# A Novel Graph Neural Network for Zone-Level Urban-Scale Building Energy Use Estimation

Eren Gökberk Halaçlı
Dept. of Comp. Eng., METU
ODTÜ-GÜNAM
Ankara, Türkiye
gokberk.halacli@odtugunam.org
gokberk.halacli@metu.edu.tr

İlkim Canlı
Dept. of Arch., METU
ODTÜ-GÜNAM
Ankara, Türkiye
ilkim.canli@odtugunam.org
ilkim.canli@metu.edu.tr

Orçun Koral İşeri
Dept. of Arch., METU
Ankara, Türkiye
koral.iseri@metu.edu.tr

Feyza Yavuz
Dept. of Comp. Eng., METU
Ankara, Türkiye
feyza.yavuz@metu.edu.tr

Çağla Meral Akgül
Dept. of Civil Eng., METU
Ankara, Türkiye
cmeral@metu.edu.tr

Sinan Kalkan
Dept. of Comp. Eng., METU
Ankara, Türkiye
skalkan@metu.edu.tr

İpek Gürsel Dino
Dept. of Arch., METU
Ankara, Türkiye
ipekg@metu.edu.tr

## ABSTRACT

Buildings are highly responsible for total energy consumption in cities; therefore, accurate estimation of building energy consumption is essential for developing energy-efficient strategies on an urban scale. Data-driven urban building energy models can predict energy use with high precision and low computational cost. In recent years, machine learning, especially neural networks, emerged as the prominent method for predicting energy load for buildings. These models typically use different input features on building form, occupancy and operation. However, they remain inadequate in capturing the complex inter-dependencies (i.e., heat transfer) between units in multi-zone buildings, as they do not have an explicit representation of the neighborhood relations between zones. The precision of data-driven models can be improved using Graph neural networks (GNN) that can capture the underlying relationships and dependencies between different building elements. In this paper we propose a novel GNN model that surpasses the current state-of-art methods for the prediction of zone-level heating energy use. We applied the methodology in a residential neighborhood consisting of 5866 buildings and 64462 zones. We use zone-level features regarding their geometry, material thermal characteristics, internal loads as node features, inter-zone parameters as edge features (total area and U value of the adjacent surfaces) and weather parameters. The results showed that our proposed model provides improvements over alternative approaches for precise prediction of urban building energy consumption.



*BuildSys '23, November 15–16, 2023, Istanbul, Turkey*

ACM ISBN 979-8-4007-0230-3/23/11.
https://doi.org/10.1145/3600100.3623747

## CCS CONCEPTS

• **Applied computing → Environmental sciences**; • **Computing methodologies → Neural networks**.

## KEYWORDS

Urban Building Energy Modeling, Graph Neural Network, Urban-scale Building Energy Consumption

**ACM Reference Format:**
Eren Gökberk Halaçlı, İlkim Canlı, Orçun Koral İşeri, Feyza Yavuz, Çağla Meral Akgül, Sinan Kalkan, and İpek Gürsel Dino. 2023. A Novel Graph Neural Network for Zone-Level Urban-Scale Building Energy Use Estimation. In *The 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '23), November 15–16, 2023, Istanbul, Turkey*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3600100.3623747

## 1 INTRODUCTION

Cities consume about 75% of the world's energy and cause more than 70% of greenhouse gas emissions [31]. Reducing the energy consumption of buildings therefore plays an important role in reducing global carbon emissions [30]. Accurate prediction of building energy use at the urban level is essential for developing energy-efficient strategies, and assessing the impact of climate change on buildings to accelerate cities' clean energy transition [16]. In this context, urban building energy modelling (UBEM) is a viable approach to make informed decisions to reduce energy consumption and CO2 emissions [11].

Urban building energy models can be physics-based and data-driven models [8]. Physics-based models are based on fundamental thermophysical mechanisms, providing reliable estimation of energy performance through simulations. However, the development of energy models is time consuming especially in an urban context, due to requirements of the large amount of data and high-level of modelling expertise [21]. Alternatively, data-driven UBEM approaches can learn from existing data without the need for physics-based models and simulations [4]. Data-driven UBEMs, unlike physics-based models, focus on capturing patterns in existing data. They are also preferable due to their highly accurate performance results at lower computational costs compared to physics-based UBEMs [22].

Despite their advantages, data-driven models still have some limitations. Most data-driven UBEMs capture target outcomes based on building-level parameters, (i.e., thermal properties of building elements, internal loads, occupant behaviours, energy systems), assuming that buildings are independent [17]. However, building energy consumption depends not only on building or zone characteristics, but also on inter-dependencies between zones or buildings [17].

In reality, buildings in close proximity affect each other in several ways. Especially when the buildings are interconnected, the interactions between adjacent surfaces have a significant impact on energy consumption and indoor thermal conditions [15]. However, there is a knowledge gap on how building interactions should be handled in data-driven UBEM models. In order to fill this gap, this paper proposes a novel graph neural network (GNN) model for the prediction of urban-scale building energy consumption. We applied the methodology in a residential neighbourhood consisting of 5866 buildings and 64462 zones. We present a comparative analysis of the performance of our proposed model with different machine learning (ML) models.

