DEVELOPING A DATA-DRIVEN FRAMEWORK FOR MULTI-SCALE INTEGRATED URBAN BUILDING AND TRANSPORTATION ENERGY MODELING

Author

Narjes Abbasabadi
Illinois Institute of Technology

Keywords

Urban energy use modeling, artificial intelligence, integrated building and transportation energy, urban socio-spatial context

Abstract

This article proposes an integrated data-driven framework for urban energy use modeling (UEUM) that enables providing a holistic image of urban energy use at multiple scales. The UEUM allows aggregating across end-uses, building, and transportation. With considering urban socio-spatial context, it gives insight into the multifaceted and intricate relationships between urban key attributes, and building and transportation energy performance. This model helps predict urban energy performance more precisely by reducing the simulation uncertainties through using disaggregated and spatially explicit data and applying artificial intelligence (AI) techniques. In addition to increasing the accuracy, the model facilitates reducing the execution time for an urban scale energy modeling. The framework was evaluated using Chicago, Illinois, a major city in the US, as a case study. The results for Chicago demonstrate the feasibility of this approach. Among the tested AI algorithms, k-nearest neighbor performed as the best model in terms of accuracy for a single-output model while artificial neural network algorithm showed the best overall performance for the integrated building and transportation energy use modeling.

Introduction

Buildings are the most significant contributor to urban energy use, followed by the transport sector (US EPA, 2017). The urban buildings and transportation energy performance are interrelated at various levels (Abbasabadi & Ashayeri J. K., 2019), and are influenced by the urban socio-spatial context (Liu et al., 2019). Hence, an integrated urban energy model is required that unifies building and transportation in one model and contextualizes the model with the actual urban socio-spatial context. This model can provide architects, urban designers and planners, and policymakers with tools to predict the urban energy and environmental impacts associated with alternative scenarios of city development. However, the existing methods and tools often have limitations in giving a realistic assessment of urban energy flows and aggregating across multi-scales and end-uses (Abbasabadi & Ashayeri J. K., 2019). Moreover, existing literature tends to examine a few aspects of urban energy determinants such as urban form, building characteristics, and human-related aspects. However, they lack a holistic approach.

The development of artificial intelligence (AI) algorithms, specifically those based on machine learning (ML) models, with rising data availability and quality, provide new possibilities for improving the precision and complexity of urban energy use models (Abbasabadi & Ashayeri, 2019). An AI simulation framework can provide the opportunity to predict and understand urban energy demand patterns and to explore the complex interrelations between urban energy determinants. This research proposes an integrated urban energy use modeling (UEUM) framework that applies AI approach, localizes the model, and considers the urban socio-spatial context. Then the results are visualized through using Geographic Information Systems (GIS) platform. The framework simulates energy use at multi-scale representing an individual building, block, neighborhood, and city scales. Chicago, Illinois, has been selected to evaluate the model. However, the model is replicable to all other cities.

Methodology

The UEUM framework employs artificial intelligence simulation approach for an integrated building and transport energy modeling. It combines building operational energy use prediction in the urban context with a travel demand model for transport energy use prediction. It uses the existing local data at building, neighborhood, and city levels, along with the GIS data, which are available for the major cities in the US. The framework explores actual socio-spatial patterns of the city to learn and extract new features and add localized variables in the model, and predicts urban energy use through learning the mathematical relationship between variables and tests different ML techniques and algorithms to propose an enhanced predictive model.

The key urban energy determinants related to the scope of this research include: 1) Building characteristics consisting of variables, such as building type, building height, building size, and construction year; 2) Urban attributes functioning as density, accessibility, connectivity, and land-use mixed, which are captured via urban sprawl index; 3) Occupancy characteristics including total population, household size in residential buildings, worker density in commercial buildings, and percentage of occupied units; 4) Socioeconomic indicators including income, unemployment, poverty level, dependency, education, and crowded housing variables; and 5) Mobility and travel patterns, which include neighborhood characteristics such as transit-oriented, walkability, bikeability indices, and mode of travel and travel distance.

Various datasets were used including urban spatial data, building characteristics, occupancy and socioeconomic indicators, building operational energy, and travel patterns data. The building characteristics are captured from the Chicago building footprints (CBF) dataset ("Building Footprints," 2015) which is a GIS-based dataset and represents a compilation of land use and geographic data for Chicago. The Chicago boundaries and zoning districts ("Boundaries-Zoning Districts," 2019) and Property Tax Data from the Assessor's Office ("Cook County Assessor Data," 2019) were used as supplementary datasets providing information such as land use and building age. The U.S. Urban Sprawl Data ("Updated Urban Sprawl Data for the United States," 2010) was also applied which represents multidimensionally geo-referenced urban attributes such as density, land use, activity centering, and accessibility factors. Chicago socioeconomic indicators dataset ("Census Data-Selected socioeconomic indicators in Chicago, 2008–2012," 2019) was used to represent households socioeconomic variables.

