

A Dynamic Model Averaging for the Discovery of Time-Varying Weather-Cycling Patterns

Guanying Jiang, Ronghui Zhang^{ID}, Xiaobo Qu^{ID}, and Dezong Zhao^{ID}, *Senior Member, IEEE*

Abstract—It has been well recognized that weather variations significantly impact cycling experiences of users. However, the weather-cycling dynamic relationship over time is not well studied in the literature. In this paper, in order to bridge this gap, we propose a Dynamic Model Averaging and Dynamic Model Selection (DMA and DMS) to reveal the characteristics of time-varying responses and the associated influencing factors for young people’s shared bike trips. Without loss of generality, dynamic models with unknown observational variances are also proposed. We take New York City as an instance and analyze the drifts of patterns of New York CitiBike trips under six weather factors from various aspects. The results suggest that the bike trips’ responses to some weather factors fluctuate dynamically while others maintain at a relatively stable level. It is concluded that a few main influencing factors are adequate to represent the travel patterns. It is observed that dynamic models, with the strength of alleviating multicollinearity, present better forecast performance than classic models. This work can facilitate the decision makers and managers to oversee and optimise travel experience of users in real time.

Index Terms—Dynamic model averaging, time-varying responses, unknown observational variances, weather effects, shared bikes, young people.

I. INTRODUCTION

WITH the booming of shared mobility services in urban cities, people’s travel behaviours have a gradual change and shared bike becomes an ideal alternative for daily trips. Especially, young people’s dependence on car ownership begins to decline and the most direct reason is that efficient

Manuscript received January 1, 2019; revised September 27, 2019 and December 22, 2019; accepted February 10, 2020. Date of publication February 25, 2020; date of current version May 3, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant U1811463 and Grant 51775565, in part by the Fundamental Research Funds for the Central Universities under Grant 18lgpy83, and in part by the Engineering and Physical Sciences Research Council of U.K., under the EPSRC Innovation Fellow scheme under Grant EP/S001956/1. The Associate Editor for this article was A. Hegyi. (*Corresponding author: Ronghui Zhang*)

Guanying Jiang is with the Guangdong Key Laboratory of Intelligent Transportation System, School of Intelligent Systems Engineering, Sun Yat-sen University, Guangzhou 510275, China, and also with the College of Economics, Jinan University, Guangzhou 510632, China (e-mail: gyjiang@stu2017.jnu.edu.cn).

Ronghui Zhang is with the Guangdong Key Laboratory of Intelligent Transportation System, School of Intelligent Systems Engineering, Sun Yat-sen University, Guangzhou 510275, China (e-mail: zhangrh25@mail.sysu.edu.cn).

Xiaobo Qu is with the Department of Architecture and Civil Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden (e-mail: drxiaoboqu@gmail.com).

Dezong Zhao is with the Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough LE11 3TU, U.K. (e-mail: d.zhao@lboro.ac.uk).

Digital Object Identifier 10.1109/TITS.2020.2974930

public transportation and shared mobility meet travel needs [1]. Recent studies concerned with the demand and supply of shared bikes [2], [3], but they seldom discussed the patterns of different cyclist cohorts, such as young and old, annual customers and temporary subscribers. As an instance, New York City (NYC) is a multi-vibrant metropolis where a great number of young people gather for further study or better employment opportunity. Its largest bike sharing system, CitiBike, has attracted numerous studies [4]–[7]. Based on the projected population available from NYC open data in 2015, people aged 15–35 account for 31.32%. However, the CitiBike trip productions at the same age range take up 53% from January to October in 2018. Young people are the main force of shared bike trips. Under this circumstance, improving the travel experience of shared bikes becomes more essential than ever before.

However, owing to completely exposure to the external environment, taking a bike is susceptible to external changes, especially weather conditions [8]. It draws significant attention and interest in measuring the weather-cycling relationship. Though many studies have revealed an apparent relationship between weather factors and bike trips, the association may change over time. For instance, when hanging out in summer, we mainly focus on the highest temperature and sunlight intensity rather than the possibility of snowing. On the other hand, in winter, our focus switches to the opposite. It seems that there are major weather factors during different time periods. In other words, weather factors play different roles over time. Thus, our objective is to address the following research issues: (1) How does the relationship between weather factors and shared bike trips vary over time? (2) What are the main influencing factors and how do they change during different time periods?

For the first problem, multivariate time series model with time-varying parameters can be conducted. It has been extensively investigated and applied, especially in macroeconomics (see also [9]). However, one drawback is that all variables are taken into account, which cannot help solve the second issue. With regards to the second problem, dummy variables and submodeling are commonly used. Unfortunately, they can only model very limited cases, such as four seasons, several months, etc. To measure the model uncertainty, another model, Bayesian Model Averaging (BMA), can be applied. It was firstly proposed by [10] and then comprehensively reviewed by [11]. However, these models cannot fetch time-varying features, so cannot address the first issue. To address two

issues simultaneously, we utilize Dynamic Model Averaging and Dynamic Model Selection in this study, first proposed by [12] and then extended by [9], [13], [14]. DMA and DMS enable varying parameters and shifting submodels, which is helpful to verify all the assumptions.