**Contributions**. The main contributions of this paper can be summarised as follows:

- To be able to consider different types of zone inter-dependencies, we extend a widely-used GNN model to consider relations between zones while making predictions. To be specific, we form a GNN that includes not only node embeddings but also edge embeddings, and augment the GraphSAGE [14] model by incorporating edge embeddings while passing messages between nodes.
- We show on a large-scale urban setting that the proposed approach performs better than an approach using multi-layer perceptron or GNN models.

## 2 RELATED WORK

### 2.1 Urban Building Energy Modelling (UBEM)

Data-driven UBEM have gained popularity since model development requires less laborious processes and benefit from the use of smart sensors and data mining approaches [17]. They provide important insights into the energy consumption patterns of buildings in urban areas by analysing building stock and developing strategies to reduce energy consumption. Data-driven UBEM can be grouped as statistical models and machine learning (ML) models [25]. Statistical models utilise explicit mathematical functions to forecast building energy consumption based on input data [22]. ML models, on the other hand, learn from existing data a mapping to estimate building energy consumption [27]. random forest (RF), support vector machines (SVMs) and artificial neural networks (ANNs) are the most common ML algorithms. Many surveys (e.g., [4, 27, 35] and [18]) reviewed existing studies that use machine learning models in prediction of building energy consumption.

### 2.2 Graph Neural Networks for UBEM

Nevertheless, the majority of data-driven UBEMs capture results based on the building level features, assuming that buildings are independent of one another. A prominent solution is to use graph neural networks (GNNs), which are deep learning algorithms based on graph representations that can model a set of objects and their relationships [37]. GNNs can use node and edge features to describe the relationships between buildings or zones. The representation of each node in a graph network is automatically defined by its attributes and the aggregate of the surrounding nodes because GNNs use a message passing mechanism [37]. Edges can be used to indicate the relationships between the interconnected buildings or zones, while nodes can be used to represent the particular building features, such as thermal characteristics, occupant behaviour, and internal loads. GNN models may learn about a building's energy consumption based on its own characteristics and propagate the energy impact of nearby buildings via graph structure by information propagation [17].

Hu et al. [17] proposed a spatio-temporal GNN model to predict hourly energy consumption of buildings based on graph representation. Their study considered solar-based inter-dependencies between 26 buildings. They concluded that graph-based data-driven models can greatly reduce building energy usage on an urban scale by comparing the performance of the proposed model with other time-series machine learning models. Lu et al. [21] developed a GNN model to learn the energy load patterns of each basic block. After applying their methodology on 800 buildings and comparing different data-driven models (ANNs, support vector regression (SVR), random forest (RF), and gradient boosting tree), their results indicated that the proposed GNN showed the highest performance accuracy. Chen et al. [9] developed a GNN-based cooling load prediction method. They compared the results with onsite data, concluding that the proposed GNN model is effective in cooling load prediction.

### 2.3 Comparative Summary

As reviewed above, we note that recent studies [9, 17, 21], have started exploring the use of GNN for UBEM, demonstrating the promise of representing relations among entities for UBEM. Differently from these studies:

- Existing approaches consider zones in isolation, ignoring the heat transfer between neighboring zones through shared internal surfaces. Especially when the indoor conditions (internal loads, set points, schedules) of neighboring zones are sufficiently different from one another, heat transfer between zones can constitute a significant portion of the total heat gains and losses.
- For more accurate estimation of energy use, we utilize GNNs that incorporate zone-level inter-dependencies. To be able to integrate zone-level inter-dependencies, our approach extends a GNN by taking into account the inter-dependency representations while making predictions.

- Prior GNN-based studies considered only small-scale neighborhoods with small number of zones such as 800 buildings with 4 to 7 zones per building in (Lu et al. [21]) whereas we validated our approach on an urban-scale with 5866 buildings with 64462 zones.

## 3 METHODOLOGY: GNN-BASED UBEM (GUBEM)

Our approach first constructs a graph by analyzing the inter-dependencies between zones (Section 3.1). For this, a region in the city of Ankara is selected and energy consumption values are obtained (Section 4). The graph constructed after the first step is used in a GNN and trained for UBEM (Section 3.2). See Figure 1 for an overview.

### 3.1 Zone Inter-dependency Analysis

Modeling zone inter-dependencies is an integral component of our contribution and has an effect on capturing heat transfer between neighboring zones. While representing relations between neighboring zones, we use the U value and the area of the wall connecting the zones as well as the floor difference between the zones. Using these parameters and features representing the zones, we model each building as homogeneous bidirectional graphs. Each node represents a zone and each bidirectional edge represents relations between neighboring zones.