For building energy data, two datasets were used, including Chicago Energy Benchmarking dataset (City of Chicago, 2016) (2,717 buildings greater than 50,000 ft^2) and Chicago Energy Usage dataset (City of Chicago, 2010) (65,378 buildings of all sizes). The urban transport energy model is developed upon Chicago Regional Household Travel Tracker Survey by CMAP ("Household Travel Survey," 2016), Fuel Economy data ("Fuel Economy," 2019) by the U.S. Department of Energy (DOE), and Average passenger transportation energy intensity per mile travel from the U.S. Department of Transportation (DOT) Bureau of Transportation Statistics (BTS) ("National Transportation Statistics," 2019) datasets. Other mobility and travel patterns were extracted from neighborhood walkability, bikeability, and transit-oriented indices ("Chicago neighborhoods on Walk Score," 2019).

Several promising AI algorithms including Artificial Neural Networks (ANNs), K-Nearest Neighbors (k-NN), Random Decision Forest (RDF), and Regression Trees (RT) in addition to Nonlinear Regression (NLR) and Multiple Linear

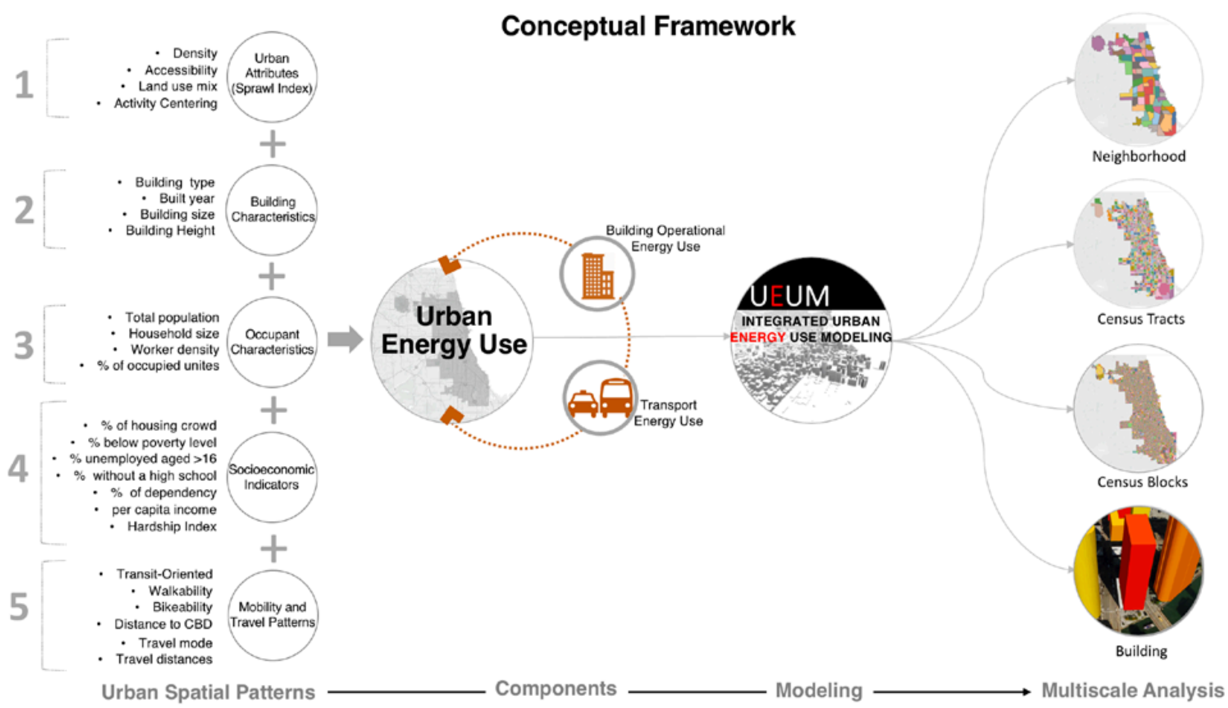

Figure 1: The UEUM Conceptual Framework.

