Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Interdependencies between electricity consumption and transit and their potential to forecast transit demand in urban settings

Juan Acosta-Sequeda

jacost42@uic.edu

University of Illinois Chicago

Ömer Verbas

Argonne National Laboratory, Transportation and Power Systems (TAPS)

Joshua Auld

Argonne National Laboratory, Transportation and Power Systems (TAPS)

Sybil Derrible

University of Illinois Chicago

Research Article

Keywords: Interrelated Urban Systems, Travel Demand, Electricity Demand, Time Series Forecasting, Machine Learning

Posted Date: August 14th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6959831/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Interdependencies between electricity consumption and transit and their potential to forecast transit demand in urban settings

Juan Acosta-Sequeda (Corresponding author)
Department of Civil, Materials, and Environmental Engineering
University of Illinois Chicago
1200 W Harrison St, Chicago, Illinois, United States, 60611
Email: jacost42@uic.edu

Ömer Verbas
Argonne National Laboratory, Transportation and Power Systems (TAPS)
9700 S. Cass Avenue, Lemont, IL, United States, 60439
Email: omer@anl.gov

Joshua Auld
Argonne National Laboratory, Transportation and Power Systems (TAPS)
9700 S. Cass Avenue, Lemont, IL, United States, 60439
Email: jauld@anl.gov

Sybil Derrible
Department of Civil, Materials, and Environmental Engineering
University of Illinois Chicago
1200 W Harrison St, Chicago, Illinois, United States, 60611
Email: derrible@uic.edu

Abstract

This study explores the interdependencies between electricity consumption and public transit demand to forecast transportation demand in urban settings. The premise is that changes in residential electricity usage reflect human activity that can be used to predict local transportation needs. By leveraging high-resolution electricity usage data from ComEd (Chicago's electricity provider) and transit ridership data from the Chicago Transit Authority (CTA), the study evaluates whether transit boardings can be predicted based on electricity usage changes. Three forecasting models are tested: Kalman Filter Forecasting, Long Short Term Memory Network, and Light Gradient Boosting Machine. The results indicate that LightGBM performs best, effectively capturing workday and weekend patterns. Including electricity usage as a covariate generally improves forecasting accuracy, although the benefits vary by model type. Moreover, we showcase the suitability of this approach when urban transit demand suffers from system shocks due to unforeseen phenomena such as the COVID-19 pandemic. The study concludes that electricity usage data can be a reliable predictor of transportation demand, particularly in high-activity zones. It highlights the potential for enhancing accuracy by integrating other transportation modes and non-residential electricity data.

Keywords: Interrelated Urban Systems, Travel Demand, Electricity Demand, Time Series Forecasting, Machine Learning

1 Introduction

Transportation, telecommunication, electricity, water, and other infrastructure networks are crucial elements of the functioning of cities (Derrible, 2025). They are also deeply interdependent and interrelated, not only in their design and operation but also in their fundamental goal of serving humans, who are ultimately agents in contact with one or more of these networks at different times (Derrible, 2019). A person who places a call enters in contact with the telecommunication and electricity networks; someone turning on the light to use the sink at night is making simultaneous use of the electricity grid, the water distribution system, and the wastewater network.

One of the main issues with operating such systems lies in demand management. Allocating resources effectively so the demand is met with the available infrastructure is central to urban sustainability and resilience (Derrible, 2018). Hence, demand modeling is critical and constantly in development in transportation and other fields (Cirillo and Axhausen, 2001; Ma *et al.*, 2018; Wang *et al.*, 2022; Zhong *et al.*, 2023; Xiong *et al.*, 2024; Chauhan *et al.*, 2025). Given the interaction between urban systems – since the driving force is human behavior – it is safe to assume that the demand for different systems is not independent (Amini *et al.*, 2018). In fact, they are likely interrelated to the point that, given the demand for one urban system, we can estimate the demand for another. A person leaving their home to go to the office will stop using electricity at home and then start a trip, from which we can see that an increase follows a decrease in energy demand in transportation (Acosta-Sequeda *et al.*, 2023; Movahedi *et al.*, 2023). Similarly, a person who just ended a trip is likely to enter a building and start consuming electricity (Lim and Kang, 2022).

Thanks to the adoption of smart meters in many cities, electricity usage data have a high quality compared to those of other infrastructure systems (Berger *et al.*, 2022). Smart meters allow providers to have detailed usage data at the account level with a time resolution of minutes (Tureczek and Nielsen, 2017). In this work, we use electricity usage data to estimate transit demand. We do so by using Chicago as a case study, where Commonwealth Eddison (ComEd) and the Chicago Transit Authority (CTA) are the electric provider and transit agency, respectively.

The relationship between transportation and land use has been extensively studied through various theoretical frameworks and models(Acheampong and Silva, 2015). The classical model of land use and transportation interaction suggests that transportation infrastructure influences land use patterns and vice versa (Mills, 1967; Alonso, 2013). Recent advancements have incorporated concepts from complexity theory and network science, recognizing that urban systems are not only interconnected but also dynamically adaptive (Batty, 2013). Many studies have highlighted the importance of integrated planning approaches. For example, the integration of

transit-oriented development (TOD) strategies has been shown to enhance accessibility and reduce traffic congestion in metropolitan areas (Cervero and Kockelman, 1997). In cities like Portland and Hong Kong, coordinated planning between land use and transportation has led to more sustainable urban growth patterns (Newman and Kenworthy, 1999; Lee and Vuchic, 2005). Urban economies are influenced by a range of factors, including the distribution of resources, labor markets, and economic policies (Chen, Matsuoka and Liang, 2018).

Social interactions and demographic trends significantly impact urban systems. Studies on social networks and community dynamics reveal how social cohesion and neighborhood characteristics influence urban resilience and quality of life (Putnam, 2000). Another study found that social interaction in the digital space, such as social media, can be useful to estimate telecommuting rates at the state level (Acosta-Sequeda, 2024). Urban systems are deeply intertwined with environmental factors. The impact of urbanization on ecosystems, air quality, and climate change is well-documented (Seto, Güneralp and Hutyra, 2012). Research on “sustainable urbanism” focuses on reducing environmental footprints through green infrastructure, energy-efficient buildings, and sustainable transportation options (Beatley, 2011). The concept of "resilience" is central to understanding how urban systems can adapt to environmental stresses and disasters (Davoudi *et al.*, 2012; Derrible, 2019).

In this study, we build upon previous research where we found early evidence of interrelationships between electricity consumption patterns in the Chicago area and the travel behavior of city dwellers evidenced in their mode choice and their commute time to work (Acosta-Sequeda *et al.*, 2023). Leveraging more granular data in terms of transportation demand, we push this line of research forward by proposing an application for transportation demand modeling. Our main hypothesis is that changes in electricity usage in a given zone in a city are indicators of human activity that can be used to forecast local transportation demand. If residential electricity usage decreases it may be an indication of people leaving a building, and, if it increases, it may indicate people arriving. In general, electric usage and transit demand follow the same 24-hour pattern with two asymmetrical demand peaks corresponding to morning and evening demand maximums. However, there is often a slight time shift between the electricity and transit demand peaks (Acosta-Sequeda *et al.*, 2023), which is one of our main motivations behind this work. Our main goal is to determine if such an apparent association can be used to estimate zonal transportation demand from electricity usage readings.

