symp. Ind. Electron; 2017. p. 1483–8. doi:10.1109/ISIE.2017.8001465.
[43] Lü X, Lu T, Kibert CJ, Viljanen M. Modeling and forecasting energy consumption for heterogeneous buildings using a physical-statistical approach. Appl Energy 2015. doi:10.1016/j.apenergy.2014.12.019.
[44] Robinson C, Dilkina B, Hubbs J, Zhang W, Guhathakurta S, Brown MA, et al. Machine learning approaches for estimating commercial building energy consumption. Appl Energy 2017. doi:10.1016/j.apenergy.2017.09.060.
[45] Kontokosta CE, Tull C. A data-driven predictive model of city-scale energy use in buildings. Appl Energy 2017. doi:10.1016/j.apenergy.2017.04.005.
[46] Abbasabadi N, Ashayeri M, Azari R, Stephens B, Heidarinejad M. An integrated data-driven framework for urban energy use modeling (UEUM). Appl Energy 2019. doi:10.1016/j.apenergy.2019.113550.
[47] Xu X, Wang W, Hong T, Chen J. Incorporating machine learning with building network analysis to predict multi-building energy use. Energy Build 2019;186:80–97. doi:10.1016/j.enbuild.2019.01.002.
[48] Wang B, Lu J, Yan Z, Luo H, Li T, Zheng Y, et al. Deep uncertainty quantification: a machine learning approach for weather forecasting. In: Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min; 2019. p. 2087–95.
[49] Zhou B, Rybski D, Kropp JP. The role of city size and urban form in the surface urban heat island. Sci Rep 2017;7:4791. doi:10.1038/s41598-017-04242-2.
[50] Silva MC, Horta IM, Leal V, Oliveira V. A spatially-explicit methodological framework based on neural networks to assess the effect of urban form on energy demand. Appl Energy 2017. doi:10.1016/j.apenergy.2017.05.113.
[51] 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. doi:10.1016/j.enbuild.2018.03.020.
[52] Cerezo Davila C, 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. doi:10.1016/j.energy.2016.10.057.
[53] 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;154:321–34. doi:10.1016/j.enbuild.2017.08.029.
[54] Kontar R El Polly B, Charan T. URBANopt: an Open-source Software Development Kit for Community and Urban District Energy Modeling. 2020 Build Perform Anal Conf SimBuild; 2020.
[55] Chen Y, Deng Z, Hong T. Automatic and rapid calibration of urban building energy models by learning from energy performance database. Appl Energy 2020;277:115584. doi:10.1016/j.apenergy.2020.115584.
[56] Coakley D, Raftery P, Keane M. A review of methods to match building energy simulation models to measured data. Renew Sustain Energy Rev 2014;37:123–41. doi:10.1016/j.rser.2014.05.007.
[57] Nagpal S, Reinhart CF. A comparison of two modeling approaches for establishing and implementing energy use reduction targets for a university campus. Energy Build 2018. doi:10.1016/j.enbuild.2018.05.035.
[58] 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. doi:10.1016/j.enbuild.2016.10.050.
[59] Booth AT, Choudhary R, Spiegelhalter DJ. A hierarchical bayesian framework for calibrating micro-level models with macro-level data. J Build Perform Simul 2013. doi:10.1080/19401493.2012.723750.
[60] Fonseca JA, Nguyen T-A, Schlueter A, Marechal F. City Energy Analyst (CEA): integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts. Energy Build 2016;113:202–26. doi:10.1016/j.enbuild.2015.11.055.
[61] Dogan T, Reinhart C. Shoeboxer: an algorithm for abstracted rapid multi-zone urban building energy model generation and simulation. Energy Build 2017;140:140–53. doi:10.1016/j.enbuild.2017.01.030.
[62] Chung W. Review of building energy-use performance benchmarking methodologies. Appl Energy 2011;88:1470–9. doi:10.1016/j.apenergy.2010.11.022.