Regression (MLR) algorithm as the most common statistical method used in the previous studies, were utilized in this research to propose the most accurate model for urban building and transportation energy use modeling. The models were validated using five-fold Cross-validation (CV) (Borra & Di Ciaccio, 2010), which is an effective validation method. As the next phase, the models were evaluated and compared in terms of their prediction performance as defined as metrics to calculate the errors between the predicted and actual values. The most widely used metrics for assessing the predictive models including the mean square error (MSE), mean absolute percentage error (MAPE) and R-squared (R^2) were used in this research. The lower the values of MSE, and MAPE and higher values of R^2 demonstrate the better performance of the model.

The travel demand modeling was developed to estimate transportation energy use per household for various modes of travel, including car and public transit across different neighborhoods. The urban transportation EUI was predicted by incorporating miles traveled, fuel economy of each travel modes, and energy intensity factors per the mode of travel which was adopted from previous studies as it is shown to be an effective method (Lindsey, Schofer, Durango-Cohen, & Gray, 2011; Norman, MacLean, M.ASCE, & Kennedy, 2006). As the next step, an integrated urban building and transportation model was developed which enables capturing complex relationships and contributions of urban energy key variables in the model. Since data for the urban transportation energy model is available only at the household level, and residential and non-residential buildings have different energy behaviors, the predictive and explanatory models for residential and non-residential buildings are simulated separately. Finally, GIS-based visualizations were developed to illustrate the predicted urban energy use across scales including building, block, neighborhood, and city levels. Figure 1 illustrates the UEUM conceptual framework.

Results

The scatterplots in Figure 2 illustrate the results of the performance of different AI simulation algorithms in terms of obtained actual versus predicted energy use values. The models were tested across the merged dataset of 58,205 observations to predict building EUIs for the 820,606 out-of-sample buildings in Chicago for which energy use data is not available. Ideally, all points should be close to a regressed diagonal line. For instance, if the Actual is 3, the predicted should be reasonably close to 3. The models were executed for predicting the EUI of all building types, including both residential and non-residential buildings. The results demonstrate the effectiveness of the k-NN algorithm for predicting building operational EUI in terms of accuracy for a single-output model, for only building energy use compared with the other AI algorithms tested in this research including ANNs, RDF, and RT, algorithms, and statistical methods including the MLR and NLR models. The results show that MLR has the weakest predictive performance in this case. As Figure 2 illustrates, the k-NN algorithm represents a MAPE of 1.832% while RT shows a MAPE of 4.8%, ANN, RDF, and NLR have a MAPE of 5.1%, and MLR has a MAPE of 5.3% in this model. The results indicate that k-NN is able to decrease the error by 65%, compared with the MLR method, which has been used in previous studies as a common method for energy prediction modeling. In terms of R^2, the k-NN, RT, ANNs, RDF, NLR, and MLR models showed R^2 (as calculated based on actual vs. predicted energy use) values of 0.75, 0.42, 0.33, 0.33, 0.29, and 0.28 respectively. It indicates that the best and weakest models are k-NN and MLR with explaining 75% and 28% of the variance in building operational EUI.

The integrated model combines the residential building EUI and household transportation EUI across 46,843 observations. The results of integrated building and transportation energy use model reveal that ANN performs the best compared with other tested algorithms. The results indicate that the urban attributes examined here using the ANN algorithm explain 41% and 96% of the variance in building and transportation energy use intensity, respectively. Figure 3 illustrates multiple-scale building and transportation energy modeling including a predictive model for energy consumption at individual building level (Figures 3A & B) and travel demand (Figure 3C), and transport energy use modeling (Figure 3D) at neighborhood scales.

The UEUM predictive model provides essential energy information at a multi-level. The framework has the potential to aid in understanding energy use patterns across the city with a high level of resolution to meet sustainable built environment plans. The model informs existing energy demand patterns and predicts future needs, which can help in retrofit programs and early-stage planning and design in line with energy efficiency policies. Through aggregating across multi-scales, it can be implemented as a more comprehensive decision-making tool. It identifies patterns of energy use and evaluates the trade-offs between different strategies or scenarios of development and their impacts on energy for building and transportation sectors.

