Journal of Cleaner Production 434 (2024) 139895

Contents lists available at ScienceDirect

Journal of Cleaner Production

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

Using neural networks to forecast marginal emissions factors: A CAISO case study

Stepp Mayes [a], Nicholas Klein [b], Kelly T Sanders [a,*]

[a] *University of Southern California, Department of Civil and Environmental Engineering, Los Angeles, 90089, CA, USA*

[b] *University of Southern California, Department of Computer Science, Los Angeles, 90089, CA, USA*

ARTICLE INFO

Handling Editor: Xin Tong

Keywords:
Marginal Emissions Factors
Demand-side management
Variable renewable energy
Load-shifting
Machine learning
Electric grid

ABSTRACT

Marginal Emissions Factors (MEFs) quantify the time-dependent changes in CO_2 emissions resulting from changes in electricity consumption. Accurate MEFs are critical for calculating the emissions impact of demand-side management (DSM) activities and programs, but current methods of calculating MEFs are limited by their temporal resolution, accuracy (particularly in grids with high penetrations of variable renewable energy), and ability to predict MEFs ahead of time, reducing their utility for DSM. We improve upon existing techniques by introducing a novel multi-layer perceptron to linear composite model that uses publicly available grid data to calculate historical MEFs and predict day-ahead MEFs. We test our model on publicly-available data from the California Independent System Operator over the period of 2019–2021, a grid with high daytime VRE generation. Results indicate that our model produces more accurate and more granular demand-based MEF estimations than comparable regression techniques and maintains high accuracy when use to forecast future MEFs. Our MEF framework can be applied to other regional grids to evaluate and design DSM strategies that leverage CO_2 emissions-reductions as motivation for altering electricity consuming behaviors.

1. Introduction

As electric grids in the United States achieve increasingly high fractions of wind and solar generation, the relationship between amount of electricity demanded and the CO_2 emissions associated with producing that electricity is decoupling and highly time-dependent. While grids with primarily fossil-fuel-based generation typically exhibit a close relationship between the amount of electricity demanded and the associated emissions, grids with high fractions of wind and solar Photovoltaic power (referred to collectively as variable renewable energy sources, or VRE) have an emissions-intensity that varies significantly throughout the day and over the course of a year. For example, wind and solar power accounted for 24.9% of total generation in 2021 in California (Commission, 2022), and as a result the California Independent System Operator (CAISO, the organization responsible for managing the majority of California's bulk power system Operator, 2023c) experiences diurnal patterns of solar generation and total demand for electricity that often create a low net load (total load minus VRE generation) in the middle of the day and a high net load in the evening. As a result, the grid tends to be less emissions-intense in the middle of the day when there is an abundance of solar power and more emissions-intense in the afternoon and evening when CAISO relies heavily on natural gas generators to meet demand (Operator, 2023b; Denholm et al., 2015). An implication of this phenomenon is that shifting an electricity load from one time period to another can reduce the amount of emissions associated with the load, even if the magnitude of demand is unaffected. Quantifying the exact change in emissions associated with modifying load at a specific time becomes a challenging but important task.

One tool used to quantify the changes in emissions associated with changes in demand are emissions factors (EFs). Researchers typically draw a distinction between Marginal Emissions Factors (MEFs) and Average Emissions Factors (AEFs) (Ryan et al., 2016; Hawkes, 2014). AEFs describe the relationship between a region of interest's total generation for a period of time and the total emissions produced from that generation, which is dependent on all of the generators supplying electricity to that region's grid. MEFs describe the relationship between changes in generation (or demand) and changes in emissions and depend only on the resource or resources that respond to changes in demand at a specific time, making them useful for evaluating the impact of adding, removing, or altering loads (Regett et al., 2018; Samaras and Meisterling, 2008).

Several studies have used AEFs for evaluating the emissions changes caused by demand-side management (DSM) strategies that change the timing of electricity loads, including air-conditioning usage (Mayes

* Corresponding author.
E-mail address: ktsanders@usc.edu (K.T. Sanders).

https://doi.org/10.1016/j.jclepro.2023.139895
Received 10 August 2023; Received in revised form 12 October 2023; Accepted 22 November 2023
Available online 23 November 2023

and Sanders, 2022), EV charging (Noussan and Neirotti, 2020), and electricity usage in the water industry (Zohrabian and Sanders, 2021). However, using MEFs for DSM applications has become more common in recent years under the justification that additional loads or changes in load impact electricity production by marginal generators (as opposed to changing production across the whole generation fleet). Recent studies have used MEFs to evaluate the impact of EV charging patterns (Gai et al., 2019; Holland et al., 2022; Huber et al., 2021; Kamiya et al., 2019), electricity storage policy (Li et al., 2017; Pimm et al., 2021; McKenna et al., 2017; Braeuer et al., 2020), air-conditioning timing (Stopps and Touchie, 2022), and societal damage factors (Donti et al., 2019).

One common method for estimating MEFs is through grid modeling. Grid models work by representing a fleet of power plants and creating an order in which they respond to demand, with the order often being cost- or merit-based and less frequently dependent on factors such as the location of the demand and/or the individual power plants (Hawkes, 2014; Deetjen and Azevedo, 2019; Gagnon and Cole, 2022; Sengupta et al., 2022; Zheng et al., 2015). A downside to the grid-modeling approach is that these models often require simplifications that prevent them from fully representing grid behavior. For example, Gagnon and Cole (2022) point out that their model does not account for maximum ramp rates, or minimum up or down times, which can significantly impact MEF estimates (Zheng et al., 2015). Other models ignore trades of electricity between regions, which can be a significant portion of supply and demand (Deetjen and Azevedo, 2019). The accuracy of these models can also vary significantly depending on location and data availability (Sengupta et al., 2022).