Our work contributes to the existing studies from the following four ways: (1) We utilize an econometrics framework with considering both parameter varying and model switching to measure shared bike trips' patterns under various weather conditions. (2) For different time periods, we specify a few different major influencing factors of shared bike trips. (3) Dynamic models present a more accurate forecast for shared-bike trips. The assumption of unknown observational variances can further enhance model performance. (4) Our work can be used as a possible early warning system for monitoring real-time bike trips production under changeable weather conditions. It can be extended to other cities with different climate features or other modes of transportation.

The rest of this paper is organized as follows. Section II reviews the research in time-varying weather-cycling patterns whereas Section III introduces two illustrative datasets. Methodology and model structures are provided in Section IV. Section V presents the results and findings, and Section VI reports the conclusions and directions for further research. Detailed calculation flows are displayed in Appendices.

II. LITERATURE REVIEW

Weather factors are critical factors for cycling [15], [16]. Weather variations make more effects on bicyclists than on pedestrians [17] while cycling is more likely and vulnerable to be influenced by weather changes than the motorized modes [8]. Different weather conditions will bring various impacts, such as colder weather, precipitation, and excessive heat in the Washington, DC. [18], a rainy or very humid day in NYC [5], temperature in Toronto [19] and Montreal [20]. In particular, adverse weather conditions affect travel behaviours predominantly and thus some Road Weather Information Systems (RWIS) are exploited to minimize the impacts, reduce the economic output and maintain the public transportation system [21]. Even, extreme weather is the main key to operate and maintain a shared bike system, such as the Pronto band [22].

Research measured the weather-cycling relationship from the temporal dimension. Models used to forecast shared bikes usage include a series of regression models [23], a novel sinusoidal model [16], distributed lag models and multilevel mixed effects model [19]. Classic time series models, like the autoregressive model [8] seasonal ARIMAX model [24], multivariate regression model [25], simple linear model [26] and neurowavelet prediction algorithm [27], have also been conducted to explore this problem.

Instead of discussing time-varying features in one big model, dummy variables and submodeling are utilized to unveil the differences in travel patterns during different time periods. [19], [24] compared different cycling patterns during weekdays and weekends whereas [8] found that weekend cycling is less likely to be influenced by weather. Studies

presented weather-cycling patterns in monthly [4], [6] and seasonal scales [8], even different time periods of the day [4], [6]. [6], [8] took time of the day, day of week, several months and seasons all into account. Other than these, hourly data [7], [17], [18], [23] and daily data [8], [16] are also commonly used. [8], [23] further regarded time factors as predictors.

Time-varying models are often used in travel pattern discovery and transit prediction. [28] proposed a time-varying parameters vector auto-regression model (TVP-VAR) to assist flexibly adjusting the toll scheme over time. As a special case of DMA and DMS, TVP-VAR measures the variation of parameters and thus reflects the uncertainty of variables relationship. Besides, [29] considered time-varying features in spatiotemporal forecast of traffic states. Time series models with time-invariant parameters are also used in traffic flow forecast [30] or pattern discovery [31], but a time-varying analysis framework can provide more detailed patterns and help solve short-term real-time tasks [32]–[34].

Apart from the temporal dimension, studies also incorporate weather impacts as part of explanatory variables to acquire the spatial heterogeneity of shared-bike trips. Generally, spatial econometrics models and panel mixed ordered logit model are utilized for analysis [5], [20], [35]. The impacts of temporal-spatial factors on bike trips, especially weather conditions, have also been revealed [36]. Based on a comparison among 75 Bicycle Sharing Systems (BSS), [15] suggested that changes in weather conditions have a more crucial impact on promoting BSS performance, compared with the increasing number of stations and bikes around.

In this section, we have reviewed the weather-cycling association from both spatial and temporal dimensions. Unveiling the time-varying pattern of this relationship is increasingly popular and highly needed, which is the main problem to be solved in this work.

III. DATA DESCRIPTION

A. CitiBike Data

New York CitiBike historical records are provided per month and cover from July 2013 till now. Every dataset contains fifteen variables, including the start time and stop time of each trip, its duration, the geographical information of the origin and destination (ID, place name, latitude, and longitude) and users' information (ID, year of birth, gender, and user type). According to the paid package, users have to give extra payment if their trip duration exceeds the specific ride time. Thus, we set 135 minutes as an upper limit and remove the trips with an extremely long duration. The minimum time cost is not required since the open data has already limited it to a sensible 60 seconds.

After being preprocessed, the trips production is aggregated per day and is treated as the response variable. We divide CitiBike trips from August 2017 to October 2018 into groups based on demographic attributes: ages, genders and user types. Table I presents basic descriptive statistics of trips production. The definitions of age cohorts are inconsistent in different studies and agencies. Referring to UN Habitat (Youth Fund) definition, we limit young people at a wider age range as

TABLE I
DESCRIPTIVE STATISTICS FOR CITIBIKE TRIPS PRODUCTION

Attribute	Category	N	Pct.	Mean	Sd.	Max	Min
Age cohorts	Young	7497527	0.48	20541.17	8499.87	35310	797
	Middle-aged	6340674	0.41	17371.71	7092.93	30000	883
	Elderly	1647571	0.11	4513.89	1687.41	7674	219
User types	Subscribers	6331182	0.93	23191.14	7977.11	35310	1493
	Customers	474806	0.07	1739.22	1530.69	10677	12
Genders	Male	4846186	0.72	17751.60	6018.73	26362	1235
	Female	1915638	0.28	7016.99	2791.52	11725	264

Note: Bike trips among age cohorts only contain annual subscribers from September 2017 to August 2018. Bike trips for two genders and two types of users only cover young people from February 2018 to October 2018. Data in August 2017 and January 2018 will be used as prior for model computation. Pct. means the proportion of the category in its attribute. Mean denotes daily average bike trips. Sd. denotes standard deviation.

15 to 34 years old (born in 1983-2002). Thus, the age range of middle-aged people (born in 1963-1982, aged 35-54) and that of elderly people (born in 1943-1962, aged 55-74) are also provided. Among age cohorts, young people contribute most to shared bike trip productions. Young people are the main force of CitiBike trips, thus bike trips of young annual subscribers from August 2017 to August 2018 are used most in the following discussions to provide a year-round analysis.

Since the temporary customers before 2018 are not required to fill in their year of birth, we utilize young people's data from January 2018 to October 2018 to explore travel patterns of different user types or genders. In Table I, the proportion of trip productions for annual subscribers, approximate to one, significantly outstrips that of temporary customers. The number of female bike trips accounts for around 40% of male's. A significant gap in travel choices exists among these groups, but their bike trip patterns may not differ since we can only analyze those who actively choose to cycle. Bike trip density of the four seasons is presented in Fig. 1, covered with 2010 census tracts map from NYC open data website.

B. Weather Data

NYC locates in cold temperate zone and belongs to temperate continental humid climate. The four seasons in NYC are distinct and it is perennially affected by humid weather. Also, NYC is greatly affected by two air masses, a warm and humid air mass from the southwest and a cold and dry one from the northwest. These basic climate features can help explain subsequent outcomes. Daily weather data from August 2017 to October 2018 are available from Weather Underground. Since other weather recording stations had different degrees of missing data, we only obtain weather records from Central Park Station as a representation of CitiBike area. There are thirteen weather factors, including temperature (max, min, avg), dew point (max, min, avg), humidity (max, min), wind speed (max, min), pressure (max, min) and precipitation (avg). In this dataset, records from March 19 to 23, 2018 and other two days are missing. Considering the strong storm at the beginning of 2018 did not last till mid-March, we implement spline imputation to fill in the blanks directly. We also limit the wind speed to non-negative value.

Most studies utilize the observations of weather factors available directly [5], [7], [8], [18], [37] or use the mean

TABLE II
DESCRIPTIVE STATISTICS FOR WEATHER FACTORS AT THE AVERAGE LEVEL IN FOUR SEASONS

Variable	Fall		Winter		Spring		Summer	
	Mean	Sd.	Mean	Sd.	Mean	Sd.	Mean	Sd.
Precipitation (in)	0.09	0.34	0.11	0.25	0.15	0.40	0.21	0.46
Temperature (°F)	60.48	12.14	36.06	11.72	52.41	12.93	75.84	5.86
Dew Point (°F)	47.77	14.13	23.46	14.47	36.18	14.77	63.19	7.43
Humidity (%)	65.09	12.94	61.68	14.31	58.98	16.40	68.75	11.79
Wind Speed (mph)	6.34	1.97	7.97	2.30	7.81	2.72	5.35	1.37
Pressure (Hg)	30.08	0.22	30.15	0.24	29.99	0.20	30.01	0.13

Note: Data ranges from September 2017 to August 2018. Sd. denotes standard deviation.

TABLE III
DESCRIPTIVE STATISTICS FOR WEATHER FACTORS AT THE EXTREME LEVEL

Variable	Scale	Mean	Sd.	Variable	Scale	Mean	Sd.
Precipitation	Max	-	-	Humidity	Max	80.95	16.08
	Min	-	-		Min	46.32	15.58
Temperature	Max	63.45	18.80	Wind Speed	Max	13.08	3.76
	Min	49.14	17.72		Min	0.67	1.65
Dew Point	Max	48.87	19.01	Pressure	Max	30.17	0.21
	Min	36.61	20.95		Min	29.95	0.24

Note: Data ranges from September 2017 to August 2018.

ones [15], [22]. Similarly, we obtain weather factors at the average level using the midrange, namely the average of the maximum and minimum observations. The descriptive statistics in Table II indicate that NYC maintains a high relative value for humidity throughout the year. Temperature, dew point and wind speed fluctuate most in spring and least in summer whereas the volatility of precipitation is highest in summer and lowest in winter. Other than the average level, two more weather recording levels are discussed. As in Table III, there is a divergence between maximum and minimum observations, especially for wind speed and humidity. We call it the extreme level, which is likely to provide a more comprehensive reflection of weather conditions. Additionally, bike trips can also be influenced by weather conditions in both current and previous periods [8], [38], thus the lag effects will also be included.

In this work, the necessity of using different scales of weather factors will be discussed in forecast comparisons while the average level will be used in the remaining case studies. To meet the model requirement, some weather factors are transformed in the first order difference while others keep the original series. We have carried out the Augmented Dickey-Fuller test for all factors and they all rejected the hypothesis and were stable with statistically significant. Note that some variables we used in the following parts are differential terms.

IV. METHODOLOGY

A. Model Framework and Forgetting Factors

In this section, we describe how to use DMA and DMS to measure time-varying response of shared bike trips on weather factors. We write them in linear Gaussian state space form as follows. The measurement equation is

$$y_t = x_t^{(k)\top} \beta_t^{(k)} + \varepsilon_t^{(k)} \quad (1)$$

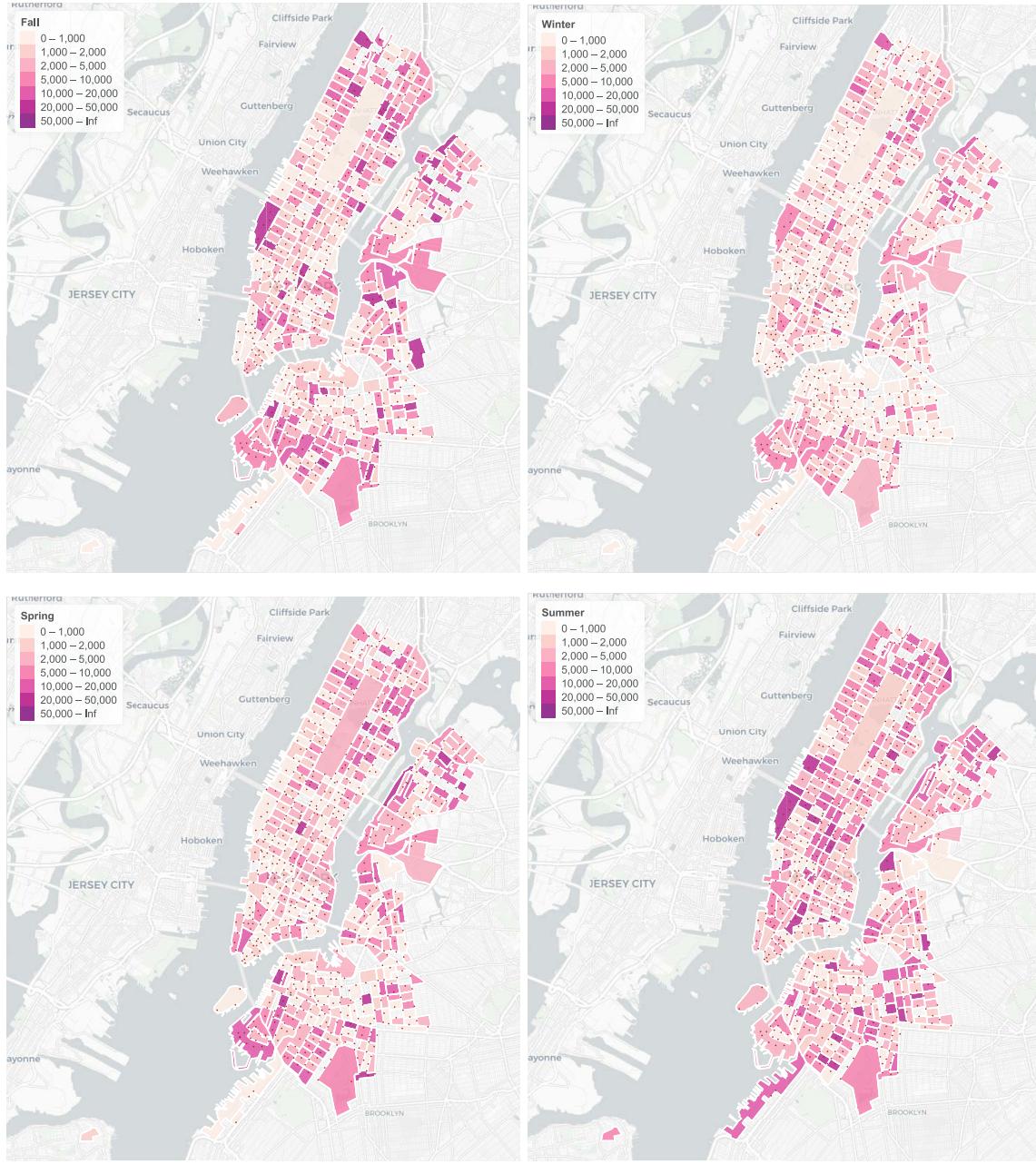


Fig. 1. CitiBike trip productions of young annual subscribers, from September 2017 to August 2018. The colour shades represent bike trip productions of census tracts. The dots are the station sites of CitiBike. It is observed that bike trip productions are mainly in summer and fall and least in winter. It can be assumed that the spatial bias for bike trips and weather conditions is minor enough to be neglected during a short time period.

and the state equation is

$$\beta_t^{(k)} = \beta_{t-1}^{(k)} + \eta_t^{(k)} \quad (2)$$

The superscript $k \in \{1, 2, \dots, K\}$ denotes the k -th alternative model; the superscript T represents matrix transpose; the subscript t means the t -th time stamp in a time series sequence. The error term of the measurement equation $\varepsilon_t^{(k)}$ \sim $N(0, V_t^{(k)})$ and that of the state equation $\eta_t^{(k)}$ \sim $N(0, U_t^{(k)})$; $N(\cdot)$ denotes normal distribution and the abbreviation *ind.* means independent. Besides, $\beta_t^{(k)}$ is a vector of time-varying coefficients; y_t denotes daily bike trip productions; $x_t^{(k)}$ is

a subset of weather factors. If we take six weather factors into consideration, there will be $2^6 - 1$ alternative models. Since we include an intercept in the predictors, 2^6 submodels exist totally. $x_t^{(k)}$ can transit among these $K = 2^6$ alternative models.

Though Markov Chain Monte Carlo Algorithm (MCMC) can be utilized for solving this state-space form model, the massive number of alternative models will bring an evident gain in both computational complexity and time assumption [12], [13]. To simplify it, we follow [12], [13] to embed three forgetting factors into the recursive process.

The first is model switching factor α . Let the posterior inclusion probability of model switching be

$$\begin{aligned}\pi_{t|t-1,k} &= p(L_t = k|Y_{t-1}) \\ &= \sum_{l=1}^K \pi_{t-1|t-1,l} q_{kl}\end{aligned}\quad (3)$$

conditionally on the past $Y_{t-1} = \{y_1, y_2, \dots, y_{t-1}\}$. $\pi_{t|t-1,k}$ denotes the probability that model k is applied to measure the travel pattern y_t at time t . L_t denotes the index of the selected model at time t and $q_{kl} = p(L_t = l|L_{t-1} = k)$ is the transition probability. The dimension of transition matrix is $2^6 \times 2^6$ which brings computational burdens. Reference [12] introduced forgetting factor α to let $\pi_{t|t-1,k}$ rely on its past, namely

$$\pi_{t|t-1,k} = \frac{\pi_{t-1|t-1,k}^\alpha}{\sum_{l=1}^K \pi_{t-1|t-1,l}^\alpha} \quad (4)$$

We sort all alternative models by ascending order of the model size and give each model an index. The model with a larger index includes more factors.

The second is parameter varying factor λ . In Kalman Filter, the initialization is

$$\beta_{t-1}^{(k)}|L_{t-1} = k, Y_{t-1} \sim N(\hat{\beta}_{t-1}^{(k)}, \Sigma_{t-1}^{(k)}) \quad (5)$$

and the prediction is

$$\beta_t^{(k)}|L_t = k, Y_{t-1} \sim N(\hat{\beta}_{t-1}^{(k)}, W_t^{(k)}) \quad (6)$$

in which we are required to update the variance $W_t^{(k)}$ based on $\Sigma_{t-1}^{(k)}$. Given Eq. (2), we have

$$W_t^{(k)} = \Sigma_{t-1}^{(k)} + U_t^{(k)} \quad (7)$$

$U_t^{(k)}$ in Eq. (7) is unknown and needs to be specified by users. It will result in more uncertainty and computation. Reference [12] applied another forgetting factor λ to simplify it:

$$W_t^{(k)} = \lambda^{-1} \Sigma_{t-1}^{(k)} \quad (8)$$

Note that $\beta_t^{(k)}$ may stay on a stable level when few distinctions exist between $W_t^{(k)}$ and $\Sigma_{t-1}^{(k)}$. It means the coefficient $\beta_t^{(k)}$ may be a constant, so does its posterior inclusion probability. In reality, almost all weather factors have a constant period throughout a year (see Fig. 2 and 6).

The last forgetting factor is stochastic volatility factor κ . Reference [13] used an Exponentially Weighted Moving Average (EWMA) form to estimate the error variance in Eq. (1), namely

$$V_t^{(k)} = \kappa V_{t-1}^{(k)} + (1 - \kappa) \hat{\varepsilon}_t^{(k)} \hat{\varepsilon}_t^{(k)T} \quad (9)$$

where $\hat{\varepsilon}_t^{(k)} = y_t - x_t^{(k)T} \beta_{t|t}^{(k)}$. $V_t^{(k)}$ relies on its lag term and current forecast residual. It can not only let $V_t^{(k)}$ shift stably, but also resist drastic fluctuations caused by sudden changes in weather.

These three factors, α, λ, κ , are specified by users. Reference [12] used $\alpha = \lambda = 0.99$ to allow models and parameters shifting smoothly and slowly. Reference [9] gave a sensitivity analysis letting $\alpha, \lambda = 0.95, 0.99, 1$. [13] selected optimal values for λ, κ based on the best DMS forecast. Distinct from

forecasts in macroeconomics, weather variations contain more uncertainties. In this paper, we specify α, λ, κ as 0.95 to fetch more sudden changes and the sensitivity analysis will also be presented.

In addition, Kalman Filter requires prior information for initialization. We specify the posterior inclusion probability as $\pi_{0|0,l} = 1/K, l = 1, 2, \dots, K$. We utilize the first month of the dataset to calculate $\beta_0^{(k)}, \Sigma_0^{(k)}$, following [13]. Detailed calculation process is presented in Appendix A.

B. Unknown Observational Variances

In some applicable scenarios, like small sample size and non-normal observations, the models with the strong hypothesis of normal distribution are not appropriate. Following [39], we discuss dynamic models with unknown observational variances and then extend this case into the DMA and DMS framework.

With the model structure Eq. (1) and Eq. (2), we re-define the error terms of the measurement equation and the state equation:

$$\varepsilon_t^{(k)} \stackrel{\text{ind.}}{\sim} N(0, V_t^{(k)}) \quad (10)$$

$$\eta_t^{(k)} \stackrel{\text{ind.}}{\sim} T_{n_{t-1}}(0, U_t^{(k)}) \quad (11)$$

where $T_n(\cdot)$ denotes the Student's t -distribution with the degree of freedom n . Thus, the distribution of the observational variance is specified as *Gamma* distribution, labeled as $Ga(\cdot)$. Instead of using the former EWMA forgetting form to update the value of $V_t^{(k)}$, we rewrite the corresponding initialization and updating steps as follows:

$$V_{t-1}^{-1(k)}|L_{t-1} = k, Y_{t-1} \sim Ga\left(\frac{n_{t-1}}{2}, \frac{n_{t-1} \hat{V}_{t-1}^{(k)}}{2}\right) \quad (12)$$

$$V_t^{-1(k)}|L_t = k, Y_t \sim Ga\left(\frac{n_t}{2}, \frac{n_t \hat{V}_t^{(k)}}{2}\right) \quad (13)$$

$$\hat{V}_t^{(k)} = \hat{V}_{t-1}^{(k)} + \frac{\hat{V}_{t-1}^{(k)}}{n_t} \left(\frac{\hat{\varepsilon}_t^{(k)} \hat{\varepsilon}_t^{(k)T}}{\hat{V}_{t-1}^{(k)} + x_t^{(k)T} W_t^{(k)} x_t^{(k)}} - 1 \right) \quad (14)$$