[63] Roth J, Lim B, Jain RK, Grueneich D. Examining the feasibility of using open data to benchmark building energy usage in cities: a data science and policy perspective. Energy Policy 2020;139:111327. doi:10.1016/j.enpol.2020.111327.
[64] Kontokosta CE, Tull C. Energyviz: web-based eco-visualization of urban energy use from building benchmarking data. Proc. Int. Conf. Comput. Civ. Build. Eng. 2016.
[65] Park JY, Yang X, Miller C, Arjunan P, Nagy Z. Apples or oranges? Identification of fundamental load shape profiles for benchmarking buildings using a large and diverse dataset. Appl Energy 2019;236:1280–95. doi:10.1016/j.apenergy.2018.12.025.
[66] Baasch GM, Evins R. Targeting Buildings for Energy Retrofit Using Recurrent Neural Networks with Multivariate Time Series. Neural Inf. Process. Syst. 2019:2019.
[67] Ali U, Shamsi MH, Bohacek M, Hoare C, Purcell K, Mangina E, et al. A data-driven approach to optimize urban scale energy retrofit decisions for residential buildings. Appl Energy 2020;267:114861. doi:10.1016/j.apenergy.2020.114861.
[68] Seyedzadeh S, Pour Rahimian F, Oliver S, Rodriguez S, Glesk I. Machine learning modelling for predicting non-domestic buildings energy performance: a model to support deep energy retrofit decision-making. Appl Energy 2020;279:115908. doi:10.1016/j.apenergy.2020.115908.
[69] Yang Z, Becerik-Gerber B. A model calibration framework for simultaneous multi-level building energy simulation. Appl Energy 2015;149:415–31. doi:10.1016/j.apenergy.2015.03.048.
[70] Deru M., Field K., Studer D., Benne K., Griffith B., Torcellini P., et al. U.S. Department of Energy commercial reference building models of the national building stock. Publ 2011.
[71] Loga T, Stein B, Diefenbach N. TABULA building typologies in 20 European countries—Making energy-related features of residential building stocks comparable. Energy Build 2016;132:4–12. doi:10.1016/j.enbuild.2016.06.094.
[72] Rahman A, Srikumar V, Smith AD. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. Appl Energy 2018;212:372–85. doi:10.1016/j.apenergy.2017.12.051.
[73] Laib O, Khadir MT, Mihaylova L. Toward efficient energy systems based on natural gas consumption prediction with LSTM Recurrent Neural Networks. Energy 2019;177:530–42. doi:10.1016/j.energy.2019.04.075.
[74] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Comput 1997;9:1735–80. doi:10.1162/neco.1997.9.8.1735.
[75] Kim T-Y, Cho S-B. Predicting residential energy consumption using CNN-LSTM neural networks. Energy 2019;182:72–81. doi:10.1016/j.energy.2019.05.230.
[76] Solcast. Solcast 2020. https://www.solcast.com (accessed November 17, 2020).
[77] County of Sacramento. Assessor Parcel Viewer 2020. https://assessorparcelviewer.saccounty.net/jsviewer/assessor.html (accessed November 17, 2020).
[78] Bergmeir C, Benítez JM. On the use of cross-validation for time series predictor evaluation. Inf Sci (Ny) 2012;191:192–213. doi:10.1016/j.ins.2011.12.028.
[79] ASHRAE. Advanced Energy Design Guides n.d. https://www.ashrae.org/technical-resources/aedgs (accessed March 9, 2021).
[80] ASHRAE. Measurement of energy and demand savings. ASHARE Guidel 14-2002 2002.
[81] EVO. International Performance Measurement and Verification Protocol. 2012.
[82] Federal Energy Management Program. Federal Energy Management Program, M&V Guidelines: Measurement and Verification for Federal Energy Projects Version 3.0. 2008.
[83] Sevlian RA, Rajagopal R. A model for the effect of aggregation on short term load forecasting. 2014 IEEE PES Gen. Meet. | Conf. Expo.; 2014. p. 1–5. doi:10.1109/PESGM.2014.6938899.
[84] Neukomm M., Nubbe V., Fares R. Grid-Interactive Efficient Buildings. 2019.