Alternatively, regression techniques that estimate MEFs based on historical grid data may do a better job of incorporating grid constraints, although they are also limited by data availability and vary in accuracy. Many regression-based analyses follow a methodology similar to that established by Hawkes (2010), where changes in demand or generation are regressed on changes in emissions (single-factor linear regression) (Holland et al., 2022; Huber et al., 2021; Kamiya et al., 2019; Li et al., 2017; Pimm et al., 2021; McKenna et al., 2017; Braeuer et al., 2020; Donti et al., 2019; Seckinger and Radgen, 2021; Siler-Evans et al., 2012; Thind et al., 2017). While this methodology works well for primarily fossil-fuel-based grids, it is less appropriate for grids that have both carbon and non-carbon emitting resources changing generation throughout the day. For example, Siler-Evans et al. (2012) regress changes in emissions on changes in fossil-fuel generation to determine the MEF, but this assumes that renewable energy sources are never the marginal resource, which is not the case for grids with significant generation of hydropower or that experience VRE curtailment (Zohrabian et al., 2023; Operator, 2023a). Li et al. (2017) expand on this approach by regressing hourly changes in emissions on hourly changes in generation from all sources. While this better incorporates renewables as marginal generators, it fails to capture grid dynamics during VRE ramping periods. Under this methodology, rapid increases in VRE generation can artificially give the impression that additional demand is emissions-free, or even create situations where hourly emissions drop despite hourly demand increasing (due to the displacement of fossil fuel generators). This could give the false conclusion that the MEF is zero or even negative, when, in the absence of curtailment, additional demand would actually increase emissions. This failure to identify the true marginal emissions could lead to large inaccuracies when estimating the emissions impacts of potential DSM strategies by incorrectly encouraging shifting load to high-renewable hours when, in reality, the marginal resource for some of those hours would be fossil-fuels.

To improve upon these single-factor regression methodologies, more recent studies have introduced models that incorporate both changes in total demand and changes in generation from VRE sources. These two variables can be thought of as having separate and opposite effects on changes in emissions and can be represented through a multiple-linear regression model, which was first introduced by Thomson et al. (2017) to capture the influence of wind generation in Great Britain. This approach was also utilized by Zohrabian et al. (2023) for VRE generation in California, where the coefficient on the VRE term was called the marginal displacement factor (MDF). These models better isolate the impact of changes in demand (instead of changes in generation or a subset of generation) on emissions, and thus are more appropriate for DSM applications that seek to determine the impact that shifting, shedding, or adding loads has on emissions. While this approach is an improvement on single-factor regression techniques, the application of the resulting MEFs to DSM is still limited by the temporal properties of these models.

Beyond being demand-based, MEFs that are intended for evaluating DSM strategies like demand response and load-shifting should also be highly temporally resolved and capable of being estimated in advance. (We refer to these qualities as “granularity” and “forecastability”, respectively, in this manuscript.) Accurate, ahead-of-time MEF estimations would enable the creation of data-driven DSM programs that encourage behavioral changes to reduce emissions. Traditional linear regression calculates the model coefficients for groups of data points, which results in assigning the same MEF to multiple points. This lack of granularity is a concern because MEFs can vary significantly hour-to-hour and day-to-day. Coarse-resolution MEFs, such as those calculated at the annual level (Donti et al., 2019; Holland et al., 2022; Seckinger and Radgen, 2021; Thomson et al., 2017; Siler-Evans et al., 2012; Thind et al., 2017), can be useful for understanding the evolution of the grid and high-level changes in the mix of marginal resources, but are poorly-suited for evaluating the emissions implications of DSM applications, which modify loads on sub-daily time intervals. Other studies group their analyses by hour of the day, calculating MEFs at the year-hour level (24 values for the whole year) (Thomson et al., 2017; Kamiya et al., 2019; Siler-Evans et al., 2012; Thind et al., 2017; Braeuer et al., 2020), or month-hour level (288 values for the whole year) (Gai et al., 2019; Li et al., 2017; Donti et al., 2019; Zohrabian et al., 2023). MEFs calculated with these methods are better suited for demand-side applications that focus on time of day, but fail to capture the significant changes in grid and consumer behavior that can occur across months, or even days. Factors such as the demand level and amount of renewable energy generation can depend strongly on weather variables that change significantly day-to-day.

To increase granularity, Beltrami et al. (2020) developed a novel methodology using an auto-regressive integrated moving average model and found that it compared favorably to traditional linear-regression approaches, achieving higher granularity and more accurate predictions of changes in emissions for historical data, though they did not explore using this method for forecasting. Other analyses group their regressions by load-level (Pimm et al., 2021; Siler-Evans et al., 2012; Thind et al., 2017; McKenna et al., 2017; Huber et al., 2021), which could theoretically improve both granularity (i.e., any time at which load is known can be assigned an MEF) and allow for MEF forecasting by using projections of load that are commonly done for electricity providers. Of the listed studies, this forecasting method was explored only by Huber et al. (2021) who used MEF forecasts to determine optimal EV charging times. The main limitation of this approach is that binning by demand prior to regression forces all points with a similar level of demand to have the same MEF; in reality, other factors such as the time of day and amount of renewable generation can strongly impact MEFs.

New methodologies are needed for calculating demand-based MEFs given the growing importance of DSM strategies for managing the challenges associated with grids with high penetrations of variable renewable energy and the rapid adoption of new grid resources and grid management strategies. The proposed method must be effective at (1) isolating the impact of demand on emissions, (2) temporally resolving the emissions factors (“granularity”), and (3) predicting these emissions factors ahead of time (“forecastability”).

We improve upon previous models by creating a composite model composed of a multi-layer perceptron (MLP) and a linear model and then apply it to the grid overseen by CAISO, a grid that has high fractions of renewable generation, a variety of resources that contribute to marginal generation, and good data availability. While (to the authors' knowledge) MLP models have not previously been used to calculate MEFs, there is precedent for using neural networks in the electric grid research space to forecast electricity demand and prices that achieve high degrees of accuracy and granularity (Park et al., 1991; Singhal and Swarup, 2011; Catalão et al., 2007; Chae et al., 2016). The advantages offered by the composite model developed in this study, including higher accuracy and granularity than traditional regression models, are maintained through model forecasting, resulting in MEFs that are well suited for DSM applications.

2. Material and methods

2.1. Preparation of dataset

We prepared a dataset of CAISO generation, demand, and emissions for the 2019–2021 period with the following data collection and manipulation process. The key data used in this methodology and their sources include:

1. data on CAISO's demand, variable renewable energy generation, and total imports from CAISO's website (Operator, 2023b).
2. hourly emissions data for individual power plants from the U.S. Environmental Protection Agency's Clean Air Markets Program Data (Agency, 2023) (it should be noted that the EPA Air Markets Program Data is only for power plants with capacity greater than 25 MW, so the emissions calculated in this study do not include fossil fuel generators under this size)
3. data identifying which power plants belong to which balancing authorities from US Energy Information Administration (EIA) 860 data (U.S. Energy Information Administration (EIA), 2019)
4. hourly data for individual BA generation by source and hourly trades between BAs from the EIA (trades are reported bidirectionally) (EIA, 2022).

For all of the above sources, the data used are publicly available and free to download. These datasets required several processing steps to fill in missing values and ensure temporal agreement between sources. First all data from time zones other than PST were shifted to the PST time zone. Second, the CAISO load data (reported in MW) reported at 5-min intervals were aggregated to the hourly demand level (in MWh) via averaging. Third, for hours where CAISO's electricity trade with another BA was not reported in the EIA data, we first checked to see if these values were present in the dataset from the other BA's perspective. If values were missing in both directions we filled in the missing values with the average of the values on the 5 closest days at the same hour of the day. Then, the sum of all hourly electricity trades between CAISO and other regions reported in the EIA exchange data was scaled to match CAISO's hourly total reported net imports (essentially treating CAISO's total reported value as the ground truth). This scaling was done by taking the difference between CAISO's reported value and the summed value and distributing this difference to each BA in proportion to the magnitude of their reported electricity trade. Finally EIA 860 data was used to aggregate the hourly powerplant emissions data to the balancing authority level, creating an estimate of total hourly emissions associated with CAISO generation and the total hourly emissions associated with the generation of each balancing authority with whom CAISO trades electricity.

These datasets were then combined to create a final dataset describing the hourly demand, VRE, and CO_2 emissions in the CAISO region (specifically, the hourly emissions are those associated with demand that occurs within the CAISO region). Calculating these hourly, demand-based emissions for CAISO requires accounting for trades of electricity between CAISO and neighboring BAs and the emissions associated with these trades. This emissions accounting process generally follows the procedure outlined by de Chalendar et al. (2019), which was also utilized by Zohrabian et al. (2023) from which we adapt Eq. (1).

$$E_h = E_h^C - E_h^X + E_h^I = E_h^C - (\sum_M \frac{E^C}{G^C} \cdot X^{C\rightarrow M})_h + (\sum_M \frac{E^M}{G^M} \cdot I^{M\rightarrow C})_h \quad (1)$$

In Eq. (1), the total emissions for CAISO at a specific hour (E_h) are calculated as the total emissions from power plants within the CAISO region (E_h^C) plus the emissions associated with imports from neighboring BAs (E_h^I) minus the emissions associated with exports to neighboring BAs (E_h^X). The emissions associated with exports are calculated as the sum over all neighboring BAs of the average emissions for CAISO (C) at a specific hour ($\frac{E^C}{G^C}$) times the amount of exports ($X^{C\rightarrow M}$) to each BA (M) during that hour. The emissions associated with imports are calculated as the sum over all BAs of the average emissions for each BA (M) at a specific hour ($\frac{E^M}{G^C}$) times the amount of imports from that BA ($I^{M\rightarrow C}$) during that hour. This method assumes that the electricity traded between BAs is reflective of the overall grid mix of the BA in question at a that specific time. After aligning these calculated CAISO hourly emissions values with hourly demand and hourly VRE generation, we then differentiate the demand, VRE, and emissions at the hourly level to create three additional features representing the hourly changes in these variables.

2.2. Composite model design

The model developed to calculate MEFs and predict the hourly changes in CAISO emissions is depicted in Fig. 1. The composite model consists of a MLP network followed by a multi-variable linear model. The MLP takes input features for a given hour and predicts the coefficients of the linear model. The linear model has a term for changes in demand (ΔD) with corresponding coefficient MEF, changes in VRE generation (ΔVRE) with corresponding coefficient MDF, and an intercept term c. These predicted terms are then plugged in to the linear model to calculate a predicted change in hourly emissions. Comparing this value to the ground truth change in emissions data, a loss term is calculated and used to train the MLP model. This structure provides a way to learn MEF, MDF, and intercept coefficients despite the lack of ground-truth values for these variables by using the error from the predicted change in emissions to inform the predictions of the latent coefficient space.

The input features used for the MLP model are hourly demand in the CAISO region, hourly variable renewable generation in CAISO, hour of the day, day of the year, and time since 2018 in number of hours. These features were selected to capture as many of the dynamics of the electric grid as possible (without introducing unnecessary noise) and because they can be known or estimated ahead of time for forecasting. The level of demand and amount of VRE generation provide information on the type of marginal generator at a given point in time because many resources are used for specific levels of net load. Hour of the day can also be a good predictor of which resource or powerplant is next in the generation queue, and additionally may capture scheduled grid behaviors such as electricity trade commitments or hydropower operation. The day of the year was included to capture seasonal effects, such as hydropower generation, and time since 2018 was incorporated to capture general trends MEFs and MDFs, such as changes caused by the mix of grid resources getting cleaner over time (California's in-state generation increased from 21% in 2019 to 23% in 2020 and 25% in 2021 Commission, 2023).

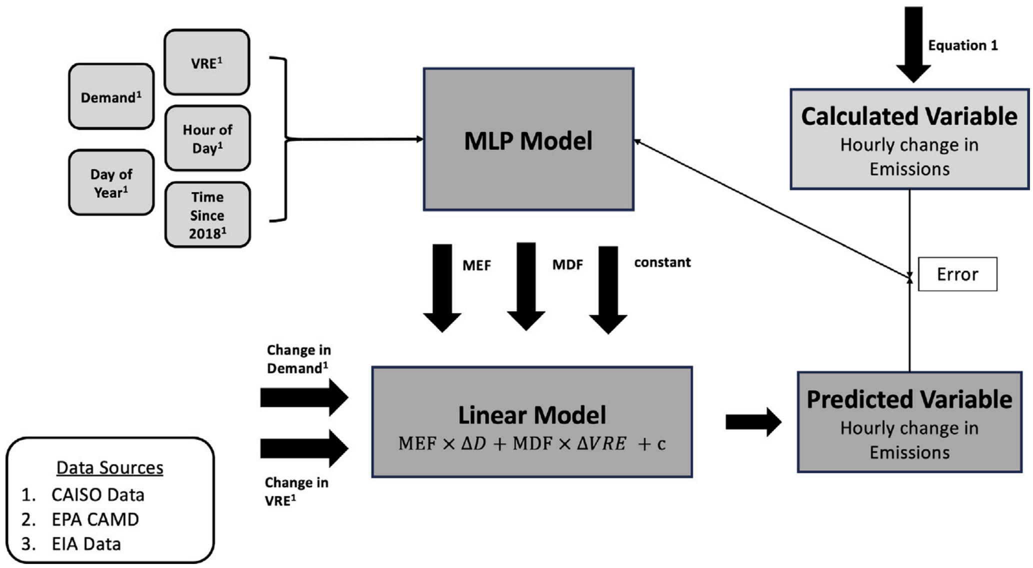

Fig. 1. Conceptualization of the data sources used, the MLP-linear composite model, and the learning process. The model predicts hourly changes in the emissions associated with demand in the CAISO region.

2.3. Model implementation

The three years of hourly data was first randomly split into train, validation, and test sets with a 60/20/20 split. All of the data was then standardized based on the mean and variance of the train set. During training, R-squared was computed on the validation set after each epoch and the model with the highest R-squared seen during this process was chosen as the final model. This model was then evaluated on the test set.

The structure and key parameters of the MLP model were determined via a standard grid search and manual tuning. The MLP has two hidden layers, the first of size 512 followed by 256. After each hidden layer there is a batch normalization layer, a Rectified Linear Unit activation, and a dropout layer with a dropout probability of .5. The model was trained for 40,000 epochs using a full batch size and the AdamW optimizer with a learning rate of .01 and a weight decay of .003. The loss function used was mean squared error with a regularization term that penalized MEF and MDF values that fell outside of a reasonable range.

In the process of developing this model, we explored additional features (such as weather variables) as well as more complex neural networks (long short-term memory models and attention-based models) but found that these alternatives did not significantly increase the accuracy of our model, and often decreased ease-of-use, especially for forecasting. The final version of the model used in this analysis, as well as the data used for model training, testing, and validation, is available in an online data repository (https://github.com/S3researchUSC/MEF-Regression), with further data available upon request.

2.4. Model outputs for historical data

Using the final trained model, we determined historical MEFs, MDFs, and intercepts at the hourly level for CAISO for 2019–2021. We assess the accuracy of our model by calculating the R-squared values and mean absolute error (MAE) between our model's predicted change in emissions and historical actual change in emissions, and compare these results to predictions made with a multi-variable (MV) regression model. The MV regression model used for comparison is based on the work of Zohrabian, Mayes, and Sanders and provides a reference point for the accuracy level of these more granular MEFs. After the prediction step, we perform correlation analyses and a feature importance analysis using Shapley Values calculated by Shap.DeepExplainer (Lundberg, 2023) and discuss the significance of the results. (Shapley Values have been used for feature importance of neural networks in power systems in previous research Zhang et al., 2020.) Shapley Values are a way of measuring the relative importance of each feature and are calculated by removing a specific feature, finding all permutations of the remaining features (including the null set), and then evaluating the marginal contribution of the feature to the estimate (i.e., how much does including this feature increase the accuracy of the model across all combinations of features). This process is repeated for each feature and the results are used to assign a value that represents the importance of each feature (Lundberg and Lee, 2017).

To assess the accuracy of our model for forecasting tasks, we use historical day-ahead forecasts of demand and VRE available on CAISO's website (CAISO, 2023) to create hypothetical demand and VRE forecasts for our test set. First, we split historical demand and VRE data by quintile and calculate the distribution of forecast errors in that quintile for the concurrent forecasts of demand and VRE. Then, we use the historical demand and VRE data from our test set and randomly sample the error distributions to derive hypothetical forecasts of VRE and demand for the test set hours. By using this sampling method, we can assess the sensitivity of our model to the typical level of inaccuracy present in CAISO's demand and VRE forecasting data, determining how accurate our model would be if it relied on forecasted data for these features (the remaining features – day of the year, hour of the day, and time since 2018 – are exactly defined for all future times and dates). We apply this method to the test set so that the model is being tested on forecasts of data that were not included in the training phase, though we note that true forecasting would expand beyond the range of trained values for the time since 2018 feature. We use the constructed day-ahead forecasts of demand and VRE to predict the hourly change in emissions and compare the accuracy of these predictions to those made using actual demand and VRE generation data.

3. Results and discussion

3.1. Hourly MEFs

The hourly MEFs estimated by our composite model using historical data for 2019, 2020, and 2021 are shown in Fig. 2. In all three years, MEFs vary significantly both throughout the day and over the course of the year. Higher MEFs are consistently seen in the late afternoon/early evening hours, but MEFs can also be high during the middle of the day,

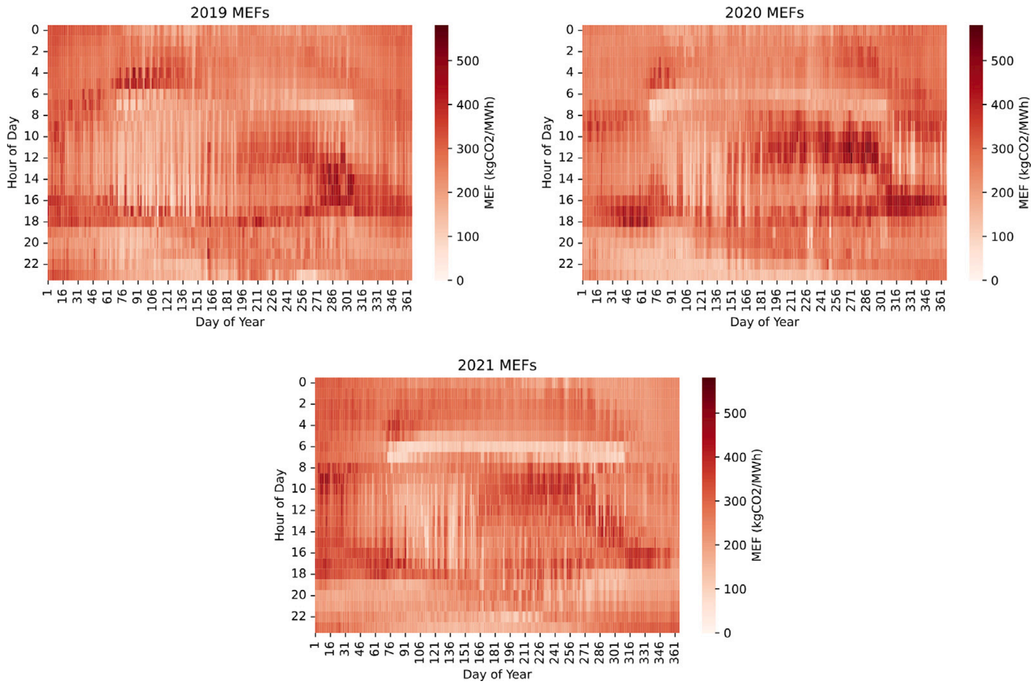

Fig. 2. Hourly marginal emissions factors calculated with the composite model for CAISO for 2019–2021.

Table 1
Accuracy of different MEF models when using actual historical data and historical forecasted data as inputs. Only the MLP composite model is structured to use forecasted data.

Metric	2019–2021 historical data		Day-ahead forecast of test set
	Binned MV regression	Composite model	Composite model
R-squared	.85	.91	.88
MAE	156,000	122,000	133,000

especially during fall months. MEFs tend to be low in the early morning throughout the year and during the middle of the day in spring and summer months. In the early morning (6–8 am), there is typically a rapid increase in VRE generation and a small or no increase in demand; this may lead to an underestimation of MEFs as it is difficult to attribute changes in emissions to changes in demand. Low MEFs in the middle of the day in spring and summer months are expected when daytime solar generation is high, and particularly low during the times when hydropower, wind output and solar generation are simultaneously high (typically in spring months) (CAISO, 2023). These hours of very high daytime solar often experience curtailment, and at times, marginal generation may be effectively emissions free (Operator, 2023a).

The MEFs calculated with this proposed methodology are notably more accurate for historical data than those calculated with the multi-variable (MV) linear regression model as shown in Table 1. Table 1 includes the accuracy of the forecasting portion of this analysis, which is discussed in Section 3.2. Our composite model outperforms the multi-variable regression model at predicting changes in emissions as measured by both R-squared and mean absolute error. Considering the significant increase in granularity achieved with our model, this result suggests that MEFs calculated with this methodology are preferable for DSM applications.

The 22% reduction in MAE is significant for DSM applications, where the specific strategies depend on accurate estimates on the emissions impacts. This effect is magnified by the increased granularity, with a year being represented by 8760 distinct MEFs as opposed to the 288 MEFs created when binning by month and hour. With this level of information, DSM strategies can take advantage of MEF variations that occur on specific days but do not present themselves as diurnal patterns over longer time periods.

A number of factors may contribute to uncertainty in MEFs and explain the remaining inaccuracy in the predictions of changes in emissions. For example, limitations in the quality of the data, which was combined and merged across multiple sources, and data pre-processing, which relied on a hierarchy of data sources and a degree of temporal aggregation, may create inaccuracies when estimating MEFs. Additionally, events such as powerplant maintenance or planned in-operation may impact the queue of generation resources in a way that is difficult for the MLP model to predict or recognize.

3.2. Forecasting results

The results of the forecasting analysis show that day-ahead forecasts of demand and VRE generated from our composite model produce reliable estimates of MEFs. The percent difference between MEFs predicted with actual versus forecasted data are shown at the hourly level for 2021 in Fig. 3, with a mean absolute difference of 9% for the entire year.

The difference between the MEFs based on forecasted data and those based on actual data is a function of inaccuracies in CAISO's forecasts of electricity demand and VRE generation. While CAISO's forecast of demand is highly accurate (mean absolute error of less than 1%), its forecast of VRE is more uncertain (mean error of approximately 15%). Fig. 3 shows that an underestimation of the MEF is more common in the middle hours of the day when using forecasted data. While there are occasionally large differences in forecasted MEFs versus those calculated with actual data, the difference is less than 10% for 72% of the hours of the year, providing large utility to DSM planning.

3.3. Drivers of MEFs and feature importance

The MEFs predicted by the MLP model show many non-linear behaviors, with the correlations between MEFs and variables such as

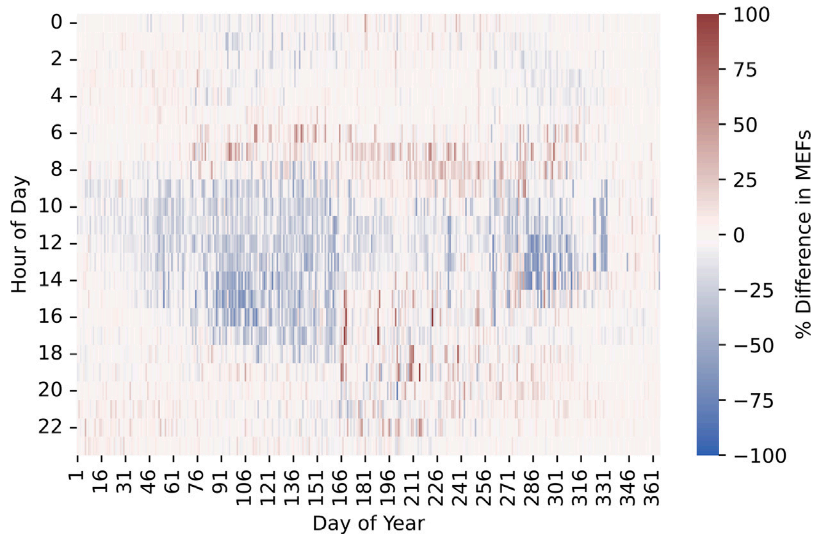

Fig. 3. Hourly percent difference between the MEFs predicted by the composite model using actual versus forecasted CAISO demand and variable renewable energy for 2021.

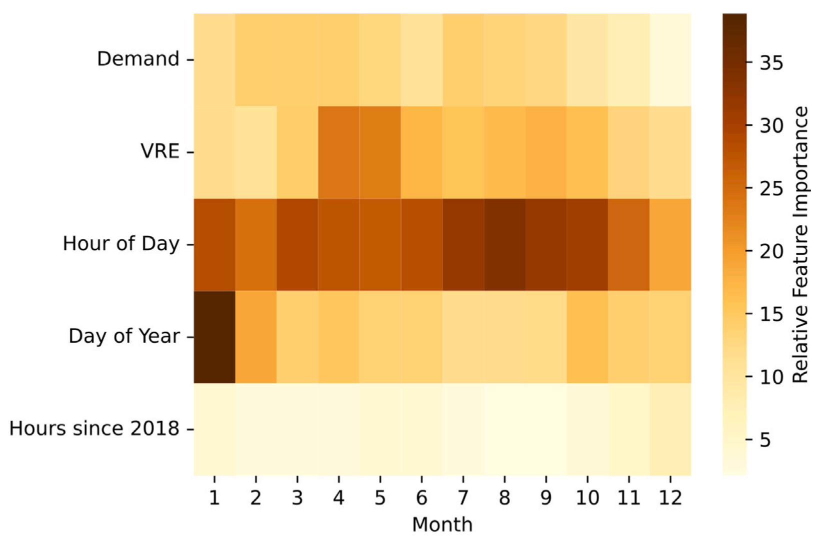

Fig. 4. Mean Shapley value by month of 2021.

demand, net demand, change in demand, and change in net demand being weak for all portions of the studied period (see Table 2). This non-linear behavior emphasizes the need for sophisticated models that can capture multiple contributing factors and the complex relationship between marginal emissions and grid behavior. Additionally, demand-related variables show reducing correlation with MEFs from 2019 to 2021, while hour of the day exhibits a stronger relationship.

To examine the relative importance of each input feature to our MLP model, we calculated Shapley values and show them by month for 2021 in Fig. 4. Hour of the day was generally the most valuable feature for reducing model error (difference between predicted and actual changes in emissions plus regularization term). Day of the year, which was included to capture seasonal variations, was an important feature, especially in months with relatively flat MEF and MDF levels, where simply knowing the time of year is enough to make an accurate estimate. This suggests that DSM planners could make use of just these variables to plan load-shifting and DR strategies that reduce emissions, though our results show this would be less effective than using the granular, forecasted MEFs developed in this study. VRE was most important in April and May, months with significant solar and wind production that frequently reached curtailment levels in the middle of the day (over 500 GWh of VRE production was curtailed throughout April and May of 2021) (Operator, 2023a). The amount of demand was of medium importance throughout the year, though generally lower in winter months with flatter levels of demand. The least important feature was found to be the total time elapsed since 2018, which was included to capture small general trends, such as a grid mix that is becoming higher percentage renewable energy over the three-year

Table 2
Correlation between MEFs and measures of grid load (i.e., hourly demand and net load), hourly changes in measures of grid load, and hour of the day. Correlations are shown for each year of the study period as well as for the entire period. Correlation is measured by Pearson's rho for continuous variables, and pseudo-rho (square root of goodness-of-fit R-squared) for categorical variables.

	Hourly demand	Hourly net load	Hourly change in demand	Hourly change in net load	Hour of day
2019 MEFs	0.19	0.25	0.23	0.33	0.38
2020 MEFs	0.20	0.11	0.30	0.31	0.47
2021 MEFs	0.05	−0.04	0.18	0.18	0.50
2019–2021 MEFs	0.15	0.10	0.24	0.27	0.41

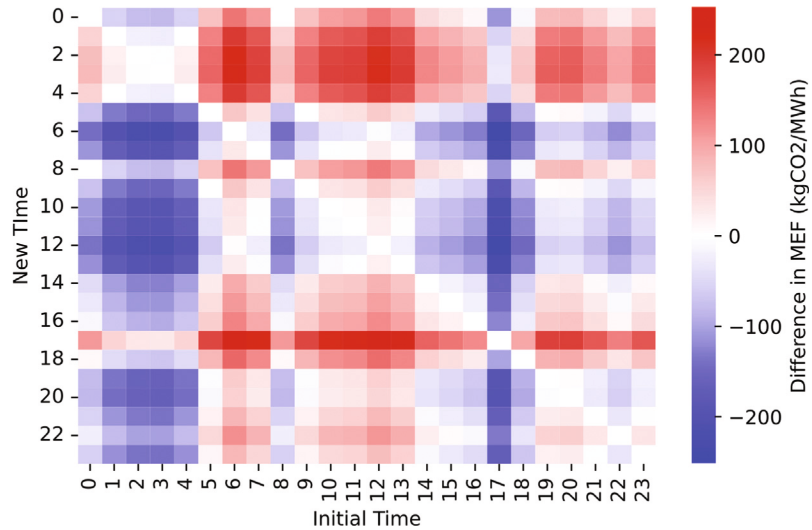

Fig. 5. Difference in MEFs between an initial hour and a new hour on April 17th, 2021 that an activity could occur in. Shades of blue represent switches in the timing of an activity that would reduce emissions.

period. It should be noted that these Shapley values pertain to the importance of each feature in reducing prediction error, and therefore are relevant to MEF, MDF and intercept; the MEFs themselves may have a different relative feature importance.

3.4. Application to DSM

MEFs that can be forecasted 24 h in advance offer a range of possibilities for using DSM to reduce CO_2 emissions, helping to meet decarbonization goals without relying on new technologies or conservation. The large variations in these granular MEFs for all timescales, and the lack of direct correlation with variables such as demand or hour of the day, make it difficult to devise general strategies for DSM. Instead, forecasted MEFs can be made available to the public and used to create short-term advice for load shifting. For example, flexible loads could be shifted from higher MEF to lower MEF hours, reducing emissions without reducing demand. Fig. 5 illustrates the difference in MEF caused by changing the timing of an electricity consuming behavior from one hour to another, using April 17th, 2021 as an example. On this day, it would be particularly valuable to shift demand away from 5 pm, and generally valuable to shift demand from the 4 to 6 pm window (i.e., peak demand hours) towards the 9 am to 1 pm window, when solar resources are highest.

Beyond load-shifting, researchers, policy-makers, and industry members could use forecasted MEFs to devise lower emissions solutions for adding a flexible load to the grid (e.g., EV charging). Hourly MEFs can be used to minimize the amount of emissions associated with a flexible load, subject to constraints such as a window of time during which it is desirable for the task to be completed. This concept could be implemented by a variety of electricity providers and consumers. Time-of-use plans, where the price per unit of electricity varies by time of day (Wang and Li, 2015), or real-time pricing, where consumers experience a rate based on the price of electricity in the whole-sale market (Allcott, 2009), are utility rate structure options designed to inform the timing of electricity usage to save utilities money, reduce peak demand, and increase grid stability. Similarly, electricity providers could incentivize consumption patterns that reduce the total amount of emissions associated with electricity consumption. While these reliability and climate mitigation goals may align during certain times of the year, Fig. 2 shows that MEFs can also be high outside of typical grid peak hours. Further, Table 2 shows that MEFs are only slightly correlated with demand, so new plans or incentives could be designed to encourage emissions-reducing consumption patterns in addition to patterns that improve grid reliability. Emphasis could be put on the periods when these benefits align, or there may be situations where prioritizing reducing peak demand is a priority (e.g., during hot summer months) and others where grid reliability is a smaller concern and the focus should be on reducing emissions. Utilities could also create plans that weight both emissions and grid benefits or give customers the opportunity to choose their priority. On the end-user side, smart home technologies (Marikyan et al., 2019) could be designed so that consumers have the ability to choose flexible appliance and EV charging schedules that prioritize CO_2 emissions reductions, in addition to cost savings.

3.5. Limitations and future improvements

Through this methodology we were able to increase the accuracy and granularity of historical estimates of MEFs as well as accurately forecast day-ahead MEFs. Though these MEFs are well suited for DSM tasks, several potential improvements remain. The temporal resolution of MEFs is limited by the least granular dataset used in the emissions calculation, which is reported at the hourly level. More granular MEFs may be able to better describe the behavior of CAISO's dynamic grid. This model may also benefit from additional inputs such as spatial information about the location of demand and generation that allows the model to better capture transmission constraints, which are expected to have a significant impact on grid-decarbonization (Brown and Botterud, 2021).

The emissions accounting methodology used in this study does not account for the emissions associated with grid-level storage. Although storage was used minimally in the years covered by this project, CAISO has rapidly expanded battery capacity in recent years from less than 1 GW in mid 2021 to over 6 GW in 2023 (CAISO, 2023). The use of storage should be addressed via shifting the emissions associated with charging a battery to the time that the battery is discharging (when the demand "actually" occurs), as was done in Thomson et al. (2017). Shifting emissions in this manner would likely increase MEFs for hours during which batteries are discharged and decrease MEFs for hours during which batteries are being charged. However, co-located renewable energy and storage plants may need to be treated differently, as this storage would be emissions-free.

A remaining question is how to use MEFs for DSM given the influence of flexible, emissions-free resources. These resources are sometimes strategically reserved for high-demand periods (for example, the use of hydropower on CAISO's grid Zohrabian et al., 2023) leading to a low MEF during a period of high net demand. Under these conditions, shifting demand to this period may not actually decrease emissions and could increase grid stress.

As mentioned in the methodology, we replicated the MEF forecasting task via historical forecasts of demand and VRE for our model test set, but future forecasting would feature predictions for hours outside of the temporal range of the training data. We believe this would have minimal impact on the accuracy of the model as this would only impact the "time since 2018" feature, and the testing range of this feature would only be slightly beyond the training range (24 h ahead).

Lastly, applying this methodology to other regions depends on the amount and quality of data made available by other Independent System Operators (ISOs). This project used data from CAISO for demand, VRE generation, forecasted demand, forecasted VRE generation, and imports. Some ISOs share sufficient information on generation, demand, and trades to replicate this methodology exactly, but others might require modifications, such as using EIA data alone to calculate electricity trades or requiring the researcher to develop their own method of forecasting demand and VRE generation. While this method was designed for grids with high fractions of renewable energy, this method should maintain high accuracy for primarily fossil-fuel grids due to the isolation of the demand impact and renewable energy impact achieved in the multivariate model.

4. Conclusions

The novel methodology introduced in this study produces accurate and high-granularity estimates of both historical and day-ahead MEFs, filling a gap in the existing literature on statistical MEFs. Accurately forecasted MEFs can be utilized for a variety of DSM applications that aim to quantify the CO_2 emissions associated with changing the demand for electricity. Hence, they can be leveraged to reduce CO_2 emissions in addition to the traditional aims of DSM such as reducing costs for electricity producers and consumers. This study found significant variations in hourly MEFs both between days, and throughout the hours of the day, suggesting that shifting the timing of a flexible load can significantly change the CO_2 emissions associated with that consumption. Our results also show that MEFs are not highly correlated with grid-level demand or net demand, underscoring the need for a flexible, non-linear MEF model, such as a neural network. This methodology will become increasingly applicable as grids across the US integrate more renewables and rely more heavily on DSM.

CRediT authorship contribution statement

Stepp Mayes: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Nicholas Klein:** Methodology, Validation, Formal analysis, Writing – original draft. **Kelly T Sanders:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

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

Data availability

The github repository is linked in the manuscript.

Funding

This work was supported by the National Science Foundation, United States [NSF CAREER Award: 184593: Coordinating climate change mitigation and adaptation strategies across the energy-water nexus: an integrated research and education framework] and the USC Viterbi College of Engineering's Ershaghi Center for Energy Transition.

References

Agency, U.E.P., 2023. Today's outlook: Emissions. URL http://www.caiso.com/TodaysOutlook/Pages/emissions.html.
Allcott, H., 2009. Real Time Pricing and Electricity Markets, Vol. 7. Harvard University.
Beltrami, F., Burlinson, A., Giulietti, M., Grossi, L., Rowley, P., Wilson, G., 2020. Where did the time (series) go? Estimation of marginal emission factors with autoregressive components. Energy Econ. 91, 104905.
Braeuer, F., Finck, R., McKenna, R., 2020. Comparing empirical and model-based approaches for calculating dynamic grid emission factors: An application to CO2-minimizing storage dispatch in Germany. J. Clean. Prod. 266, 121588.
Brown, P.R., Botterud, A., 2021. The value of inter-regional coordination and transmission in decarbonizing the US electricity system. Joule 5, 115–134.
CAISO, 2023. Open acess same-time information system. URL http://oasis.caiso.com/mrioasis/logon.do.
Catalão, J.P.d.S., Mariano, S.J.P.S., Mendes, V., Ferreira, L., 2007. Short-term electricity prices forecasting in a competitive market: A neural network approach. Electr. Power Syst. Res. 77 (10), 1297–1304.
Chae, Y.T., Horesh, R., Hwang, Y., Lee, Y.M., 2016. Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. Energy Build. 111, 184–194.
Commission, C.E., 2022. Total system electric generation. URL https://www.energy.ca.gov/data-reports/energy-almanac/california-electricity-data/2021-total-system-electric-generation.
Commission, C.E., 2023. Total system electric generation. URL https://www.energy.ca.gov/media/7311.
de Chalendar, J.A., Taggart, J., Benson, S.M., 2019. Tracking emissions in the US electricity system. Proc. Natl. Acad. Sci. 116 (51), 25497–25502.
Deetjen, T.A., Azevedo, I.L., 2019. Reduced-order dispatch model for simulating marginal emissions factors for the United States power sector. Environ. Sci. Technol. 53 (17), 10506–10513.
Denholm, P., O'Connell, M., Brinkman, G., Jorgenson, J., 2015. Overgeneration from Solar Energy in California. A Field Guide to the Duck Chart. Tech. Rep., National Renewable Energy Lab.(NREL), Golden, CO (United States).
Donti, P.L., Kolter, J.Z., Azevedo, I.L., 2019. How much are we saving after all? Characterizing the effects of commonly varying assumptions on emissions and damage estimates in PJM. Environ. Sci. Technol. 53 (16), 9905–9914.
EIA, U., 2022. Electricity data. URL https://www.eia.gov/electricity/data.php.
Gagnon, P., Cole, W., 2022. Planning for the evolution of the electric grid with a long-run marginal emission rate. Iscience 25 (3), 103915.
Gai, Y., Wang, A., Pereira, L., Hatzopoulou, M., Posen, I.D., 2019. Marginal greenhouse gas emissions of ontario's electricity system and the implications of electric vehicle charging. Environ. Sci. Technol. 53 (13), 7903–7912.
Hawkes, A.D., 2010. Estimating marginal CO2 emissions rates for national electricity systems. Energy Policy 38 (10), 5977–5987.
Hawkes, A.D., 2014. Long-run marginal CO2 emissions factors in national electricity systems. Appl. Energy 125, 197–205.
Holland, S.P., Kotchen, M.J., Mansur, E.T., Yates, A.J., 2022. Why marginal CO2 emissions are not decreasing for US electricity: Estimates and implications for climate policy. Proc. Natl. Acad. Sci. 119 (8), e2116632119.
Huber, J., Lohmann, K., Schmidt, M., Weinhardt, C., 2021. Carbon efficient smart charging using forecasts of marginal emission factors. J. Clean. Prod. 284, 124766.
Kamiya, G., Axsen, J., Crawford, C., 2019. Modeling the GHG emissions intensity of plug-in electric vehicles using short-term and long-term perspectives. Transp. Res. D 69, 209–223.
Li, M., Smith, T.M., Yang, Y., Wilson, E.J., 2017. Marginal emission factors considering renewables: A case study of the US Midcontinent independent system operator (MISO) system. Environ. Sci. Technol. 51 (19), 11215–11223.
Lundberg, S., 2023. Shap.DeepExplainer. URL https://shap-lrjball.readthedocs.io/en/latest/generated/shap.DeepExplainer.html.
Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, Vol. 30.
Marikyan, D., Papagiannidis, S., Alamanos, E., 2019. A systematic review of the smart home literature: A user perspective. Technol. Forecast. Soc. Change 138, 139–154.
Mayes, S., Sanders, K., 2022. Quantifying the electricity, CO2 emissions, and economic tradeoffs of precooling strategies for a single-family home in Southern California. Environ. Res.: Infrastruct. Sustain. 2 (2), 025001.
McKenna, E., Barton, J., Thomson, M., 2017. Short-run impact of electricity storage on CO2 emissions in power systems with high penetrations of wind power: A case-study of Ireland. Proc. Inst. Mech. Eng., A 231 (6), 590–603.

Noussan, M., Neirotti, F., 2020. Cross-country comparison of hourly electricity mixes for EV charging profiles. Energies 13 (10), 2527.

Operator, C.I.S., 2023a. Managing oversupply. URL http://www.caiso.com/informed/Pages/ManagingOversupply.aspx.

Operator, C.I.S., 2023b. Today's outlook: Emissions. URL http://www.caiso.com/TodaysOutlook/Pages/emissions.html.

Operator, C.I.S., 2023c. Understanding the ISO. URL http://www.caiso.com/about/Pages/OurBusiness/Default.aspx.

Park, D.C., El-Sharkawi, M., Marks, R., Atlas, L., Damborg, M., 1991. Electric load forecasting using an artificial neural network. IEEE Trans. Power Syst. 6 (2), 442–449.

Pimm, A.J., Palczewski, J., Barbour, E.R., Cockerill, T.T., 2021. Using electricity storage to reduce greenhouse gas emissions. Appl. Energy 282, 116199.

Regett, A., Böing, F., Conrad, J., Fattler, S., Kranner, C., 2018. Emission assessment of electricity: Mix vs. marginal power plant method. In: 2018 15th International Conference on the European Energy Market (EEM). IEEE, pp. 1–5.

Ryan, N.A., Johnson, J.X., Keoleian, G.A., 2016. Comparative assessment of models and methods to calculate grid electricity emissions. Environ. Sci. Technol. 50 (17), 8937–8953.

Samaras, C., Meisterling, K., 2008. Life cycle assessment of greenhouse gas emissions from plug-in hybrid vehicles: implications for policy.

Seckinger, N., Radgen, P., 2021. Dynamic prospective average and marginal GHG emission factors—Scenario-based method for the german power system until 2050. Energies 14 (9), 2527.

Sengupta, S., Spencer, T., Rodrigues, N., Pachouri, R., Thakare, S., Adams, P.J., Tongia, R., Azevedo, I.M., 2022. Current and future estimates of marginal emission factors for Indian power generation. Environ. Sci. Technol. 56 (13), 9237–9250.

Siler-Evans, K., Azevedo, I.L., Morgan, M.G., 2012. Marginal emissions factors for the US electricity system. Environ. Sci. Technol. 46 (9), 4742–4748.

Singhal, D., Swarup, K., 2011. Electricity price forecasting using artificial neural networks. Int. J. Electr. Power Energy Syst. 33 (3), 550–555.

Stopps, H., Touchie, M.F., 2022. Load shifting and energy conservation using smart thermostats in contemporary high-rise residential buildings: Estimation of runtime changes using field data. Energy Build. 255, 111644.

Thind, M.P., Wilson, E.J., Azevedo, I.L., Marshall, J.D., 2017. Marginal emissions factors for electricity generation in the midcontinent ISO. Environ. Sci. Technol. 51 (24), 14445–14452.

Thomson, R.C., Harrison, G.P., Chick, J.P., 2017. Marginal greenhouse gas emissions displacement of wind power in Great Britain. Energy Policy 101, 201–210.

U.S. Energy Information Administration (EIA), 2019. Annual Electric Generator data - EIA-860 data file. URL www.eia.gov/electricity/data/eia860/.

Wang, Y., Li, L., 2015. Time-of-use electricity pricing for industrial customers: A survey of US utilities. Appl. Energy 149, 89–103.

Zhang, K., Xu, P., Zhang, J., 2020. Explainable AI in deep reinforcement learning models: A shap method applied in power system emergency control. In: 2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2). IEEE, pp. 711–716.

Zheng, Z., Han, F., Li, F., Zhu, J., 2015. Assessment of marginal emissions factor in power systems under ramp-rate constraints. CSEE J. Power Energy Syst. 1 (4), 37–49.

Zohrabian, A., Mayes, S., Sanders, K.T., 2023. A data-driven framework for quantifying consumption-based monthly and hourly marginal emissions factors. J. Clean. Prod. 396, 136296.

Zohrabian, A., Sanders, K.T., 2021. Emitting less without curbing usage? Exploring greenhouse gas mitigation strategies in the water industry through load shifting. Appl. Energy 298, 117194.