This study follows two steps (Figure 1). First, we find the model best suited to our application, which is framed as a time series model to leverage time series features such as seasonality, trends, and covariates, all common attributes of transportation demand data (Hunt and Ninomiya, 2003; Lee and Gwon, 2011). To select the best modeling method, we implement all the different models in a single station (for rail) and zip code (for bus). The selection of these aims to select the best model based on typical locations in terms of ridership and distance to the central business district (CBD). The electricity data is assigned to each bus stop and rail station based on proximity, where we used a distance of 660 meters based on the average walking distance to transit in American cities (Van Soest, Tight and Rogers, 2020). The hyperparameter tuning for each model is performed with the Grid Search method (Liashchynskyi and Liashchynskyi, 2019); the hyperparameters will be chosen based on the best scores on the test dataset. In addition, each model is also tested with and without electricity usage data as covariates. Second, once the models for each case (bus and rail) are selected, we proceed to present its application as a case study for the entire city of Chicago.

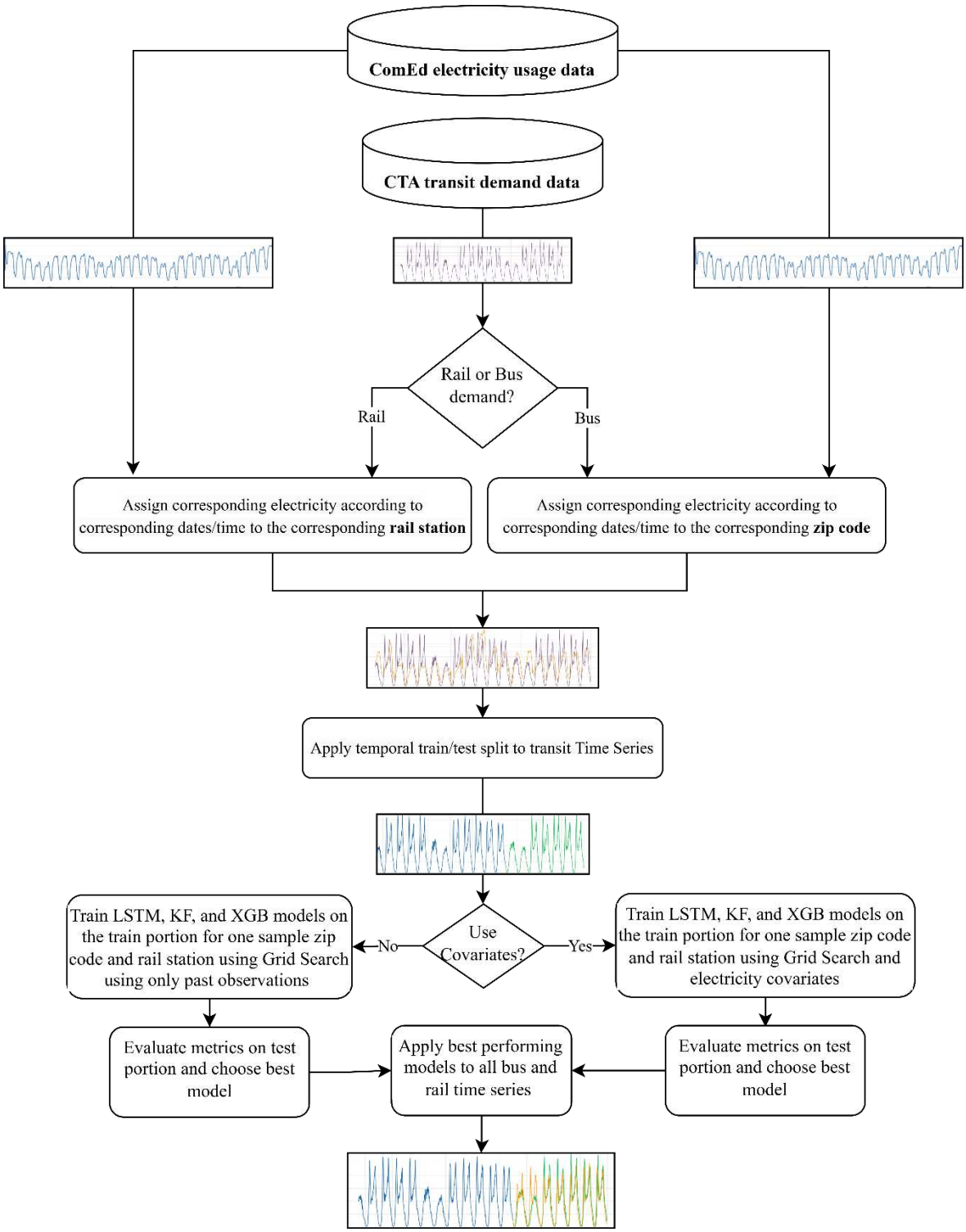


Figure 1. Flow chart depicting the methodology from the data integration process to the time series forecasting results.

2 Methods

2.1 Data

The data employed comprise two major infrastructure systems: the electricity grid and the transportation system. For electricity demand data, we use ComEd records. For transit ridership we use data provided through a partnership between Argonne National Laboratory (ANL) and the CTA.

The ComEd data consists of anonymized records of electricity consumption of all the accounts with smart meters in the ComEd area of service, which comprises Cook County and several other adjacent counties in northeastern Illinois. The records contain residential-level data for around 3.9 million accounts with a zip-code-level geographical resolution. Usage is reported in kWh every 30 minutes.

The CTA data consist of bus and rail boardings at each bus stop and train station aggregated in 30-minute intervals at the zip code level. The original boardings dataset provided by the CTA contains point data, but it was aggregated to the zip code level to match the geographic resolution in the ComEd dataset. For this specific work, we extracted 21 days worth of bus boardings. Due to data availability constraints, we could not use bus and rail data corresponding to the same time window. Hence, the bus data spans Sunday, September 1st, 2019, to Saturday, September 21st, 2019. The rail data spans Tuesday, October 1st, 2019, to Monday, October 21st, 2019. Both datasets contain full records on these dates. We chose 2019 data because it was the last pre-pandemic period when both electricity and transit data were available.

2.2 Time Series Forecasting

We use time series (TS) forecasting methods as our modeling approach using the Darts library (Herzen *et al.*, 2022). This is based on our prior knowledge of the seasonal nature of transportation demand, which is also true for electricity demand. We explore three different time series forecasting models—LSTM, KFF, and LGBM (see below)—with and without covariates and test them to identify the best-performing approach based on common metrics. The best performing method of the three is explained in detail below, after the initial method section process.

2.2.1 Long Short-Term Memory Recurrent Neural Networks (LSTM)

LSTM is a type of recurrent neural network (RNN) designed to capture long-range dependencies in sequential data. Unlike traditional RNNs, which struggle with vanishing gradients when learning long sequences, LSTMs use special gating mechanisms to regulate the flow of information. These gates—input, forget, and output—control what information is stored, forgotten, or output at each time step. This allows LSTMs to remember important information over long periods, making them particularly effective for tasks like speech recognition, language modeling, and time series prediction. LSTM retains relevant information across many time steps, addressing one of the key limitations of earlier RNN models, which makes them particularly useful for time series forecasting.