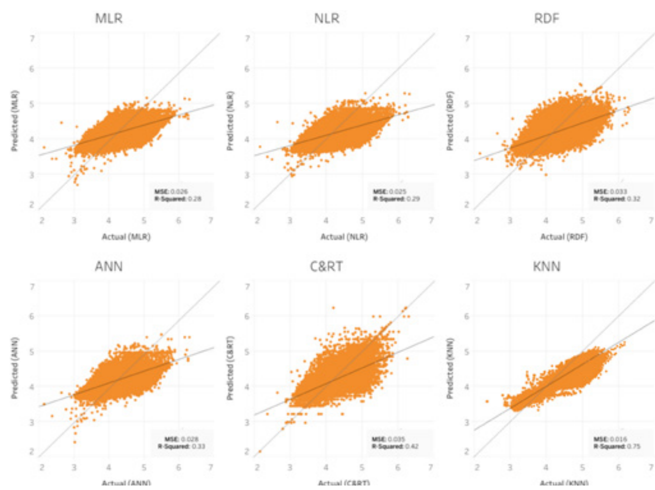

Figure 2: Scatterplots of the Actual vs Predicted Energy Use.

The results of this study demonstrate that all the urban attributes incorporated in the model (e.g., building characteristics, urban spatial patterns, occupancy, and socioeconomic factors, and mobility and travel factors) are vital predictors. These factors influence both building and transportation energy performance, however, with a different magnitude of impact. The results suggest that occupancy and socioeconomic indicators play an essential role in urban energy use models for both building and transportation EUI. Considering other influencing factors such as occupant behavioral factors and construction systems could contribute to a more comprehensive model for urban energy modeling.

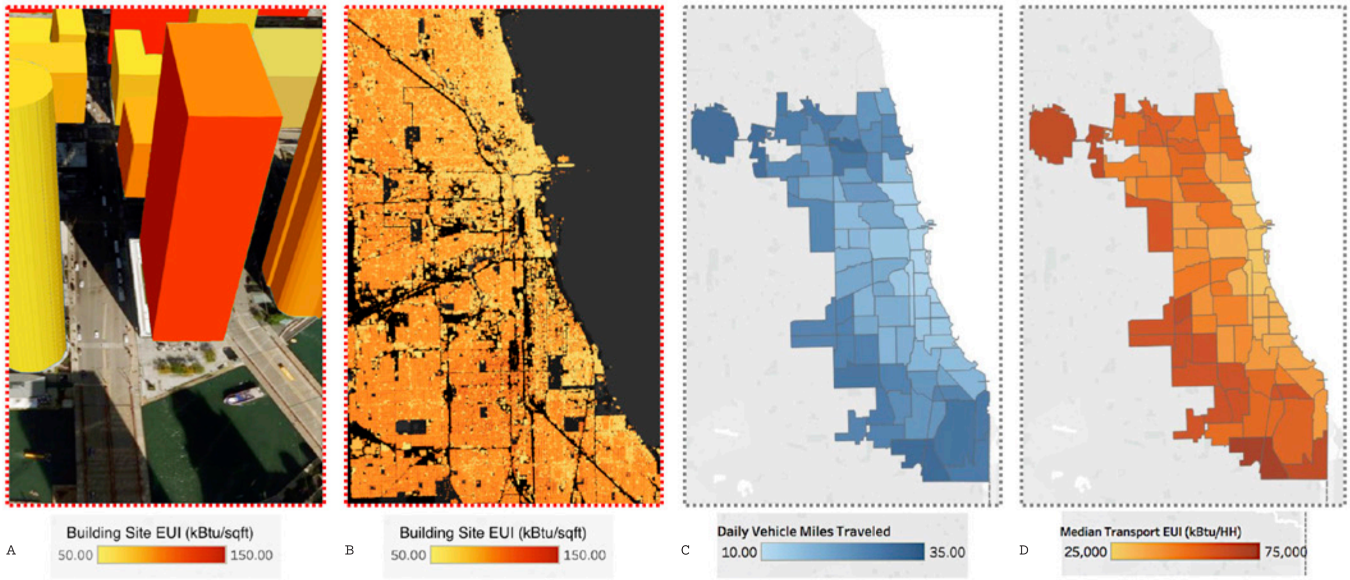

Figure 3: A Multi-Scale and Multi-Dimensional Urban Energy Use Modeling. A & B) Urban Building Energy Modeling. C) Travel Demand Modeling. D) Transportation Energy Modeling.

Conclusion

The integrated urban energy use modeling (UEUM) framework developed by this research enables predicting urban building and transportation energy at multi-scales of building, block, neighborhood, and city levels through applying an artificial intelligence (AI) based approach. The analysis results reveal that the urban energy prediction accuracy can be enhanced by utilizing building level disaggregated data and incorporating the actual urban socio-spatial factors. Also, more advanced AI algorithms enable an improved perdition model. Among the promising models tested in this research, the k-nearest neighbor (k-NN) showed the best predictive performance for single output model (e.g., building or transportation EUI as the target variable). While ANNs algorithm performed the best for an integrated model with two outputs, building and transportation simultaneously. The finding of this study also provides empirical evidence on existing energy demand profiles and how urban socio-spatial context impacts the building and transport energy performance. Finally, including urban spatial patterns along with socioeconomic and occupancy indicators can help more in-depth modeling of the integrated urban energy use. Future study suggests a quantitative analysis of the complex relationship between the urban socio-spatial patterns and building and transport EUI.