### 3.2 Zone-level Graph Neural Networks for UBEM

*3.2.1 Background on Graph Neural Networks.* Common deep learning methods are applicable on Euclidean data. However, in the real world, data can be quite irregular and non-Euclidean. A prominent solution to this problem is to represent irregular structures in data with graphs. Graph Neural Networks have been successfully applied to different fields including such as Chemistry [12], Finance [33], Natural Language Processing [29], etc. A commonly-used GNN type is Message Passing GNNs that perform convolution operation using neighbours of nodes [36]. A message passing GNN computes new node representations by "passing messages" across each node's neighbours [37].

A GNN relies on a graph which consists of nodes connected with edges. Let $G = (V, E)$ be a graph with $V$ the set of nodes and $E$ the set of edges. Each node $v$ is represented by a feature vector $h_v \in \mathbb{R}^k$. $\mathcal{N}(v)$ is the set of direct neighbours of node $v \in V$. Then, message passing can be formalised as follows for a layer $l$ [12]:

$$h_u^{l+1} = \Phi\left(h_u^l, \bigoplus_{v \in \mathcal{N}(u)} \Psi(h_u^l, h_v^l)\right), \tag{1}$$

where $\Psi(\cdot,\cdot)$ is the message function which computes the first step of pair-to-pair message between node $u$ and its neighbour. $\bigoplus$ is the permutation invariant aggregator function that transforms multiple messages from distinct neighbours into a single vector. Finally, $\Phi(\cdot,\cdot)$ is the readout function that computes the node representation using the same nodes feature vector and aggregate message vector.

*3.2.2 Our Graph Neural Network.* The conventional message-passing mechanism in Section 3.2.1 is insufficient for our problem because properties representing the edges are not included while passing messages. To address those limitations, we propose a novel message-passing GNN inspired by GraphSAGE [14].

Our main idea is to enrich node-level features with embeddings that are computed using neighbour nodes and edges between them. The graph $G = (V, E)$ in our model is represented by a node feature matrix $V \in \mathbb{R}^{|V| \times k}$ as well as an edge feature matrix $E \in \mathbb{R}^{|E| \times m}$ (Figure 2). The feature representation for a node is denoted by $h_v^l$ for a node $v$ at layer $l$, that of a directed edge between nodes $v$ and $u$ is $e_{vu}$.

One layer of our model can be formalised as:

$$h_v^{l+1} = \sigma\left(W\left(h_v^l, \bigoplus_{u \in \mathcal{N}(v)} \alpha_{vu} \Psi(h_u^l, e_{vu})\right) + b\right), \tag{2}$$

where $\sigma(\cdot)$ is the non-linearity, $\alpha_{vu}$ is the attention weight, and $b$ is the bias. With this setting, MLP message function $\Psi(\cdot,\cdot)$ can learn more complex relations between nodes giving the model a higher representative power. When using dynamic attention score $\alpha_{vu}$ as in [6], each message is multiplied with the attention score between the central node and respective neighbour nodes. This helps the model to focus on important neighbours. After computing messages from every node in the neighbourhood, the model aggregates them with a permutation invariant aggregator $\bigoplus$ such as mean or sum functions. The result is the embedding of the one-step neighbourhood. Then this embedding vector is concatenated with the central node feature vector after that the resulting vector is fed to a fully connected layer with an activation function $\sigma(\cdot)$. The result is the next layer representation of the central node. As each layer is applied, nodes will receive signals from more of their neighbours.

## 4 CASE STUDY

We demonstrate our methodology in Ankara's Bahçelievler neighbourhood, a residential area covering 0.57 km$^2$. These buildings are representative of a wide range of thermal characteristics due to the different construction years, ranging from 1950 to 2022. The study area is located in ASHRAE's 4B climate zone, which is characterised by cold winters and warm, dry summers. In order to calculate energy consumption necessary to train the GNN model, we developed an urban building energy model (UBEM) followed by simulation-based data generation (heating energy consumption). 5866 buildings were modelled to the zone level (64462 zones), where each apartment unit is assumed to be a thermal zone and is assigned different semantic data, as explained below.

### 4.1 3D Model Development and Energy Modelling

UBEM requires geometric data and semantic data for the buildings in the neighbourhood. Geometric data includes physical characteristics as floor height, number of floors, building footprint, orientation, building dimensions, and local site conditions. Semantic data includes building function, construction year, construction material properties, internal loads, HVAC properties, setpoints and occupancy patterns.

Geometric data was procured from different sources. Building footprints were gathered from the Çankaya municipality. Data on

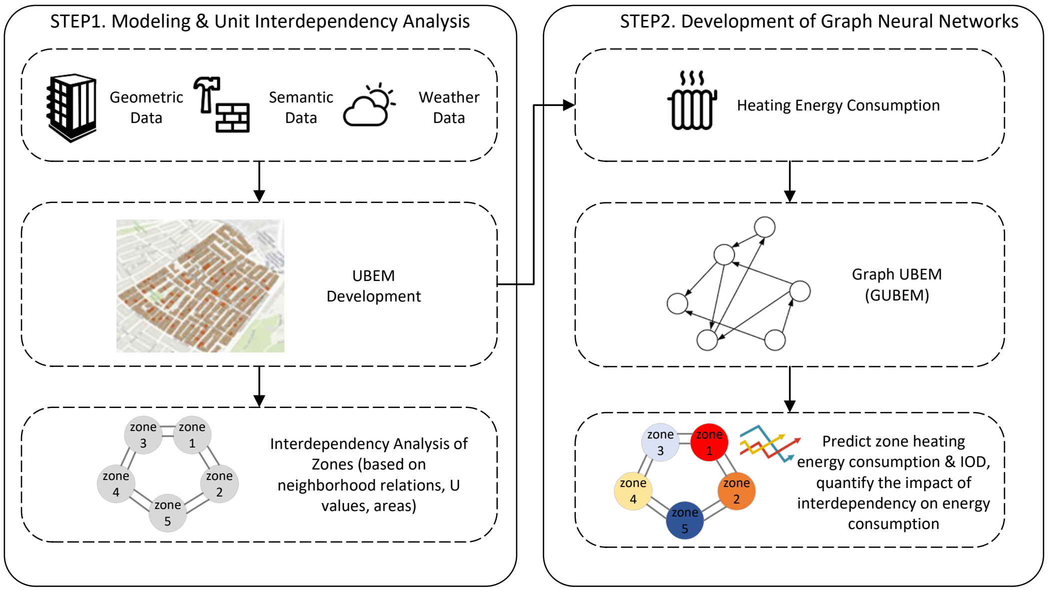



**Figure 1: An overview of the proposed methodology.**

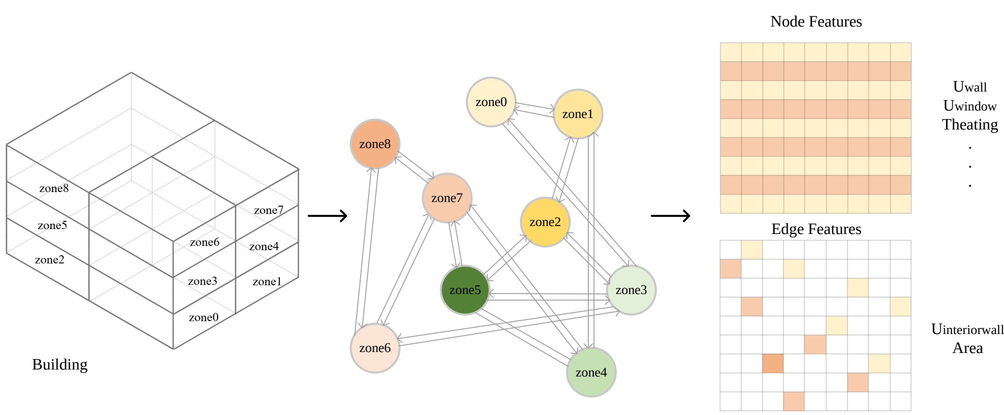



**Figure 2: A graph representation of zones in buildings.**

the number of floors, units, and their functionalities were gathered from the Ministry of Interior, General Directorate of Population and Citizenship Affairs [32]. Semantic data, on the other hand, were acquired from the Energy Performance Certificates (EPC) by the Turkish Ministry of Urbanization and Climate Change [10], which includes data on year of construction, and construction materials.

Each building and its residential units were simulated sequentially in an automated process. In parallel, each zone's attributes (which will be mostly used for GNN model training) were recorded in the UBEM dataset. During each simulation, the associated building semantic data was retrieved from the database. Subsequently, simulation results were tagged with the building's block and parcel numbers and logged in the same dataset.

In preparation of the UBEM datasets for missing data and data generation steps, two key density estimation techniques, namely parametric (PDE) and non-parametric (NPDE), are employed. Parametric Density Estimation (PDE) assumes that the density is represented by a known parametric family of distributions (e.g., normal or uniform distribution) with unknown parameters, where the missingness is equal to 100%, whereas Non-Parametric Distribution (NPDE) requires sampling from the associated distribution to generate the density function; it can be applied to data sets with

**Table 1: Node Feature Statistics**

| Abbreviation | Distribution | Source | Min | Max |
|---|---|---|---|---|
| $unit\,wwr$ (in 4 directions) | NPDE [1] | EPC | 0 | 47 |
| $U_{wall}$ | NPDE | EPC | 0.15 | 1.9 |
| $U_{roof}$ | NPDE | EPC | 0.147 | 2.42 |
| $U_{ground}$ | NPDE | EPC | 0.145 | 2.1 |
| $U_{window}$ | NPDE | EPC | 0.85 | 3.65 |
| $Q_{people}$ | NPDE | TUIK [3] | 0.002 | 0.09 |
| $T_{heating}$ | PDE | [26] | 19 | 25 |
| $Q_{lightning}$ | PDE | [34] | 2.5 | 12 |
| $Q_{equipment}$ | PDE | [7] | 1.75 | 8 |
| $SHGC$ | PDE | [13] | 0.301 | 0.849 |
| $Infiltration$ | PDE | [2] | 0.000286 | 0.0005 |
| $COP_{Boiler}$ | PDE | [28] | 0.8 | 0.95 |

**Table 2: Statistics of Simulation Results**

| | Mean | Min | Max | Std |
|---|---|---|---|---|
| $Q_H$ $(kWh/m^2)$ | 79.02 | 1.591 | 371.769 | 49.02 |

$0 < missingness < 100\%$. When the missingness of the data is greater than zero, we reproduced the missing data using EPCs with non-parametric distribution and assigned them to the buildings that have no data (data imputation). These data include UWall, URoof, UGround, UWindow, WWR and Qpeople. In the study, data on occupants (except for person density), SHGC, COPBoiler, and infiltration rate were completely missing. For these data, we generated new data based on PDE using uniform distribution (data generation). Table 1 shows the statistics of dataset based on data imputation and generation.

## 4.2 Simulation-Based Data Generation

We made our simulations using EnergyPlus 9.2 with real data calibration. Heating energy demand ($Q_H$) is calculated as a performance objective as the heating load is predominate in Ankara. Figure 3 and Table 2 show the distribution and statistics of $Q_H$. The mean of Heating End Use is calculated as 79.02 $Kwh/m^2$ and while standard deviation is calculated as 49.02. The Heating End Use ranges between 1.591 and 371.769 $Kwh/m^2$.

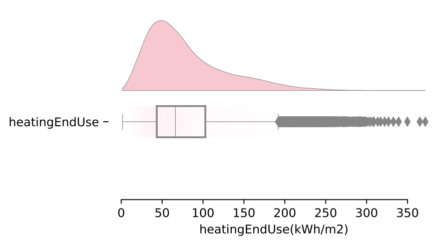



**Figure 3: Distribution of $Q_H$ (kWh/m2)**

# 5 EXPERIMENTS AND RESULTS

## 5.1 Implementation and Training Details

To predict heating demand for buildings, we used the final node feature representations produced by GNN model as an input to the MLP. In this setting, GNN acts as the "feature extractor" and MLP makes predictions on the final representations of the node features. MLP has four layers with ReLU [24] activation while GNN has Leaky ReLU [23] activation functions. We use L2 normalisation between every GCN layer and layer normalisation between every GCN layer and every MLP layer[5].

To get the best results, we made an extensive hyperparameter search for the number of GNN layers, the number of layers in the message function ($\Psi$), the output shape of the message function (embedding size), the aggregator type, the dynamic attention use, learning rate and batch size. We trained our models with the ADAM optimizer [19].

The dataset is split into three parts for training, validation and testing. Each part consists of 70, 15 and 15 percent of the dataset respectively.

We use the commonly-used $R^2$ score and Root Mean Squared Error (RMSE) as performance metrics.

## 5.2 Compared (Baseline) Models

We selected MLP, GraphSAGE [14], GCN [20] and GAT [6] as baseline models. All these models use only node (zone) features and do not consider inter-dependencies (edge features) between zones.

For a fair comparison, we trained the baseline models by performing a hyperparameter search with parameters listed in Table 3. For MLP, we used ReLU [24] non-linearity, as it performed better. For GNN baselines, we used the final representations of the nodes as input for their MLP predictors.

## 5.3 Energy Estimation: Quantitative Results

We predicted heating load using the setup from the previous section and compared our model with the baseline models described in Section 5.2. Note that the baseline models also work at the zone-level but do not take into account features (discussed in Section 3.1) between the neighboring zones.

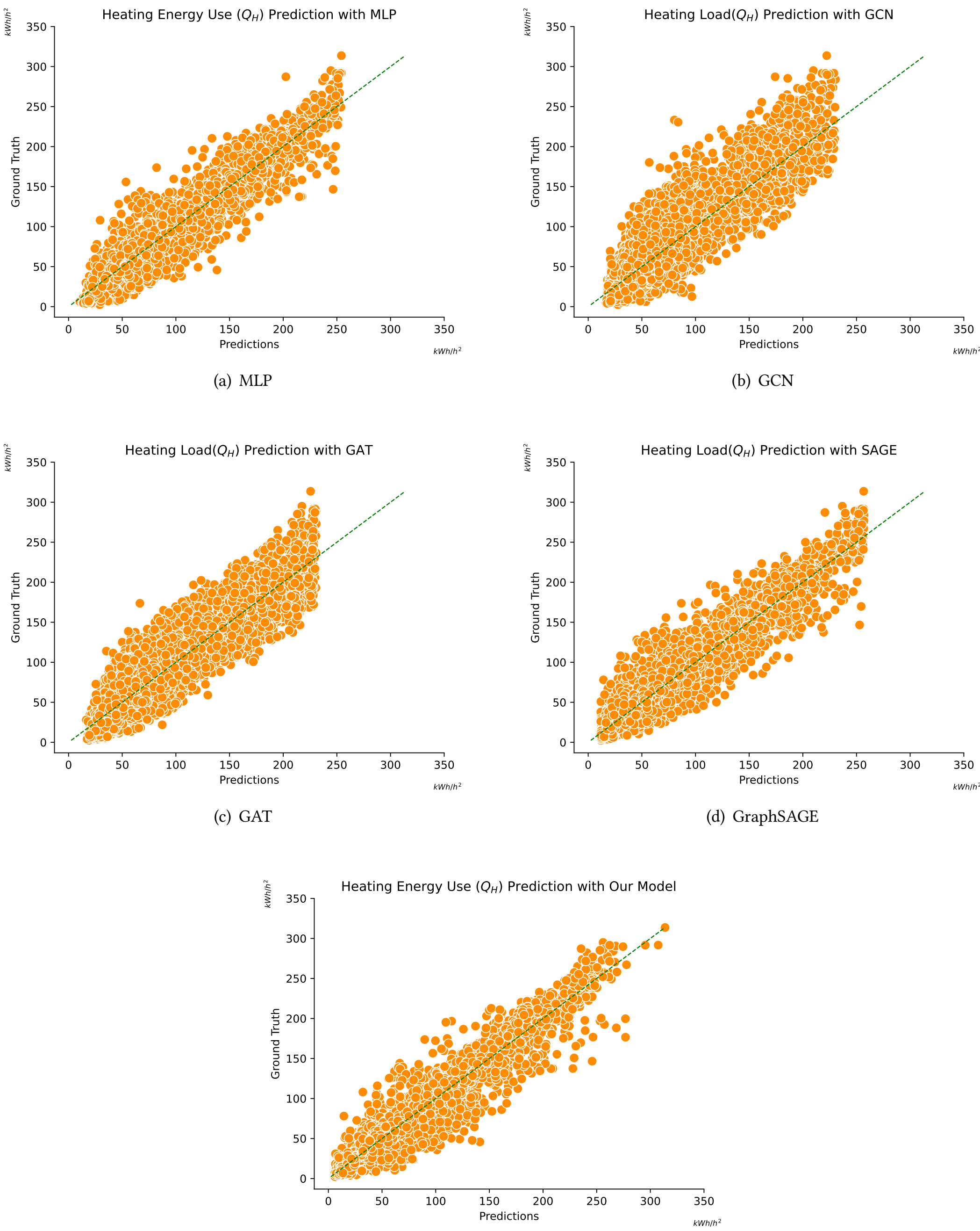
The quantitative results of all methods are provided in Table 4. The $R^2$ scores and RMSE values in the table suggest that our approach (GUBEM) provides the best values compared to not only MLP but also other GNN-based approaches. Considering that the baseline GNN approaches are also zone-based, the performance of GUBEM acknowledges the importance of capturing the interdependence between zones with explicit features.

## 5.4 Energy Estimation: Qualitative Results

We also compared the performances of the methods by plotting the predictions with respect to the ground truth values. The plots in Figure 4 show that our model generalizes better than the baseline models. We argue that our model's ability to use inter-zone relations with its message passing mechanism yields higher performance in node-level prediction.

# 6 CONCLUSION

In this research, we introduced GNN-based UBEM (GUBEM), a novel message passing GNN approach designed to predict building energy consumption at an urban level, taking into account inter-zone relationships. Our approach incorporates both node features and edge features, capturing the connections between different




(a) MLP

(b) GCN

(c) GAT

(d) GraphSAGE

(e) GUBEM (Ours)

**Figure 4: Heating load estimation results on the test set.**

zones. The significant contribution of our study lies in the explicit consideration of inter-zone relationships when predicting building energy consumption on an urban scale.

To evaluate our model, we conducted experiments on a dataset comprising 5866 buildings, totaling 64462 zones. The results demonstrated higher performance compared to multi-layer perceptrons and other GNN frameworks in predicting building energy consumption at this urban scale. The high performance of our model is

**Table 3: Hyperparameters tuned for baseline models.**

| Model | Hyperparameters |
|---|---|
| MLP | learning rate, hidden layer size, batch size |
| GraphSAGE & GAT &GCN | learning rate, hidden layer size (MLP), batch size, convolution layer number, hidden shape between convolution layers, output shape of the last convolution layer |

**Table 4: Heating Load Estimation Metrics ($R^2$ & RMSE)**

| Model | Training Set | Validation Set | Test Set |
|---|---|---|---|
| MLP | 0.9168 14.1106 | 0.9092 14.7020 | 0.9100 14.7111 |
| GCN | 0.7966 22.0621 | 0.7664 23.5834 | 0.7684 23.6100 |
| GAT | 0.8305 20.1381 | 0.8184 20.7930 | 0.8190 20.8731 |
| GraphSAGE | 0.9189 13.9312 | 0.9028 15.2116 | 0.9015 15.3952 |
| GUBEM (**Ours**) | **0.9336 12.6017** | **0.9217 13.6517** | **0.9202 13.8548** |

promising to generalise our model into a framework with different tasks (edge level prediction, graph classification, etc.) and domains.

Despite its contributions, our study is limited in that it only considers inter-zone relationships, overlooking potential valuable information from inter-building relationships. Our work can be extended to include both inter-zone and inter building relationships. In addition, it is promising to apply our approach for estimating cooling energy and indoor overheating degree, which we leave as future work. Finally, using using our model with measured data would bring more reliable results.

## ACKNOWLEDGMENTS

This study is funded by the Scientific and Technological Research Council of Turkey under grant number TUBİTAK 120M997.

## REFERENCES

[1] [n. d.]. Binalarda Enerji Performansı Uygulaması. https://beptr.csb.gov.tr/bep-web/

[2] 2013. ASHRAE Standard 90.1-2013, Energy Standard for Buildings Except Low-Rise Residential Buildings (SI Edition). Online. www.techstreet.com

[3] 2021. Türkiye Family Structure Survey. Ankara. https://data.tuik.gov.tr/Bulten/Index?p=Turkiye-Family-Structure-Survey-2021-45813&dil=2

[4] Kadir Amasyali and Nora M. El-Gohary. 2018. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews* 81 (2018), 1192–1205. https://doi.org/10.1016/j.rser.2017.04.095

[5] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. 2016. Layer Normalization. (2016). arXiv:1607.06450 [stat.ML]

[6] Shaked Brody, Uri Alon, and Eran Yahav. 2022. How Attentive are Graph Attention Networks? (2022). arXiv:2105.14491 [cs.LG]

[7] Henrik Brohus, Per Heiselberg, A. Simonsen, and K.C. Sørensen. 2009. Uncertainty of Energy Consumption Assessment of Domestic Buildings. In *Building Simulation 2009 : University of Stractclyde, Glosgow, 27th-30th July*, Paul A. Strachan, Nick J. Kelly, and Michaël Kummert (Eds.). The International Building Performance Simulation Association Conference. The International Building Performance Simulation Association Conference (IBPSA) ; Conference date: 27-07-2009 Through 30-07-2009.

[8] Yongbao Chen, Mingyue Guo, Zhisen Chen, Zhe Chen, and Ying Ji. 2022. Physical energy and data-driven models in building energy prediction: A review. *Energy Reports* 8 (2022), 2656–2671. https://doi.org/10.1016/j.egyr.2022.01.162

[9] Zhiwen Chen, Zhengrun Zhao, Qiao Deng, Peng Tang, Chunhua Yang, Xinhong Li, and Weihua Gui. 2023. A knowledge embedded graph neural network-based cooling load prediction method using dynamic data association. *Energy and Buildings* 278 (2023), 112635. https://doi.org/10.1016/j.enbuild.2022.112635

[10] Csb.gov.tr/. [n. d.]. *T.C. Çevre, Şehircilik ve İklim Değişikliği Bakanlığı.* https://csb.gov.tr/

[11] Sanam Dabirian, Karthik Panchabikesan, and Ursula Eicker. 2022. Occupant-centric urban building energy modeling: Approaches, inputs, and data sources - A review. *Energy and Buildings* 257 (2022), 111809. https://doi.org/10.1016/j.enbuild.2021.111809

[12] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. 2017. Neural Message Passing for Quantum Chemistry. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70* (Sydney, NSW, Australia) *(ICML'17)*. JMLR.org, 1263–1272.

[13] National Glass. 2019. *Glass Performance Data.* https://www.nationalglass.com.au/wp-content/uploads/2019/06/Glass-Data_v4-Low-Res.pdf

[14] William L. Hamilton, Rex Ying, and Jure Leskovec. 2017. Inductive Representation Learning on Large Graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (Long Beach, California, USA) *(NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, 1025–1035.

[15] Jiang He, Akira Hoyano, and Takashi Asawa. 2009. A numerical simulation tool for predicting the impact of outdoor thermal environment on building energy performance. *Applied Energy* 86, 9 (2009), 1596–1605. https://doi.org/10.1016/j.apenergy.2008.12.034

[16] Yuqing Hu, Xiaoyuan Cheng, Suhang Wang, Jianli Chen, Tianxiang Zhao, and Enyan Dai. 2022. Times series forecasting for urban building energy consumption based on graph convolutional network. *Applied Energy* 307 (2022), 118231. https://doi.org/10.1016/j.apenergy.2021.118231

[17] Yuqing Hu, Xiaoyuan Cheng, Suhang Wang, Jianli Chen, Tianxiang Zhao, and Enyan Dai. 2022. Times series forecasting for urban building energy consumption based on graph convolutional network. *Applied Energy* 307 (2022), 118231. https://doi.org/10.1016/j.apenergy.2021.118231

[18] Mohamad Khalil, A. Stephen McGough, Zoya Pourmirza, Mehdi Pazhoohesh, and Sara Walker. 2022. Machine Learning, Deep Learning and Statistical Analysis for forecasting building energy consumption — A systematic review. *Engineering Applications of Artificial Intelligence* 115 (2022), 105287. https://doi.org/10.1016/j.engappai.2022.105287

[19] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[20] Thomas N. Kipf and Max Welling. 2017. Semi-Supervised Classification with Graph Convolutional Networks. (2017). arXiv:1609.02907 [cs.LG]

[21] Jie Lu, Chaobo Zhang, Junyang Li, Yang Zhao, Weikang Qiu, Tingting Li, Kai Zhou, and Jianing He. 2022. Graph convolutional networks-based method for estimating design loads of complex buildings in the preliminary design stage. *Applied Energy* 322 (2022), 119478. https://doi.org/10.1016/j.apenergy.2022.119478

[22] Jun Ma and Jack Cheng. 2016. Estimation of the building energy use intensity in the urban scale by integrating GIS and big data technology. *Applied Energy* 183 (09 2016), 182–192. https://doi.org/10.1016/j.apenergy.2016.08.079

[23] Andrew L Maas, Awni Y Hannun, Andrew Y Ng, et al. 2013. Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, Vol. 30. Atlanta, GA, 3.

[24] Vinod Nair and Geoffrey E. Hinton. 2010. Rectified Linear Units Improve Restricted Boltzmann Machines. In *Proceedings of the 27th International Conference on International Conference on Machine Learning* (Haifa, Israel) *(ICML'10)*. Omnipress, Madison, WI, USA, 807–814.

[25] Alex Nutkiewicz, Zheng Yang, and Rishee K. Jain. 2018. 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. *Applied Energy* 225 (2018), 1176–1189. https://doi.org/10.1016/j.apenergy.2018.05.023

[26] Pil Brix Purup and Steffen Petersen. 2020. Research framework for development of building performance simulation tools for early design stages. *Automation in Construction* 109 (2020), 102966. https://doi.org/10.1016/j.autcon.2019.102966

[27] Ying Sun, Fariborz Haghighat, and Benjamin C.M. Fung. 2020. A review of the-state-of-the-art in data-driven approaches for building energy prediction. *Energy and Buildings* 221 (2020), 110022. https://doi.org/10.1016/j.enbuild.2020.110022

[28] Wei Tian, Jitian Song, Zhanyong Li, and Pieter de Wilde. 2014. Bootstrap techniques for sensitivity analysis and model selection in building thermal performance analysis. *Applied Energy* 135 (2014), 320–328. https://doi.org/10.1016/j.apenergy.2014.08.110

[29] Ming Tu, Jing Huang, Xiaodong He, and Bowen Zhou. 2020. Graph Sequential Network for Reasoning over Sequences. arXiv:2004.02001 [cs.CL]

[30] ubem.io 2030. *Ubem.io.* Retrieved July 25, 2023 from https://www.ubem.io/

[31] UN Habitat. 2021. Energy. https://unhabitat.org/. Accessed: March 9, 2021.

[32] Nüfus ve Vatandaşlık İşleri Genel Müdürlüğü. [n. d.]. *Nüfus ve Vatandaşlık İşleri Genel Müdürlüğü.* https://www.nvi.gov.tr/

[33] Jianian Wang, Sheng Zhang, Yanghua Xiao, and Rui Song. 2022. A Review on Graph Neural Network Methods in Financial Applications. *Journal of Data Science* 20, 2 (2022), 111–134. https://doi.org/10.6339/22-JDS1047

[34] Song Yang, Wei Tian, Eduard Cubi, QingXin Meng, YunLiang Liu, and Lai Wei. 2016. Comparison of Sensitivity Analysis Methods in Building Energy Assessment. *Procedia Engineering* 146 (2016), 174–181. https://doi.org/10.1016/j.proeng.2016.06.369 The 8th international cold climate HVAC Conference.

[35] Liang Zhang, Jin Wen, Yanfei Li, Jianli Chen, Yunyang Ye, Yangyang Fu, and William Livingood. 2021. A review of machine learning in building load prediction. *Applied Energy* 285 (2021), 116452. https://doi.org/10.1016/j.apenergy.2021.116452

[36] Ziwei Zhang, Peng Cui, and Wenwu Zhu. 2022. Deep Learning on Graphs: A Survey. *IEEE Trans. on Knowl. and Data Eng.* 34, 1 (jan 2022), 249–270. https://doi.org/10.1109/TKDE.2020.2981333

[37] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. 2020. Graph neural networks: A review of methods and applications. *AI Open* 1 (2020), 57–81. https://doi.org/10.1016/j.aiopen.2021.01.001