2.2.2 Kalman Filter Forcasting (KFF)

KFF is an algorithm designed to estimate the state of a linear dynamic system using noisy measurements. It is used to produce estimates of unknown variables over time, combining predictions and observations to refine the estimates. It operates in two steps: prediction and update. During the prediction step, the filter estimates the current state based on the previous state and a system model. In the update step, the estimate is corrected based on new observations using a weighted average that minimizes the error variance. This results in an optimal estimate of the state. KFF can handle noisy data effectively, adapts to changes over time, and is suitable for systems with underlying linear dynamics.

2.2.3 LightGBM (LGBM)

LGBM is a popular gradient boosting framework that efficiently handles large datasets and high-dimensional features. For time series forecasting, LGBM can be adapted to capture temporal patterns and

trends. Time series data needs to be transformed into a format suitable for LGBM. This consists of using a sliding window representation of the data, where the time series is converted into a supervised learning problem by using past observations as independent variables and future observations as dependent ones. Similarly to the conventional usage of Gradient Boosting Machines (GBMs), the model iteratively improves its performance by combining weak learners (decision trees) into a strong predictive model.

2.2.4 Covariates

In time series forecasting, a covariate is an additional variable used to improve the accuracy of predictions. Unlike the primary time series data, which is the main variable of interest (such as stock prices or temperature readings), covariates provide supplementary information that can help model the relationship between the primary time series and external factors that might influence it. A covariate is an external variable or feature related to the time series and can affect its behavior. It can enhance the forecasting model by providing context or additional information that the primary time series alone may not fully capture. One common example is using economic indicators such as unemployment rates or GDP as covariates for forecasting. Weather data can predict energy consumption, where weather variables like temperature or humidity can be useful covariates. This study hypothesizes that electricity consumption can be used as a covariate of the number of bus boardings, which is the target time series.

2.3 Metrics

To evaluate the time series estimations, we use the following three metrics:

- Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2} \tag{1}$$

- Coefficient of variation (CV)—a normalized measure of the variation between two time series expressed as a percentage, calculated as follows:

$$CV = 100 \cdot \frac{RMSE(y_t, \hat{y}_t)}{\bar{y}_t} \tag{2}$$

where y_t are the actual values, $\hat{y}_t$ are the predicted values, $\bar{y}_t$is the average of y_t, and RMSE is the root mean squared error.

- Mean Absolute Percentage Error (MAPE):

$$MAPE = 100 \cdot \frac{1}{T}\sum_{t=1}^{T}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \tag{3}$$

2.4 Model Selection

To select the best time series modeling approach, we applied each method to the a single zip code and rail station based on the assumption that they represent the average location in terms of geography and ridership.

The models are implemented with and without electricity covariates, giving us both an indication of which method performs best and a first look into the effect of introducing electricity data as a covariate. In each instance, the models are trained (i.e., calibrated) on the first twelve days and tested on the remaining seven days. The parameters for each model were selected by using the Grid Search approach (see Supplementary Materials). Figure 2 and Figure 3 display the results of the three different methods for bus and rail, respectively.

From the figures, we see that the LSTM and KFF perform poorly without electricity demand covariates. LGBM performs considerably better without covariates, although it struggles to capture weekend variations for both the bus and rail demand. On the versions with covariates, we can see that the three methods perform much better, demonstrating a better capacity to reproduce the two-peak daily pattern. However, the LSTM model does not capture the increased demand on weekdays and seems to reach a steady maximum demand for every day of the week. Similarly, KFF presents similar daily patterns for all days of the week, failing at representing the difference between weekend and weekday demand.

For both bus and rail, LGBM seems to capture this difference; displaying the typical two-peak daily pattern for work demand is also more accurate for the LGBM method with the covariates implementation.

Finally, Figure 1 summarizes the adopted method selection and modelling approaches followed in this document.

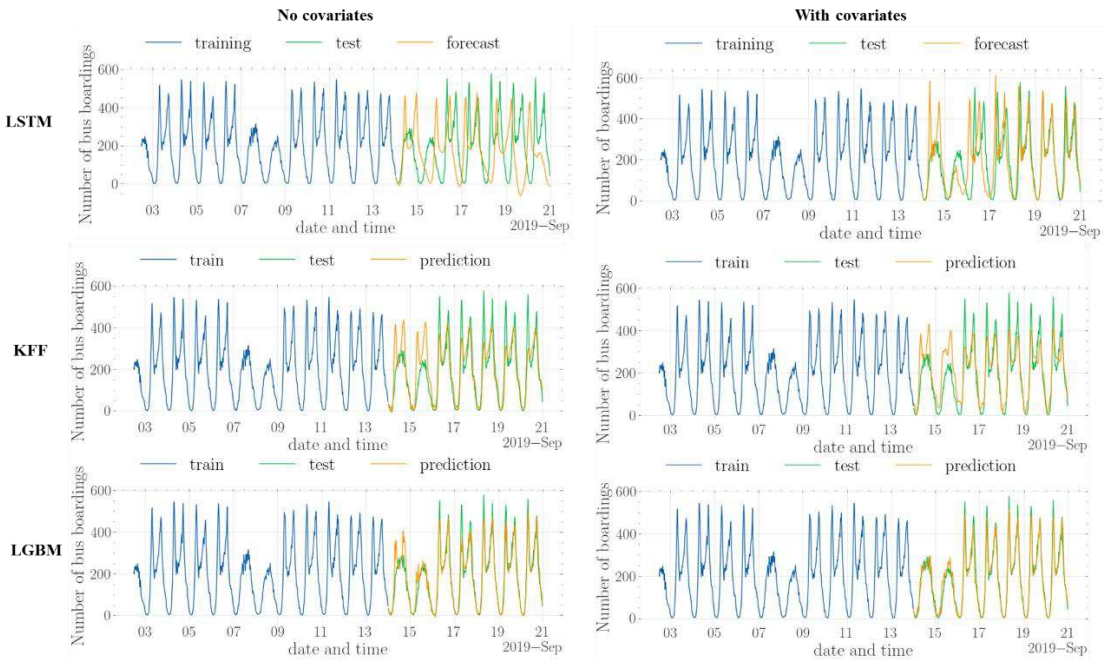


Figure 2. Results of calibrating the LR, KFF, and LGBM methods with and without covariates for the bus ridership time series on the selected zip code.