References

Abbasabadi, N., & Ashayeri, M. (2019). Urban energy use modeling methods and tools: A review and an outlook. *Building and Environment*, *161*, 106270. https://doi.org/10.1016/j.buildenv.2019.106270

Borra, S., & Di Ciaccio, A. (2010). Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. *Computational Statistics & Data Analysis*, *54*(12), 2976–2989. https://doi.org/10.1016/j.csda.2010.03.004

Boundaries-Zoning Districts (current). (2019). Retrieved February 5, 2019, from Chicago Data Portal website: https://data.cityofchicago.org/Community-Economic-Development/Boundaries-Zoning-Districts-current-/7cve-jgbp

Building Footprints (current). (2015). Retrieved February 5, 2019, from Chicago Data Portal website: https://data.cityofchicago.org/Buildings/Building-Footprints-current-/hz9b-7nh8?-category=Buildings&view_name=Building-Footprints-current-

Census Data-Selected socioeconomic indicators in Chicago, 2008–2012. (2019). Retrieved March 20, 2019, from Chicago Data Portal website: https://data.cityofchicago.org/Health-Human-Services/Census-Data-Selected-socioeconomic-indicators-in-C/kn9c-c2s2

Chicago neighborhoods on Walk Score. (2019). Retrieved April 16, 2019, from Walk Score website: https://www.walkscore.com/IL/Chicago

City of Chicago. (2010). Energy Usage 2010, City of Chicago, Data Portal. Retrieved February 5, 2019, from City of Chicago, Environment & Sustainable Development website: https://data.cityofchicago.org/Environment-Sustainable-Development/Energy-Usage-2010/8yq3-m6wp

City of Chicago. (2016). Chicago Energy Benchmarking 2016 Data Reported in 2017, City of Chicago, Data Portal. Retrieved from City of Chicago, Environment & Sustainable Development website: https://data.cityofchicago.org/Environment-Sustainable-Development/Chicago-Energy-Benchmarking-2016-Data-Reported-in-/fpwt-snya

Cook County Assessor Data. (2019). Retrieved May 6, 2019, from Cook County Government, Open Data website: https://datacatalog.cookcountyil.gov/

Fuel Economy. (2019). Retrieved February 5, 2019, from U.S. Department of Energy website: https://www.fueleconomy.gov/

Household Travel Survey. (2016). Retrieved February 5, 2019, from Chicago Metropolitan Agency for Planning (CMAP) website: http://www.cmap.illinois.gov/data/transportation/travel-survey

Lindsey, M., Schofer, J. L., Durango-Cohen, P., & Gray, K. A. (2011). The effect of residential location on vehicle miles of travel, energy consumption and greenhouse gas emissions: Chicago case study. *Transportation Research Part D: Transport and Environment*, *16*(1), 1–9. https://doi.org/10.1016/J.TRD.2010.08.004

Liu, X., Ou, J., Chen, Y., Wang, S., Li, X., Jiao, L., & Liu, Y. (2019). Scenario simulation of urban energy-related CO_2 emissions by coupling the socioeconomic factors and spatial structures. *Applied Energy*, *238*, 1163–1178. https://doi.org/10.1016/j.apenergy.2019.01.173

National Transportation Statistics. (2019). Retrieved February 5, 2019, from United States Department of Transportation, Bureau of Transportation Statistics website: https://www.bts.gov/topics/national-transportation-statistics

Norman, J., MacLean, H. L., M.ASCE, & Kennedy, C. A. (2006). Comparing high and low residential density: Life-cycle analysis of energy use and greenhouse gas emissions. *Journal of Urban Planning and Development*, (March 2006), 10–21. https://doi.org/10.1061/(ASCE)0733-9488(2006)132

Updated Urban Sprawl Data for the United States. (2010). Retrieved February 5, 2019, from National Cancer Institute, Geographic Information Systems & Science for Cancer Control website: https://gis.cancer.gov/tools/urban-sprawl/

US EPA, O. (2017, February 8). Inventory of U.S. Greenhouse Gas Emissions and Sinks [Reports and Assessments]. Retrieved June 22, 2019, from US EPA website: https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks