FairST: Equitable Spatial and Temporal Demand Prediction for New Mobility Systems

An Yan
yanan15@uw.edu
University of Washington
Seattle, WA

Bill Howe
billhowe@uw.edu
University of Washington
Seattle, WA

ABSTRACT

We present a fairness-aware model for predicting demand for new mobility systems. Our approach, called FairST, consists of 1D, 2D and 3D convolutions to learn the spatial-temporal dynamics of a mobility system, and fairness regularizers that guide the model to make equitable predictions. We propose two fairness metrics, region-based fairness gap (RFG) and individual-based fairness gap (IFG), that measure equity gaps between social groups for new mobility systems. Experimental results on two real-world datasets demonstrate the effectiveness of the proposed model: FairST not only reduces the fairness gap by more than 80%, but achieves better accuracy than state-of-the-art but fairness-oblivious methods including LSTMs, ConvLSTMs, and 3D CNN.

CCS CONCEPTS

• **Information systems** → *Spatial-temporal systems*; *Data analytics*; • **Applied computing** → *Transportation.*

KEYWORDS

fairness in machine learning, convolutional neural networks, equity, new mobility, spatial-temporal data mining

ACM Reference Format:
An Yan and Bill Howe. 2019. FairST: Equitable Spatial and Temporal Demand Prediction for New Mobility Systems. In *27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '19), November 5–8, 2019, Chicago, IL, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3347146.3359380

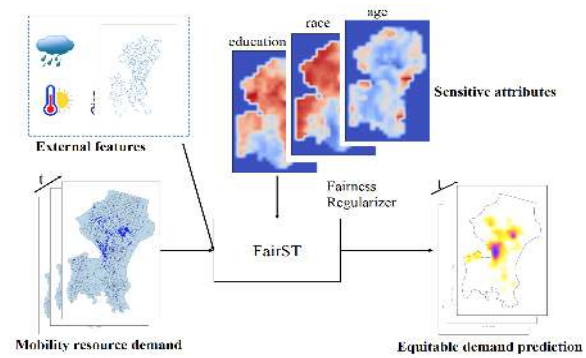

Figure 1: FairST is a Fairness-aware Spatial-Temporal model for new mobility systems. By incorporating a fairness regularizer, it makes equitable demand predictions between groups defined by, for example, race, age, or education level.

1 INTRODUCTION

New mobility systems such as bike-sharing and ride-hailing have been introduced in many cities, offering affordable and on-demand transportation options for citizens. Accurate demand prediction is crucial to guide resource allocation [8] in new mobility systems. For example, ride-hailing companies predict demand to direct drivers to high-demand areas [1]. Likewise, bikeshare operators *rebalance* bikes from low-demand to high-demand areas based on demand estimates [9]. However, new mobility services have been shown to reinforce socioeconomic inequities [7]. Underestimation of resource demand for the disadvantaged groups may result in insufficient supply to them which can produce a feedback loop: social disparities are misinterpreted in the model as lack of demand, reinforcing reduced access to services. Therefore, spatiotemporal demand prediction models for new mobility regimes must consider fairness as a first-class design requirement.

In this paper, we incorporate fairness in a model for new mobility resource demand prediction. We address three challenges to achieve this goal: accurate modeling of the spatial-temporal dynamics of mobility resource demand, defining novel fairness metrics suitable for this task, and effective integration of fairness into the model.

Modeling the spatial-temporal dynamics of mobility resource demand. Resource demand exhibits complex spatial and temporal patterns, and is influenced by many factors such as weather and road network [8]. We use a 3D convolutional neural network (3D CNN) as the core building block in our model to capture spatial-temporal dynamics [10]. To integrate exogenous features, we adopt a three-stream architecture that fuses together 1D, 2D, and 3D convolutional layers, respectively. A 1D CNN is used to extract 1D temporal features such as city-wide temperature, and a 2D CNN is used to extract 2D spatial features such as bike lanes.

Designing fairness metrics for mobility resource demand. Although fairness metrics is an active research area, there do not exist metrics for continuous spatio-temporal prediction problems. Many fairness metrics such as statistical parity are designed for classification [5] and typically require a categorical sensitive attribute (e.g., white) [2, 3]. However, in our settings, both the prediction target (e.g. bike demand) and the sensitive attributes (e.g., the percentage of white in a region) are usually numerical values. Moreover, mobility resource demand is associated with zonal population [6], so fairness metrics should be designed on a per capita basis. To

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SIGSPATIAL '19, November 5–8, 2019, Chicago, IL, USA
© 2019 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-6909-1/19/11.
https://doi.org/10.1145/3347146.3359380

address these challenges, we interpret fairness in demand prediction as the requirement that individuals of different groups have access to a similar amount of the resource. We propose two metrics: *region-based fairness gap (RFG)* and *individual-based fairness gap (IFG)*. Both assess the gap between mean per capita demand across groups over a period of time. However, RFG assumes that a distinct label (e.g., white) is assigned to a region. IFG instead is assigned a distribution (e.g., percentage white) based on demographics.

Integrating fairness into the prediction model. Fairness can be enforced during different stages of the machine learning, including data preprocessing, model training, and postprocessing [2, 5]. We propose two regularizers, corresponding to RFG and IFG, to be encoded as regularizers in the loss function during training. To the best of our knowledge, our work is the first to incorporate fairness in a spatial-temporal mobility setting using deep neural networks.

We introduce FairST, a **Fair**ness-aware **S**patial-**T**emporal model that accounts for dynamics of mobility resource demand and enforces fairness through regularizers (Figure 1).Our experiments on two real-world datasets demonstrate that FairST effectively closes the fairness gaps while outperforming state-of-the-art but fairness-oblivious models. FairST can be extended to other spatial-temporal scenarios with fairness concerns such as crime incidence prediction.

2 USE CASES

2.1 Datasets

We obtained **Seattle dockless bikeshare data** from the Transportation Data Collaborative operated by the University of Washington. The data spans from October 1, 2017 to October 31, 2018, including over 1.6 million trips. We use the number of pickup as a proxy for demand. **RideAustin** [1] is a ride-hailing service in Austin, Texas. Rides data is openly available [2]. The data used in this paper spans from August 1, 2016 to April 13, 2017, including over 1.4 million trips. We use the number of rides as a proxy for demand. **Socioeconomic data** such as population and race for Seattle and Austin at the block group level were obtained from the SimplyAnalytics database [3]. We obtained **1D weather data** such as city-level temperature for both cities from NCEI [4]. We collected **2D urban data** such as bike lanes from open data portals [5].

Data preparation. We place a bounding box around a city and partition the bounding box into equal-sized squares. We choose a grid size of 1km by 1km for Seattle Bikeshare and 2km by 2km for RideAustin. For each grid cell, resource demand forms a time series. For each hour, the study area can be likened to a frame in a video and each region can be seen as a pixel with demand as its value.

2.2 Prediction Problem Definition

We aim to build fair models to forecast next time step demand for mobility resource for a city based on the demand of previous time steps. For Seattle bikeshare and RideAustin, we aim to predict hourly demand based on the demand of the last 7 days. For Seattle bikeshare, we use the data from October 2017 to August 2018 for training and the data from September 2018 to October 2018 for testing. For RideAustin, we use the data from August 2016 to February 2017 for training and the data from March 2017 to April 2017 for testing. The prediction should balance two objectives: minimizing prediction accuracy loss and minimizing fairness loss.

3 MODEL AND FAIRNESS METRICS

3.1 Model Architecture

We design a three-stream prediction framework to 1) automatically capture the spatio-temporal context, and 2) include external features to help with accuracy. We use a 3D CNN submodel to learn from 3D historical resource demand, a 1D CNN submodel to learn from 1D time series features, and a 2D CNN submodel to extract 2D urban information. The outputs of all submodels were fused together, on top of which additional convolutional layers were applied to achieve the final prediction (See Figure 2).

The first submodel is based on 3D convolutions, which model spatial-temporal information. It consists of three 3D convolutional layers, followed by a 2D convolutional layer. The third 3D convolutional layer adopts 1 filter to achieve temporal pooling. Finally, a 2D convolution layer is used to integrate information and output the feature map for submodel fusion.

Training objectives. Our loss function is a weighted sum of an accuracy loss and a fairness loss. The fairness loss acts as a regularizer for the model. We use Mean Absolute Error (MAE) as accuracy loss. The overall loss function is defined as

$$L = L_{accuracy} + \lambda L_{fairness} \tag{1}$$

where $L_{accuracy}$ is MAE, $L_{fairness}$ is the fairness loss, and λ is the weight for the fairness loss. In the next section, we describe details of the proposed fairness loss.

3.2 Fairness Metrics and Regularizers

We consider fairness as individuals of different groups receiving equal resources. In the mobility setting, fair prediction implies adjusting the demand to reduce the difference in per capita demand among groups. Our definition adapts *group fairness* that requires the disadvantaged group to experience similar predicted outcomes as the advantaged group [5], and *vertical equity* that requires transportation policies to favor the disadvantaged groups [4].

We therefore propose two fairness metrics: a Region-based Fairness Gap (RFG) and an Individual-based Fairness Gap (IFG). Both metrics measure the gap between mean per capita demand across two groups over a certain period of time. However, for RFG, each geographic region is assigned a categorical group label (e.g., Caucasian or non-Caucasian) according to some criteria. For IFG, the group label is assigned proportionally based on the region's demographics, so that it is numeric (e.g., percentage of Caucasian). Although here we focus on a square grid partitioning, these two metrics can be used for any customized partitioning.

Notation. We start by introducing notation.

- Let s_i be the ith square region of the study area $\mathcal{S}$.
- Let p_i denote the population of square region s_i divided by the total population of the city.
- Let $\hat{y}_{i,t}$ and $y_{i,t}$ be the estimated demand and ground truth demand for region s_i at time t, respectively.

[1] http://www.rideaustin.com/
[2] https://data.world/ride-austin/ride-austin-june-6-april-13
[3] EASI/MRI Census US. Retrieved from SimplyAnalyticsdatabase
[4] https://www.ncei.noaa.gov/access/search/index
[5] https://data.seattle.gov/, https://data.austintexas.gov/

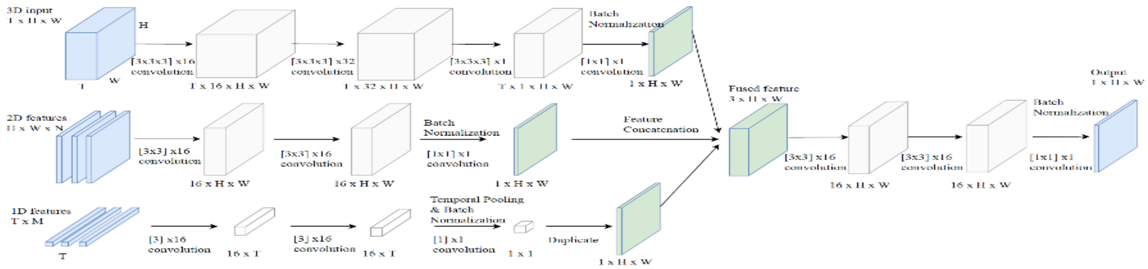

Figure 2: A three-stream network for new mobility resource demand prediction. T, H, W are the number of time steps, height of input, and width of input, respectively. N and M are the number of 2D and 1D features, respectively.

- Let $E_T[\hat{y}_{i,t}]$ be the average predicted value for the ith square region in $\mathcal{S}$ over time period T.

Region-based Fairness Gap (RFG). Let every region s_i be assigned to one of two groups (e.g., Caucasian and non-Caucasian) with regard to one sensitive attribute A (e.g., race), denoted by G^+ (the advantaged group) and G^- (the disadvantaged group). We define RFG between two groups with regard to sensitive attribute A over a period of time T as follow:

$$RFG = \frac{\sum_{i \in G^+} E_T[\hat{y}_{i,t}]}{\sum_{i \in G^+} p_i} - \frac{\sum_{j \in G^-} E_T[\hat{y}_{j,t}]}{\sum_{j \in G^-} p_j} \tag{2}$$

Individual-based Fairness Gap (IFG). Let w_i^+ denote the percentage of people in the advantaged group in region s_i and let w_i^- denote the percentage of people in the disadvantaged group. For example, if a region s_i is 65% white, then $w_i^+ = 65\%$ and $w_i^- = 35\%$. Formally, we define IFG over T as follow:

$$IFG = \frac{\sum_{i \in \mathcal{S}} E_T[\hat{y}_{i,t}] w_i^+}{\sum_{i \in \mathcal{S}} p_i w_i^+} - \frac{\sum_{j \in \mathcal{S}} E_T[\hat{y}_{j,t}] w_j^-}{\sum_{j \in \mathcal{S}} p_j w_j^-} \tag{3}$$

Fairness loss. Based on the RFG and IFG, we define two corresponding fairness loss terms to incorporate fairness into training.

The *Region-based Fairness loss (RF loss)* at time t is defined as

$$L_{RF}(t) = \frac{1}{\sum_{i \in \mathcal{S}} y_{i,t}} \left| \frac{\sum_{i \in G^+} \hat{y}_{i,t}}{\sum_{i \in G^+} p_i} - \frac{\sum_{j \in G^-} \hat{y}_{j,t}}{\sum_{j \in G^-} p_j} \right| \tag{4}$$

The first term is the estimated per capita demand in group G^+ at time t. Likewise, the second term is for group G^-. $\sum_{i \in \mathcal{S}} y_{i,t}$ is a normalizing factor.

The Individual-based Fairness loss (IF loss) at time t is defined as

$$L_{IF}(t) = \frac{1}{\sum_{i \in \mathcal{S}} y_{i,t}} \left| \frac{\sum_{i \in \mathcal{S}} \hat{y}_{i,t} w_i^+}{\sum_{i \in \mathcal{S}} p_i w_i^+} - \frac{\sum_{j \in \mathcal{S}} \hat{y}_{j,t} w_j^-}{\sum_{j \in \mathcal{S}} p_j w_j^-} \right| \tag{5}$$

4 EXPERIMENTS

We evaluate our method on the Seattle dockless bikeshare dataset and the RideAustin dataset. We compare FairST without fairness loss with baseline models in terms of prediction accuracy. We then incorporate RF or IF loss into our model, and compare against two existing fairness regularizers on a sensitive attribute (i.e. race).

Implementation: We train FairST using Adam optimizer with a batch size of 32. The learning rate starts at 0.005 and decays every 5,000 steps with a rate of 0.96. To implement RF loss, we use the overall city statistics as thresholds to discretize the continuous sensitive attributes. For example, the percentage of white of Seattle in 2018 is 65.74%, we then set the regions with more than 65.74% white as Caucasian group, otherwise as non-Caucasian group.

Baseline Models: We compare the prediction accuracy of FairST with several other models: 1) **Historical Average (HA)**. We compute $\hat{y}_{i,t}$ using the mean values of all previous observations at location s_i at the same time of the day and the same day of the week. 2) **Autoregressive Integrated Moving Average Model (ARIMA)**. We develop an independent ARIMA model for each individual grid cell. 3) **Long short-term memory Network (LSTM)**. We train the LSTM model individually for each square grid. 4) **Convolutional LSTM (ConvLSTM)** [11]. ConvLSTM can capture both spatial and temporal dependencies in one network. 5) **3D CNN**. This 3D CNN model is equivalent to FairST without external features.

Baseline Fairness Regularizers: We compare the proposed loss functions (RF loss and IF loss) with two existing fairness losses [2, 3]: 1) **Equal Means Loss (EM Loss)**. Equal Means loss [3] enforces the mean prediction to be the same for different groups. 2) **Pairwise Fairness Loss (Pairwise Loss)**. Pairwise loss is based on the idea of group fairness [2]. Comparisons of predictions across groups are based on cross pairs $i \in G^+$ and $j \in G^-$.

Evaluation Metrics: We evaluate the prediction accuracy of all models with **MAE**. We evaluate the fairness of models using **RFG** and **IFG**, as well as the **Spearman's rho**, which measures the strength of monotonic correlation between two variables. We calculate Spearman's rho between mean per capita demand over the test period of a region and the percent of advantaged population of that region. A positive Spearman's rho with a p-value less than 0.05 suggests disparities in demand. As discussed in Introduction, existing fairness metrics cannot be used directly in our setting.

5 RESULTS AND DISCUSSION

5.1 Demand Prediction Accuracy

Table 1 and Table 2 show MAE of all models on the Seattle bikeshare data and the RideAustin data, respectively. FairST ($\lambda = 0$) proposed by this paper outperforms all other methods. LSTM outperforms ARIMA and HA, due to its capability of modeling complex nonlinear relationships. ConvLSTM achives better accuracy than LSTM, as it can learn both spatial and temporal information. The 3D CNN model performs better than ConvLSTM since the 3D CNN is more powerful in terms of capturing strong local spatial-temporal correlations in our problem as compared to the recurrent architectures.

5.2 Fair Prediction

We trained FairST with Region-based Fairness loss (RF), Individual-based Fairness loss (IF), Equal Means loss (Equal Means), and Pairwise loss (Pairwise) respectively, on a single attribute, i.e. race on

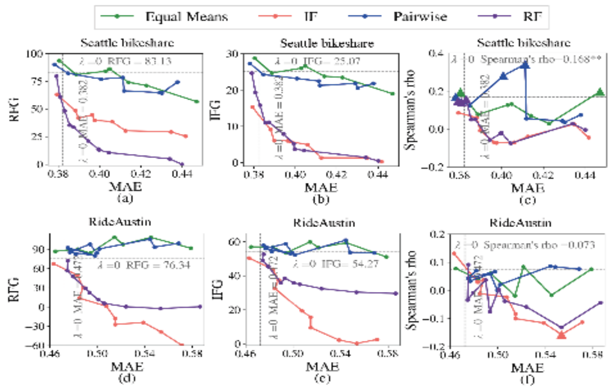

Figure 3: Accuracy vs. fairness metrics. Triangles in (c) and (f) represent statistical significance (p-value < 0.01)

two datasets. Figure 3 illustrates the relationships between MAE and fairness metrics, each point on a curve corresponds to a λ value, which increases from left to right of the curve.

Figure 3 (a), (b), (d), and (e) show that RF and IF regularizer are very effective in controlling both RF and IF gaps. The use of fairness loss actually may *improve* the MAE over the baseline (λ = 0) for small values of λ (see Table 1). For example, IF regularizer (λ = 0.2) brings IFG down while keeping better accuracy than FairST with λ = 0. The reason is that the fairness terms may provide a regularizing effect on accuracy. Figure 3 (c) and (f) show the fairness of models evaluated by Spearman's rho. The use of RF or IF loss effectively "decorrelates" the predicted demand and the percent of Caucasian. For example, in Seattle bikeshare case, FairST without fairness would result in an unfair prediction (see Table 1). Models with an IF or a RF regularizer bring down the Spearman's rho to around zero, and the predictions are no longer significantly correlated with race as λ increases. In contrast, Spearman's rhos of models with an Equal Means or a Pairwise regularizer stay positive throughout.

Table 1 shows the results of FairST for Seattle bikeshare. Both the RF and IF regularizer bring down about 85% IFG (from 31.915 to 3.363 and 4.902, respectively) while keeping better MAE than 3D CNN. Similarly, table 2 shows the results for RideAustin. Compared to 3D CNN, RF regularizer brings down about 99.5% RFG (from 62.004 to 0.347) and IF regularizer brings down 80.5% IFG (from 48.713 to 9.473) while keeping better accuracy.

In summary, FairST is able to achieve an accuracy better than the state-of-the-art baseline models while closing more than 80% of fairness gap. The proposed fairness regularizers are more effective than baseline fairness regularizers in reducing unfairness.

6 CONCLUSION

In this paper, we proposed FairST, a fairness-aware spatial-temporal model based on 3D CNN for predicting new mobility resource demand. A key feature of FairST is the integration of fairness regularizers to the model to encourage equitable prediction. We also proposed two fairness metrics that measure equity gaps between social groups for urban mobility systems. Experiments on two real-world new mobility datasets demonstrate that FairST is able to close more than 80% of fairness gap while achieving *better* accuracy than state-of-the-art but fairness-oblivious baseline methods.

Table 1: FairST compared to baselines for Seattle bikeshare

	λ	MAE	RFG	IFG	Spearman's rho
Ground Truth	/	/	112.568	38.969	0.016
HA	/	0.484	194.454	79.906	0.565
ARIMA	/	0.538	319.032	129.447	0.569
LSTM	/	0.468	280.685	116.023	0.522
ConvLSTM[11]	0.00	0.432	74.485	22.907	0.210**
3D CNN	0.00	0.408	100.878	31.915	0.091
FairST	0.00	0.382	83.127	25.073	0.168**
FairST + RF	0.02	**0.379**	79.570	24.694	0.144**
FairST + RF	0.50	0.404	**10.627**	**3.363**	-0.076
FairST + IF	0.20	**0.379**	63.130	15.281	0.085
FairST + IF	1.50	0.406	38.473	4.902	**-0.070**

**. Correlation is significant at the 0.05 level.
*. Correlation is significant at the 0.01 level.

Table 2: FairST compared to baselines for RideAustin

	λ	MAE	RFG	IFG	Spearman's rho
Ground Truth	/	/	80.120	59.742	0.120*
HA	/	0.662	48.457	33.550	0.118*
ARIMA	/	0.597	82.587	61.457	0.117*
LSTM	/	0.570	61.329	42.101	0.073
ConvLSTM[11]	0.00	0.567	66.428	46.534	0.073
3D CNN	0.00	0.532	62.004	48.713	0.051
FairST	0.00	0.472	76.340	54.274	0.073
FairST + RF	0.05	0.475	56.703	49.092	**-0.034**
FairST + RF	0.80	0.524	**0.347**	32.436	-0.059
FairST + IF	0.06	**0.463**	67.358	50.357	0.131*
FairST + IF	1.20	0.515	-27.397	**9.473**	-0.100

**. Correlation is significant at the 0.05 level.
*. Correlation is significant at the 0.01 level.

REFERENCES

[1] Franziska Bell and Slawek Smyl. 2018. Forecasting at Uber: An Introduction. *Uber Engineering* (2018). https://eng.uber.com/forecasting-introduction/
[2] Richard Berk, Hoda Heidari, Shahin Jabbari, Matthew Joseph, Michael J. Kearns, Jamie Morgenstern, Seth Neel, and Aaron Roth. 2017. A Convex Framework for Fair Regression. *CoRR* abs/1706.02409 (2017).
[3] Toon Calders, Asim Karim, Faisal Kamiran, Wasif Ali, and Xiangliang Zhang. 2013. Controlling attribute effect in linear regression. In *ICDM*. IEEE, 71–80.
[4] Alexa Delbosc and Graham Currie. 2011. Using Lorenz curves to assess public transport equity. *Journal of Transport Geography* 19, 6 (2011), 1252–1259.
[5] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness Through Awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS '12)*. 214–226.
[6] Wafic El-Assi, Mohamed Salah Mahmoud, and Khandker Nurul Habib. 2017. Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. *Transportation* 44, 3 (2017), 589–613.
[7] Yanbo Ge, Christopher R Knittel, Don MacKenzie, and Stephen Zoepf. 2016. *Racial and gender discrimination in transportation network companies.* Technical Report. National Bureau of Economic Research.
[8] Yexin Li, Yu Zheng, and Qiang Yang. [n. d.]. Dynamic Bike Reposition: A Spatio-Temporal Reinforcement Learning Approach. In *KDD '18*.
[9] Stephen J Mooney, Kate Hosford, Bill Howe, An Yan, Meghan Winters, Alon Bassok, and Jana A Hirsch. 2019. Freedom from the station: Spatial equity in access to dockless bike share. *Journal of Transport Geography* 74 (2019), 91–96.
[10] Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, and Manohar Paluri. [n. d.]. Learning Spatiotemporal Features with 3D Convolutional Networks. In *ICCV '15*.
[11] Shi Xingjian, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. 2015. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems*. 802–810.