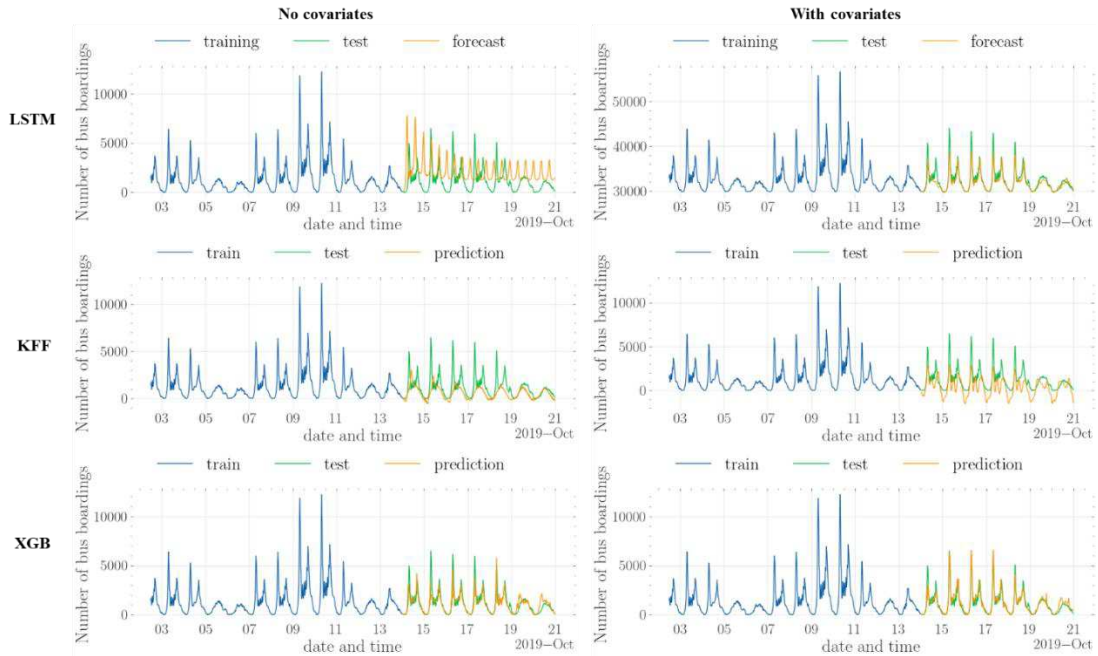


Figure 3. Results of calibrating the LR, KFF, and LGBM methods with and without covariates for the rail ridership time series on the selected zip code.

Table 1 contains the numeric results of the metrics obtained from each one of the implementations described. The results from this table are consistent with the observations made from Figure 2 and Figure 3, where the implementations with covariates outperforms the one without covariates in every case except for the KFF in the rail mode, where both models fail. In addition, Table 1 confirms that, among the three tested methods, LGBM performs best. Hence, it is the one we will use in the rest of this study.

Table 1. Reported metrics after implementing the best model specification for bus and rail with and without covariates.

Mode	Model	RMSE	MAPE	CV
Bus	**LSTM**			
	No covariates	202.63	735.28	121.08
	With covariates	127.18	67.58	63.39
	KFF			
	No covariates	81.43	45.54	40.61
	With covariates	72.40	32.61	34.25
	LGBM			
	No covariates	45.25	29.38	24.23
	With covariates	32.67	35.23	17.38
Rail	**LSTM**			
	No covariates	1776.60	64.64	86.41
	With covariates	0.04	72.82	55.29
	KFF			
	No covariates	1147.10	273.53	205.39
	With covariates	1101.56	307.83	287.11
	LGBM			
	No covariates	660.39	51.18	59.87
	With covariates	445.12	43.62	36.71

2.5 A closer look into the selected model: LGBM

The foundation of LGBM is Gradient Boosting, an ensemble learning technique that builds a strong predictive model by sequentially combining weak learners, typically decision trees. For time series forecasting, it is adapted to predict future values based on past observations and exogenous features as follows: given a univariate time series dataset:

$$\{(x_t, y_t)\}_{t=1}^{T}$$

where $x_t \in \mathbb{R}^d$ is the feature vector at time t, which may include past values $(y_{t-1}, y_{t-2}, \dots)$ and exogenous variables, $y_t \in \mathbb{R}$ is the target value at time t, and T is the total number of observed time steps, the objective is to forecast future values y_{T+h} for a given forecast horizon h, which translates into estimating the function:

$$F: \mathbb{R}^d \to \mathbb{R}, \quad \text{such that} \quad y_{T+h} \approx F(x_T)$$

Gradient Boosting constructs an additive model as:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$$

where $F_m(x)$ is the model prediction at iteration m, $F_0(x)$ is the initial prediction, often set to the mean of y_t, $h_m(x)$ is the weak learner trained at iteration m, γ_m is the learning rate, which controls the contribution of each weak learner.

To optimize the model, a differentiable loss function $L\big(y, F(x)\big)$ is minimized, where common choices are Mean Squared Error (MSE) and Mean Absolute Error (MAE).

At each iteration m, the pseudo-residuals (negative gradients) are computed as:

$$g_{im} = -\left[\frac{\partial L\big(y_i, F(x_i)\big)}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}$$

For the MSE loss, the gradient simplifies to:

$$g_{im} = y_i - F_{m-1}(x_i)$$

which represents the difference between the true value and the current model's prediction.
Each weak learner $h_m(x)$ is trained to fit the computed gradients:

$$h_m(x) = \sum_{i=1}^{N} \big(g_{im} - h(x_i)\big)^2$$

This ensures that each new weak learner focuses on correcting the mistakes of the previous iterations. After training $h_m(x)$, the optimal step size γ_m is determined by solving:

$$\gamma_m = \sum_{i=1}^{N} L\left(y_i, F_{m-1}(x_i) + \gamma h_m(x_i)\right)$$

For MSE loss, the optimal value of γ_m can be computed as:

$$\gamma_m = \frac{\sum_{i=1}^{N} g_{im}\, h_m(x_i)}{\sum_{i=1}^{N} h_m^2\,(x_i)}$$

After M boosting iterations, the final model is:

$$F_M(x) = F_0(x) + \sum_{m=1}^{M} \gamma_m \, h_m(x)$$

where each weak learner's contribution is weighted by γ_m.

For time series forecasting, this method is extended to predict future values beyond T. Two common strategies are the *Recursive Strategy*

$$\hat{y}_{T+1} = F_M(x_T)$$
$$\hat{y}_{T+2} = F_M(\hat{x}_{T+1})$$

where $\hat{x}_{T+1}$ includes the predicted value $\hat{y}_{T+1}$ as an input, and the *Direct Strategy* that trains separate models for each forecast horizon h:

$$\hat{y}_{T+h} = F_M^{(h)}(x_T)$$

where $F_M^{(h)}(x)$ is a separately trained gradient boosting model for each h.

To incorporate covariates into Gradient Boosting for TS Forecasting, we now introduce $z_t \in \mathbb{R}^{d_z}$ as the vector of exogenous covariates at time t:

$$\{(x_t, z_t, y_t)\}_{t=1}^{T}$$

The input to the model at time t now consists of the feature vector: $u_t = [x_t, z_t] \in \mathbb{R}^d$. Thus, the model at each iteration is updated as:

$$F_m(u) = F_{m-1}(u) + \gamma_m h_m(u)$$

where the weak learner $h_m(u)$ is now trained on both lagged values and covariates. The loss function remains the same but now depends on both x and z:

$$L\big(y, F(x,z)\big) = \frac{1}{N}\sum_{i=1}^{N}\big(y_i - F(x_i, z_i)\big)^2$$

and the gradients are computed as:

$$g_{im} = -\left[\frac{\partial L\big(y_i, F(x_i, z_i)\big)}{\partial F(x_i, z_i)}\right]_{F(x,z)=F_{m-1}(x,z)}$$

For MSE loss:

$$g_{im} = y_i - F_{m-1}(x_i, z_i)$$

Each weak learner is trained to fit the residuals:

$$h_m(x,z) = \sum_{i=1}^{N}\big(g_{im} - h(x_i, z_i)\big)^2$$

which ensures that the model learns both temporal dependencies from x and external influences from z.

3 Case Study: Chicago

In this section we present the result after implementing the different models to all zip codes in the study

area on both the bus and rail modes. Each model is applied in two scenarios: using only the past values of boardings to forecast future boardings and using electricity usage as covariates to forecast future boardings. Figure 4 shows the cumulative result of applying the modeling approach for the entire study area. This means that a separate LGBM model is estimated for every single zip code for which there is bus service, and then the time series are added for both the ground truth and the predicted time series.

We can see the same similarly found in the single zip code case. The prediction without covariate fails at capturing the overall difference between workday and weekday in terms of number of demand peaks and also maximum daily demand. When covariates are introduced, this problem diminishes as we can see a variation from weekend to workdays and a maximum daily demand closer to the actual test (ground truth) values.

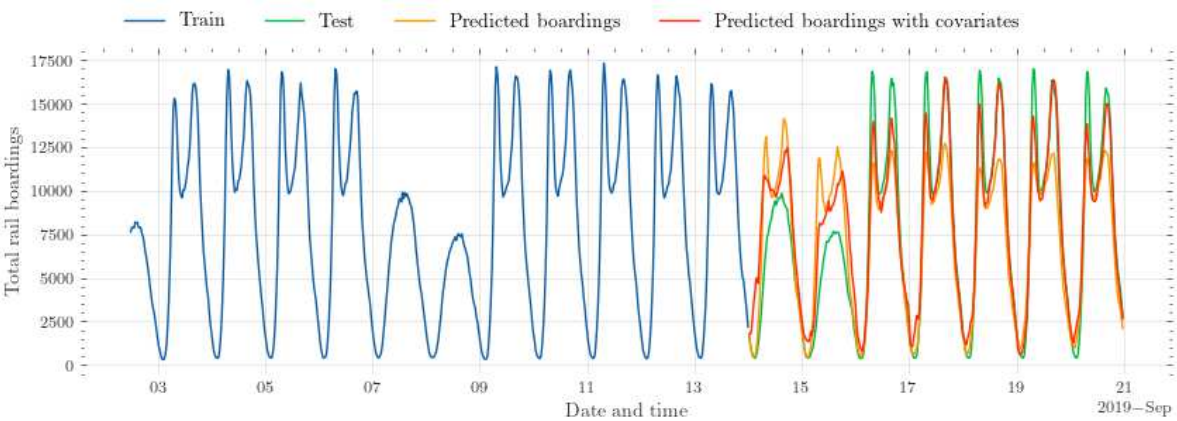


Figure 4. Predicted total bus demand with and without covariates for the entire study area.

Similarly, Figure 5 displays the equivalent results for the rail demand. In this case, the presence of atypical demand on two workdays presents an additional challenge to the model estimation, as it requires calibration including two out of the ordinary demand peaks. Even in that case, we see similar results to the ones of bus demand, where predictions using covariates outperform those made without covariates. However, there seems to be a sub-optimal representation of the demand peaks in terms of their magnitude. In order to quantify the actual zip code-based accuracy of the methods, we introduce Figure 6, where the MAPE for each zip code was calculated for both the covariates and no-covariates cases. One of the general observations we can draw from these plots is that the bus demand can be estimated with higher accuracy under this methodology for both types of implementations. The error values have a higher variance in the rail case, whereas only a few zip codes in the bus case present an outlying high error. In both cases, however, the introduction of covariates improves the demand estimation, sometimes by several orders of magnitude, but there are also a few exceptions where even the introduction of covariates leads to poor results.

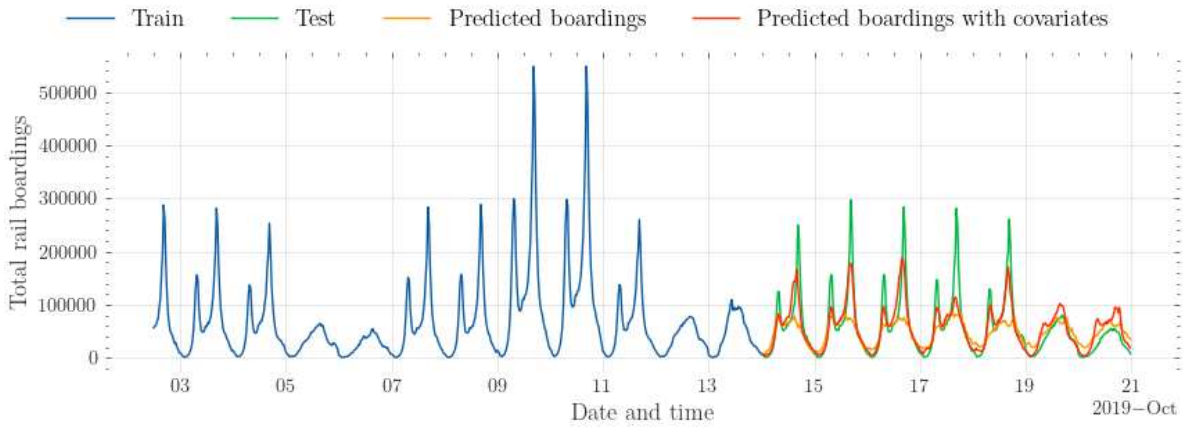


Figure 5. Predicted total rail demand with and without covariates for the entire study area.

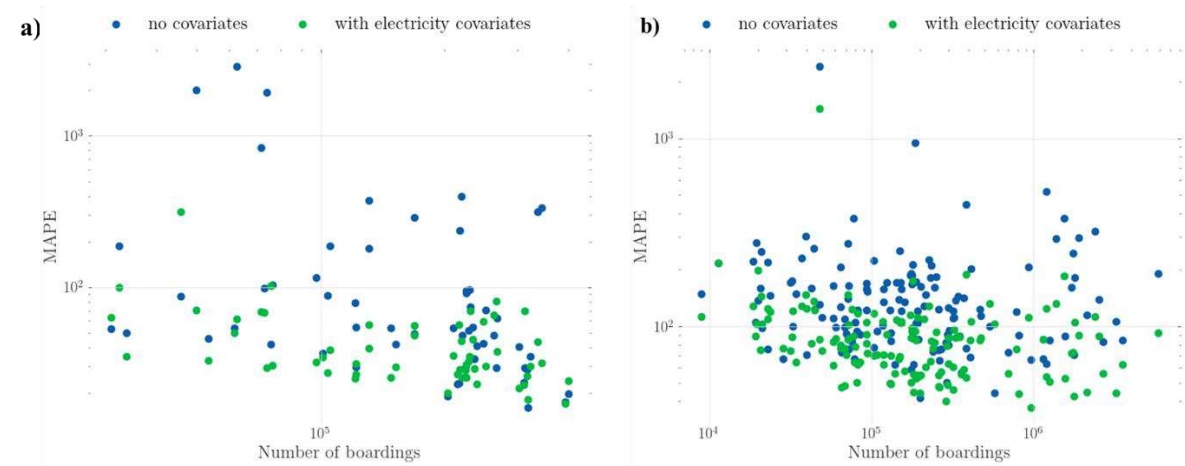


Figure 6. MAPE for each a) zip code and b) rail station obtained from the approach with and without electricity covariates as a function of the demand. Each point represents a zip code or a rail station, respectively.

To obtain spatial insights into the performance of these methods, we mapped the MAPE for every single zip code used in the bus demand (Figure 7a-b) and every rail station (Figure 7Figure 6c-d). For the bus demand, we can see that the error decreases in virtually every zip code after introducing the covariates, reaching acceptable values of percent error (below 50%) in most zip codes. The zip codes with higher error values are mostly concentrated in the downtown and west parts of the city. This indicates that these regions present a lower correlation between electricity and bus demand. The rail demand results display and overall lower performance, despite some improvement from the introduction of covariates. As opposed to bus demand, rail presents better results in the downtown stations and northern stations, which suggests the existence of underlying conditions that tie rail demand and electricity usage for these areas.

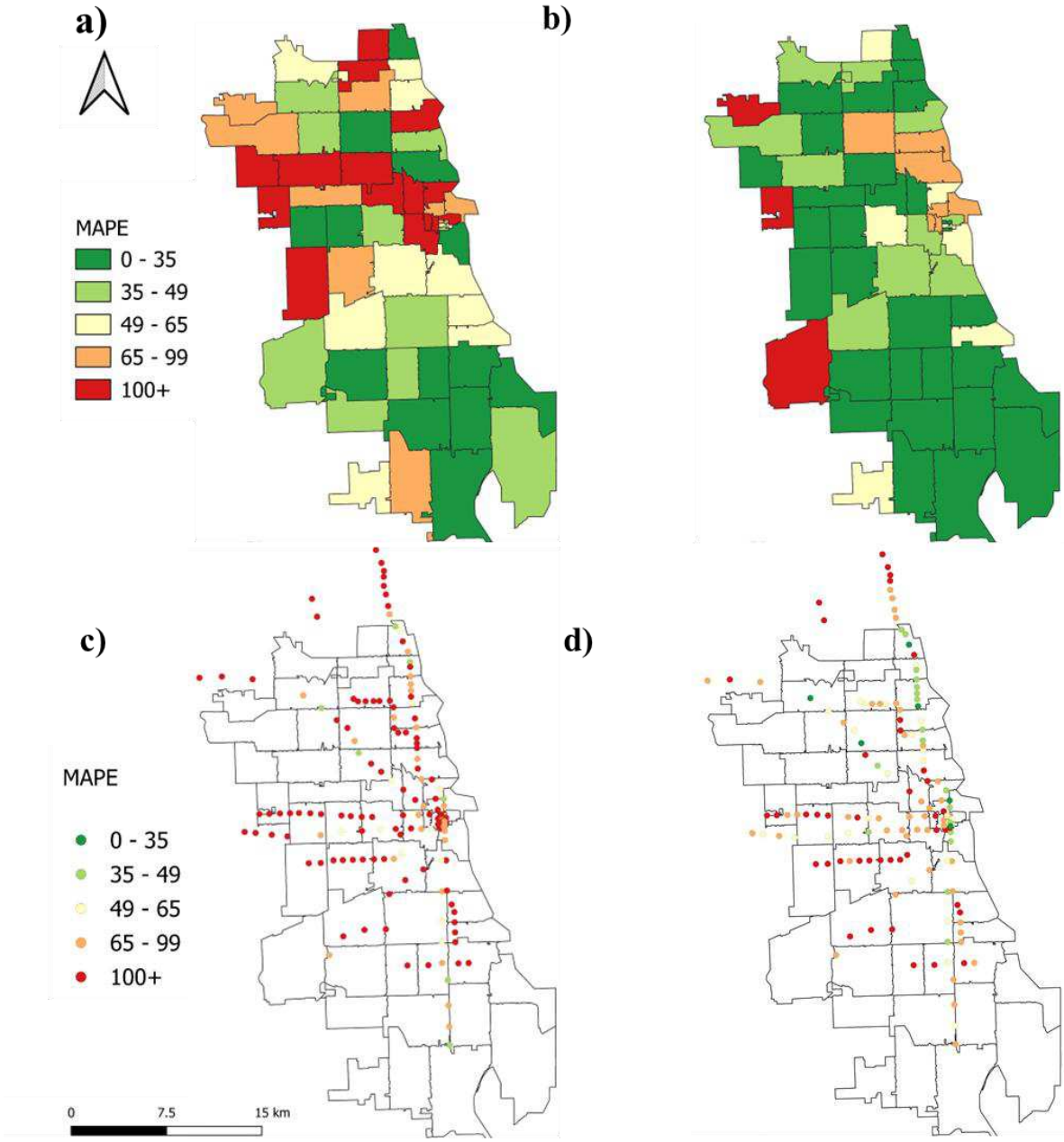


Figure 7. Geographical distribution of zip code MAPE for the bus demand (a) without and b) with covariates; and station-based rail demand (c) without and (d) with covariates.

3.1 Application case: a shock to the system

So far, we have shown how transit demand estimations improve by incorporating electricity demand covariates. In most cases, the improvement is important, but it could be argued that relying solely on past observations and the seasonal nature of transit demand is already good enough. Here, we want to demonstrate an instance where this methodology is especially useful: when the typical demand of the city is disturbed by a

major event. Major events that can disrupt the expected behavior of urban systems include natural disasters, events of social unrest, security threats, and important public health measures, such as pandemics. Incorporating transit and electricity usage data of Chicago during the start of the COVID-19 pandemic in the city, we are going to showcase how this methodology can be applied in such circumstances.

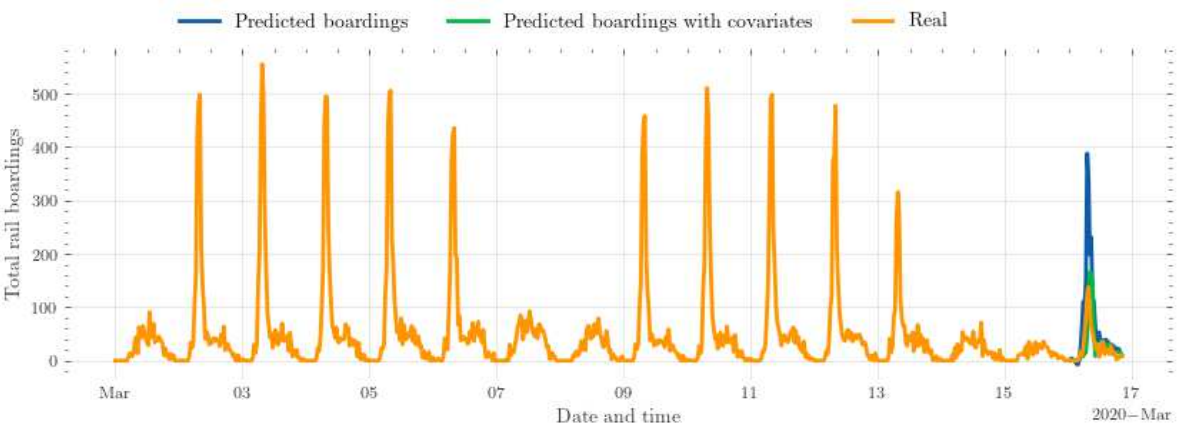


Figure 8. Predicted and real rail boardings for a selected station for March 2020, during the COVID-19 pandemic.

Figure 8 shows real rail boardings for a station in Chicago during March of 2020, a period of time that contains the official beginning of the COVID-19 pandemic and the start of the city-wide measures implemented to reduce the virus spread. The first couple weeks of March, we see a typical demand pattern consistent with weekend and weekday patterns with pronounced peaks on weekdays. However, on Monday March 16th, the demand changed from in the past two Mondays. The work from home (WFH) measures were in place after the World Health Organization (WHO) officially declared the global pandemic a few days prior. This led to a sharp decline in the rail demand, which based solely on previous demand data is almost impossible to accurately predict, as shown in the *predicted* time series values corresponding to that Monday's demand. However, if we introduce electricity data as time series covariates, we can see an immediate result in the predicted demand that closely resembles the actual atypical demand found on that day. Table 2 shows the overall metrics after accounting for the total demand of all rail stations in the city. We can see a clear improvement in both the absolute and relative metrics on the modeling approach with covariates.

Table 2. City-wide results including all rail stations when predicting the day-ahead demand after the COVID-19 outbreak impacted transit ridership.

Model	**RMSE**	**MAPE**	**CV**
No covariates	14867.49	51.20	62.88
With covariates	9216.85	27.04	32.26

4 DISCUSSION

The results suggests that the electricity consumption as covariates can help better predict transit demand. Let's recall our main hypothesis: *changes in electricity usage in a given zone in a city are indicators of human activity that can be used to forecast local transportation demand.* That hypothesis, in combination with the intuition that the way travel and electricity are related to each other has the form of a lagged relationship, might explain, at least in part, the underlying mechanism that makes this modeling exercise work. Here is

worth mentioning that with *lagged* we mean that current electricity readings are correlated with future transit ridership, given that under our hypothesis, the underlying relationship is dictated by people leaving their buildings and then taking transit. If a person that causes a change in electricity usage, then engages in a trip, there is already a lagged nature in the sense of the intrinsic time difference between the two activities. Hence, the cumulative effect of electricity demand changes might as well be perceived as a lagged cumulative cause in terms of trips. Moreover, the fact that electricity usage offers an indication of potential trips means that when the model uses past and current electricity data as input is already capturing trips time series in the form of its covariate.

We tested this modeling approach making use of the COVID-19 case, where the system is disrupted by an atypical phenomenon and solely past transit demand observations have no means of accounting for this sort of effect in advance. Hence, another variable, such as electricity usage, can serve as the proxy for several reasons: 1) widely-adopted smart-metering infrastructure makes it possible to measure in real-time the city response to such events; 2) electricity usage displays a high interrelation with urban travel based in this and past studies; and 3) electricity data is not only a proxy for transportation, but also is lagged in the sense that the changes observed in electric usage that can affect transit happen in advance, which combined with the real-time nature of metering systems has a huge potential of becoming an important input for intelligent transport systems frameworks and overall smart-city infrastructure.

From Figure 6, the extent to which our assumption holds seems to be determined at each zone and station by some underlying mechanism that ties these phenomena at different levels of correlation. In addition, the lower errors as the magnitude of demand increases in Figure 6 suggests that a higher activity in this mode of transportation is an indication of a higher interrelation to electricity usage, whereas zones with lower transit activity indicate that other modes of transportation better explain the changes in electricity usage explained by urban travel. For instance, Movahedi *et al.* (Movahedi *et al.*, 2023) explores the interrelationship between traffic counts and electricity consumption. This points to one of the possible next steps in this line of research: in order to fully cover the extent to which urban travel explains electricity patterns and vice versa, we need to incorporate other modes of transportation, including WFH decisions that vastly affect both ridership and electricity usage. Another critical aspect that this work does not consider but the authors recognize as crucial is that here, we are only considering residential electricity usage, but non-residential electricity usage in combination with incorporating land use distribution should refine the results we observed.

One final discussion remark revolves around the relevance of an approach of this nature. Despite the positive results achieved with a relatively simple approach, it does seem that a phenomenon such as transit demand is so periodic in nature that just assuming the repetition of workday/weekend patterns is enough to obtain acceptable estimates with a lot less effort. Even when this could be correct in a lot of instances, we consider that there lies precisely the value of this work in obtaining a look ahead of time at what happens when urban systems operate out of the ordinary. We can think of it in terms of protests, road closures, pandemics, or major weather events: if by means of other closely monitored urban infrastructure systems, we can obtain at least a 5-hour ahead clue of what to expect, the city response to unusual urban dynamics could benefit from an additional layer of estimations.

5 CONCLUSION

This study shows that integrating electricity consumption data with time series forecasting techniques can significantly improve the accuracy of transportation demand modeling, particularly in urban environments. We validated our hypothesis that changes in electricity usage can serve as an indicator of human activity, and by extension, forecast local transportation demand. The inclusion of electricity usage data as a covariate notably improved the model's ability to predict typical demand patterns, particularly the weekday-weekend variations, and addressed the time shift between electricity consumption and transportation demand peaks. The case study of Chicago not only provided initial evidence into the interdependence of electricity and transit demand but also illustrated the predictive power of this approach under atypical conditions, such as the onset of the COVID-19 pandemic. By incorporating electricity usage as a real-time, lagged proxy for travel behavior, the model was able to better account for disruptions in normal transit patterns that would be difficult

to predict using traditional methods based solely on historical transit data. Furthermore, this work highlights the potential of smart city infrastructure, particularly smart meters, in creating more responsive and resilient urban systems. The observed variation in model performance across different zip codes and transit stations highlights the importance of considering local context and the dynamic nature of urban networks. Future research could expand on this framework by incorporating additional urban systems, such as non-residential electricity consumption and other modes of transportation, to further refine demand forecasts. Additionally, integrating land use patterns and work-from-home behavior could offer a more comprehensive understanding of the mechanisms behind this interrelationship.

References

Acheampong, R.A. and Silva, E. (2015) 'Land use–transport interaction modeling: A review of the literature and future research directions', *Journal of Transport and Land Use*, 8(3). Available at: https://doi.org/10.5198/jtlu.2015.806.

Acosta-Sequeda, J. *et al.* (2023) 'Residential Electricity Consumption Patterns and their Relationship to Commute Times by Mode', *Findings* [Preprint]. Available at: https://doi.org/10.32866/001c.87940.

Acosta-Sequeda, J. (2024) 'Estimating Telecommuting Rates in the USA Using Twitter Sentiment Analysis'.

Alonso, W. (2013) 'Location and Land Use: Toward a General Theory of Land Rent', in *Location and Land Use*. Harvard University Press. Available at: https://doi.org/10.4159/harvard.9780674730854.

Amini, M.H. *et al.* (2018) 'A Panorama of Future Interdependent Networks: From Intelligent Infrastructures to Smart Cities', in M.H. Amini et al. (eds) *Sustainable Interdependent Networks: From Theory to Application*. Cham: Springer International Publishing, pp. 1–10. Available at: https://doi.org/10.1007/978-3-319-74412-4_1.

Batty, M. (2013) 'Big data, smart cities and city planning', *Dialogues in Human Geography*, 3(3), pp. 274–279. Available at: https://doi.org/10.1177/2043820613513390.

Beatley, T. (2011) *Biophilic Cities: Integrating Nature Into Urban Design and Planning*. Island Press.

Berger, S. *et al.* (2022) 'What motivates smart meter adoption? Evidence from an experimental advertising campaign in Germany', *Energy Research & Social Science*, 85, p. 102357. Available at: https://doi.org/10.1016/j.erss.2021.102357.

Cervero, R. and Kockelman, K. (1997) 'Travel demand and the 3Ds: Density, diversity, and design', *Transportation Research Part D: Transport and Environment*, 2(3), pp. 199–219. Available at: https://doi.org/10.1016/S1361-9209(97)00009-6.

Chauhan, R.S. *et al.* (2025) 'Causation versus prediction in travel mode choice modeling', *npj Sustainable Mobility and Transport*, 2(1), pp. 1–11. Available at: https://doi.org/10.1038/s44333-024-00022-4.

Chen, Y.-J., Matsuoka, R.H. and Liang, T.-M. (2018) 'Urban form, building characteristics, and residential electricity consumption: A case study in Tainan City', *Environment and Planning B: Urban Analytics and City Science*, 45(5), pp. 933–952. Available at: https://doi.org/10.1177/2399808317690150.

Cirillo, C. and Axhausen, K.W. (2001) 'Comparing urban activity travel behaviour', in *Arbeitsberichte Verkehrs- und Raumplanung. Annual Meeting of the Transportation Research Board*, IVT, ETH Zurich. Available at: https://doi.org/10.3929/ethz-a-004339644.

Davoudi, S. *et al.* (2012) 'Resilience: A Bridging Concept or a Dead End? "Reframing" Resilience: Challenges for Planning Theory and Practice Interacting Traps: Resilience Assessment of a Pasture Management System in Northern Afghanistan Urban Resilience: What Does it Mean in Planning Practice? Resilience as a Useful Concept for Climate Change Adaptation? The Politics of Resilience for Planning: A Cautionary Note: Edited by Simin Davoudi and Libby Porter', *Planning Theory & Practice*, 13(2), pp. 299–333. Available at: https://doi.org/10.1080/14649357.2012.677124.

Derrible, S. (2018) 'An approach to designing sustainable urban infrastructure', *MRS Energy & Sustainability*, 5, p. E15. Available at:

https://doi.org/10.1557/mre.2018.14.

Derrible, S. (2019) *Urban Engineering for Sustainability*. The MIT Press.

Derrible, S. (2025) *The Infrastructure Book*. 1st edn. Prometheus Books. Available at: https://www.prometheusbooks.com/9781493086641/the-infrastructure-book/ (Accessed: 26 February 2025).

Herzen, J. *et al.* (2022) 'Darts: User-Friendly Modern Machine Learning for Time Series', *Journal of Machine Learning Research*, 23(124), pp. 1–6.

Hunt, L.C. and Ninomiya, Y. (2003) 'Unravelling Trends and Seasonality: A Structural Time Series Analysis of Transport Oil Demand in the UK and Japan', *The Energy Journal*, 24(3), pp. 63–96. Available at: https://doi.org/10.5547/ISSN0195-6574-EJ-Vol24-No3-3.

Lee, J.-M. and Gwon, Y.-J. (2011) 'A Study on Dynamic Change of Transportation Demand Using Seasonal ARIMA Model', *Journal of Korean Society of Transportation*, 29(5), pp. 139–155.

Lee, Y.-J. and Vuchic, V.R. (2005) 'Transit Network Design with Variable Demand', *Journal of Transportation Engineering*, 131(1), pp. 1–10. Available at: https://doi.org/10.1061/(ASCE)0733-947X(2005)131:1(1).

Liashchynskyi, Petro and Liashchynskyi, Pavlo (2019) 'Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS'. arXiv. Available at: https://doi.org/10.48550/arXiv.1912.06059.

Lim, J. and Kang, M. (2022) 'The relationship between site planning and electricity consumption: An empirical analysis of multi-unit residential complexes in Seoul, Korea', *Environment and Planning B: Urban Analytics and City Science*, 49(3), pp. 971–986. Available at: https://doi.org/10.1177/23998083211039855.

Ma, X. *et al.* (2018) 'A geographically and temporally weighted regression model to explore the spatiotemporal influence of built environment on transit ridership', *Computers, Environment and Urban Systems*, 70, pp. 113–124. Available at: https://doi.org/10.1016/j.compenvurbsys.2018.03.001.

Mills, E.S. (1967) 'An Aggregative Model of Resource Allocation in a Metropolitan Area', *The American Economic Review*, 57(2), pp. 197–210.

Movahedi, A. *et al.* (2023) 'Interrelationships between urban travel demand and electricity consumption: a deep learning approach', *Scientific Reports*, 13(1), p. 6223. Available at: https://doi.org/10.1038/s41598-023-33133-y.

Newman, P. and Kenworthy, J. (1999) 'Costs of Automobile Dependence: Global Survey of Cities', *Transportation Research Record*, 1670(1), pp. 17–26. Available at: https://doi.org/10.3141/1670-04.

Putnam, R.D. (2000) *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster.

Seto, K.C., Güneralp, B. and Hutyra, L.R. (2012) 'Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools', *Proceedings of the National Academy of Sciences*, 109(40), pp. 16083–16088. Available at: https://doi.org/10.1073/pnas.1211658109.

Tureczek, A.M. and Nielsen, P.S. (2017) 'Structured Literature Review of Electricity Consumption Classification Using Smart Meter Data', *Energies*, 10(5), p. 584. Available at: https://doi.org/10.3390/en10050584.

Van Soest, D., Tight, M.R. and Rogers, C.D.F. (2020) 'Exploring the distances people walk to access public transport', *Transport Reviews*, 40(2), pp. 160–182. Available at: https://doi.org/10.1080/01441647.2019.1575491.

Wang, Y. *et al.* (2022) 'The impact of traffic demand management policy mix on commuter travel choices', *Transport Policy*, 117, pp. 74–87. Available at: https://doi.org/10.1016/j.tranpol.2022.01.002.

Xiong, J. *et al.* (2024) 'Identifying, Analyzing, and forecasting commuting patterns in urban public Transportation: A review', *Expert Systems with Applications*, 249, p. 123646. Available at: https://doi.org/10.1016/j.eswa.2024.123646.

Zhong, C. *et al.* (2023) 'Online prediction of network-level public transport demand based on principle component analysis',

Communications in Transportation Research, 3, p. 100093. Available at: https://doi.org/10.1016/j.commtr.2023.100093.

6 Supplementary Materials

6.1.1 Grid-Search space and results

Table 3. Grid Search space attributes explored and the best performing hyperparameters obtained for each model, mode, and covariates implementation.

Mode	Model	Hyperparameters	Search Values	Best Value (no Covariates)	Best Value with Covariates
Bus	LSTM	hidden_dim	5,10,12,14,15,17,18,20,21,22,23,24,25	20	20
		input_chunk_length	6,8,10,12,14,15,20,25,30,35,40,45,50,55,70,100	14	14
		training_length	10,20,30,40,50,70,80,100	20	30
	KFF	x_dim	10,20,30,40,50,60,70,80,90,100,110,120,130	130	50
	LGBM	lags	6,8,10,12,15,20,24,30,48	20	19
		lags_past_covariates	6,8,10,12,15,20,24,30,48	NA	25
		output_chunk_length	6,8,10,12,15,20	8	9
		n_estimators	10,20,30,40,50,100	100	50
Rail	LSTM	hidden_dim	5,10,12,14,15,17,18,20,21,22,23,24,25	20	20
		input_chunk_length	6,8,10,12,14,15,20,25,30,35,40,45,50,55,70,100	14	14
		training_length	10,20,30,40,50,70,80,100	30	30
	KFF	x_dim	10,20,30,40,50,60,70,80,90,100,110,120,130	130	20
	LGBM	lags	6,8,10,12,15,20,24,30,48	24	12
		lags_past_covariates	6,8,10,12,15,20,24,30,48	NA	24
		output_chunk_length	6,8,10,12,15,20	12	15
		n_estimators	10,20,30,40,50,100	30	100