In this case, the initialization, prediction and updating of the model parameters shift from *Normal* distribution to Student's t distribution. It directly affects the updating of the response variable y_t :

$$y_t|Y_{t-1} \sim T_{n_{t-1}}(x_t^{(k)T} \hat{\beta}_{t|t-1}^{(k)}, V_t^{(k)} + x_t^{(k)T} W_t^{(k)} x_t^{(k)}) \quad (15)$$

Further, the posterior inclusion probability slightly differs.

$$\pi_{t|t,k} = \frac{\pi_{t|t-1,k} f_k(y_t|Y_{t-1})}{\sum_{l=1}^K \pi_{t|t-1,l} f_l(y_t|Y_{t-1})} \quad (16)$$

Detailed calculation process is presented in Appendix B. The data and codes are provided in this repository: <https://github.com/GuanyingJiang/Dynamic-Model-Averaging-for-ITS>.

C. Measurement Metrics

In this part, we introduce the metrics we used to measure the time-varying bike trip patterns. The first is the posterior

inclusion probability of each weather factor. It is calculated by summing the updating posterior inclusion probabilities $\pi_{t|t,k}$ of the models which include this weather factor. It varies from 0 to 1 and reflects the time-varying significance of each weather factor.

The second is the updating posterior regression coefficients $\beta_t^{(k)} | L_t = k, Y_t$. We do not attempt to present the coefficients of every model alternative, since the number of submodels is enormous and each of them consists of different influencing factors. We obtain the time-varying coefficient range through the maximum, minimum and average of the updating β_t over models that include the corresponding factor.

The third is based on the predictive results of DMA and DMS. DMA is a weighted average of forecasts from all K models and it takes the posterior inclusion probability as weights. DMS is the forecast of the k -th model with the maximum posterior probability (see also Appendix A). DMS consists of the best predictive models for all time periods. The expected model size of DMA in Eq. (17), the size of the best models in DMS, and the weather factors used in DMS over time and their frequency are all utilized to specify the main influencing factors.

$$E(\text{size}_t) = \sum_{k=1}^K \pi_{t|t-1,k} \text{size}^{(k)} \quad (17)$$

The last includes two evaluation metrics, Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), to verify model effectiveness.

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (x_t - \hat{x}_t)^2} \quad (18)$$

$$MAPE = \frac{1}{T} \sum_{t=1}^T \frac{|x_t - \hat{x}_t|}{x_t} \quad (19)$$

Before starting analysing, we tend to specify that DMA and DMS can be regarded as a general form of some other models used to measure the weather-cycling patterns. For instance, if there is only one model structure and K is one, DMA and DMS degenerates into TVP (Time-Varying Parameter) models. If the variance of the error term $\eta_t^{(k)}, U_t^{(k)}$, approximates to zero, then the coefficients $\beta_t^{(k)}$ become constant. In this case, DMA and DMS only switches among K models with time-invariant parameters and can be regarded as static Bayesian Model Averaging. If these two cases occur simultaneously, DMA and DMS becomes a multiple linear regression model. All these models are special cases of DMA and DMS, which means we can fetch more information through the model framework of DMA and DMS.

V. EMPIRICAL RESULTS AND DISCUSSIONS

Section V-A takes NYC CitiBike trips as an instance to present how the weather impacts on cycling vary over time, using the posterior inclusion probability of weather factors and the posterior mean of coefficients. Section V-B helps understand the shrinkage of the predictive model structure through the expected model size, the best predictive model

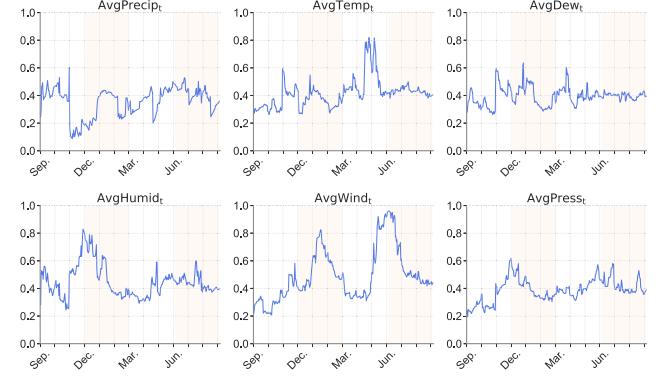


Fig. 2. CitiBike trip productions of young people: posterior inclusion probability of weather factors at the average level over time. Two background colors are adopted to distinguish four seasons' patterns (fall, winter, spring and summer in turn from left to right for each subfigure). The higher the posterior inclusion probability, the more likely the factor works as a predictor.

and its composition. Section V-C illustrates the model forecast performance between the proposed models and the classic ones whereas Section V-D discusses the main edge of DMA and DMS to solve the weather-cycling issue.

A. Time-Varying Impacts and Variable Coefficients

Fig. 2 presents the posterior inclusion probabilities of six weather factors in a whole year. It can be seen that the response of young people's bike trips on weather factors have time-varying patterns. The most volatile factors are temperature, humidity and wind speed while other factors just fluctuate around a certain level all year round. Specifically, the influence of average temperature reaches its peak mainly in spring whereas average humidity maintains a high inclusion level during the transition period of fall and winter. The two essential impact periods of average wind speed are exactly in parallel with the periods of two air masses in NYC. Though the significant impacts of these weather factors have been discussed [5], [7], the findings of how they vary over time are limited.

To better distinguish the time-varying patterns, we adopt several background colors in Fig. 2. Almost all weather factors fluctuate most in winter and spring and least in fall and summer. From fall to summer, the weather factors with the highest posterior probability shift from average humidity to average wind speed, and then from average temperature to average wind speed. Briefly, only a few main influencing factors are interlaced to affect the shared-bike trips.

Though the posterior inclusion probability illustrates the necessity of including the variable into predictors, it cannot reflect the size of its effect. In Fig. 3, the magnitude of coefficient changes is close to that of posterior inclusion probability, but the trend direction may differ, like the average wind speed. The coefficient ranges of temperature and dew point fluctuate most in winter while that of precipitation and humidity have evident fluctuations in spring. The effects of other factors basically distribute within a small range. Furthermore, a factor with a low posterior inclusion probability may also have a large impact on shared-bike trips, such as precipitation and pressure.

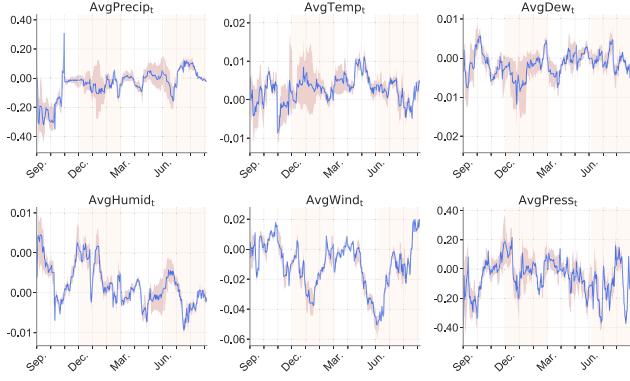


Fig. 3. Posterior means and intervals of regression coefficients over time under DMA.

TABLE IV

THE INFLUENCING FACTORS CORRESPONDING TO THE HIGH-FREQUENCY
BEST MODELS AT THE AVERAGE LEVEL

Model	Freq.	AvgPrecip	AvgTemp	AvgDew	AvgHumid	AvgWind	AvgPress
2	15						
4	16						
5	71						
14	14						
16	11						
20	23						
21	18						
38	16						
64	140 (Intercept)						
		0	29	27	66	119	41

Note: The first two columns are the model index and its selected frequency. The rest columns with ticks are the included variables of the selected model. In this work, every model includes an intercept.

B. Main Influencing Factors and the Selected Models

The characteristics of high-frequently used predictors and model structures are reported herein. For one thing, there are totally 2^6 alternative models, but only 30 models are selected and we show those with more than 10 times usage in DMS in Table IV. The univariate and bivariate models are the most high frequency used. It indicates that only a few models with a few factors are adequate for measuring young people's bike trip patterns. The predictors used at the highest frequency are also in line with that in Fig. 2. Though the main climate feature in NYC is serious humidity, it seems wind speed is more effective in measuring bike trips.

For another, the expected model size and the index of the best predictive model are presented in Fig. 4 and Fig. 5. The size of the best predictive model of DMS gathers at 2 to 3 while the expected model size of DMA fluctuates around 3.5. The full variable model is not used at any point in time. The least influencing factors are used in fall, early spring and late summer while the most is mainly in winter. The expected model size maintains its level in every middle month of the season but abruptly shifts during the transition period of seasons. Additionally, the model index shows a bit more fluctuations than the model size. It is mainly because several model structures have the same number of predictors. These two metrics present similar patterns overall.

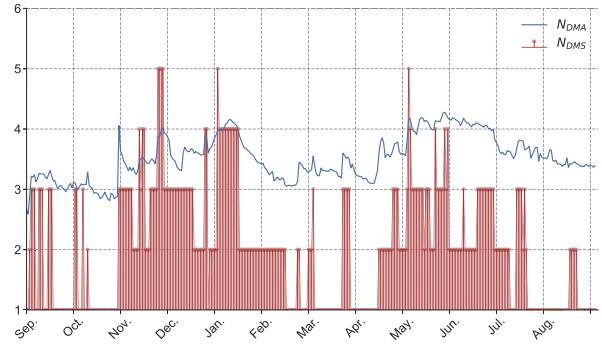


Fig. 4. The expected model size over time at the average level.

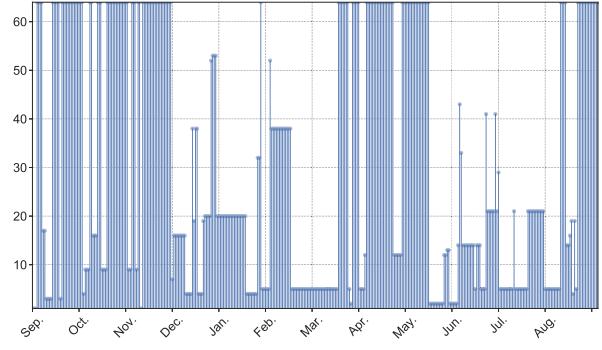


Fig. 5. The time-varying best measurement models at the average level.

TABLE V

A FORECAST COMPARISON AMONG VARIOUS GROUPS OF PREDICTORS

	N	MAPE(A)	RMSE(A)	MAPE(S)	RMSE(S)
The average level	6	0.1688	0.1438	0.1673	0.1401
The average level*	6	0.1690	0.1439	0.1658	0.1391
Average & Weekdays	7	0.0978	0.0970	0.0960	0.0921
Extreme (Max. only)	6	0.1694	0.1443	0.1649	0.1384
Extreme (Min. only)	6	0.1707	0.1433	0.1686	0.1414
The extreme level	11	0.1748	0.1483	0.1686	0.1414
Extreme & Weekdays	12	0.1005	0.0993	0.1010	0.0988
The lag effect	12	0.1604	0.1424	0.1510	0.1387
Lag & Weekdays	13	0.0965	0.0966	0.0933	0.0902

Note: N is the number of predictors; (A) and (S) denote the results of DMA and DMS separately; the asterisk indicates the case of unknown variances; ‘& Weekdays’ means adding a typical time factor into the average-level, extreme-level and lagging models.

C. Forecast Performance

Table V presents the forecast performance of dynamic models with various predictors. For the average level, DMA with known observational variances is slightly better than that with unknown variances while DMS has an opposite performance. The inclusion of the time factor improves the forecast precision nearly by half for all cases. As for the extreme level, the maximum observations provide the best forecast result while the integration of maximum and minimum factors does not perform well as expected. The accuracy of the lagging model is the highest among all alternatives without weekdays factor. However, a large number of predictors will increase computation burden significantly. Overall, the average-level

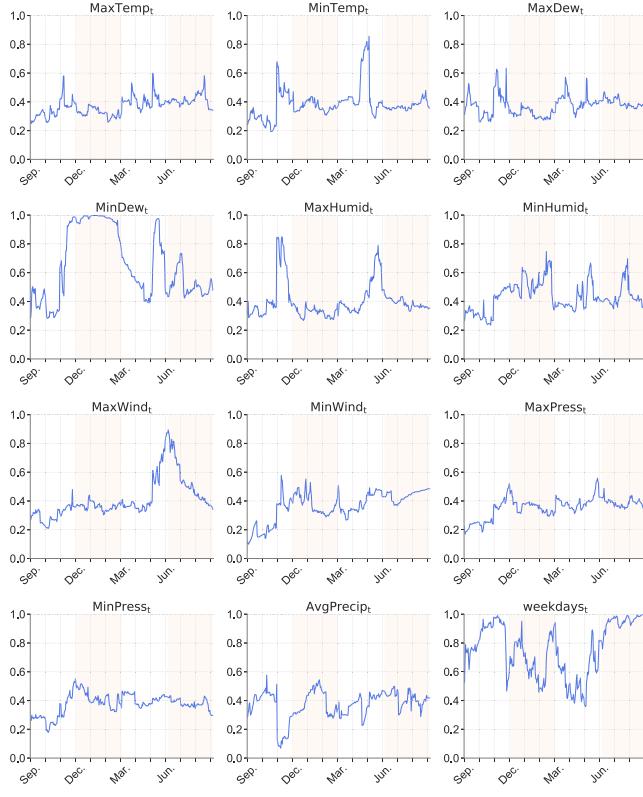


Fig. 6. CitiBike trip productions of young people: posterior inclusion probability of six weather factors at the extreme level.

dynamic models with unknown observational variances or a time factor are more suitable for shared bike trips forecast.

To illustrate how the time factor is included, Fig. 6 shows the case ‘Extreme & Weekdays’, which gains the most improvement after including the Weekdays factor. The inclusion probability of Weekdays exceeds 0.8 almost throughout summer and fall, showing a predominant role in bike trips forecast. Owing to NYC’s climate, the maximum humidity in fall, minimum dew point in winter, minimum temperature in spring and maximum wind speed in summer are also necessary. Additionally, though there are totally 2^{12} submodels in this case, the numbers of high frequency predictive model is 66, without growing exponentially. The expected model size of DMA fluctuates around 6.5 while that of DMS is 3 to 5 mainly.

Table VI illustrates the enhancement of the proposed dynamic models, compared with various classic models. Firstly, the sensitivity analysis of the forgetting rates indicates the increase of λ (parameter varying factor) weakens the model while a smoother α (model switching factor) enhances it instead. Secondly, TVP (a single full model with Time-Varying Parameters) and DM (Dynamic Models with shifting submodels but constant parameters) are carried out for comparison. Unfortunately, only TVP ($\lambda = 0.95$) guarantees the accuracy while the rest decreases forecast precisions. The third group is BMA and BMS models with all the forgetting factors fixed at 1. Even though BMA and BMS include all alternative models for forecast, they perform the worst among all. Therefore, allowing a certain degree of model switching

TABLE VI
A FORECAST COMPARISON AMONG VARIOUS MODEL STRUCTURES

	α	λ	κ	MAPE(A)	RMSE(A)	MAPE(S)	RMSE(S)
DMA/DMS	0.95	0.95	0.95	0.1688	0.1438	0.1673	0.1401
	0.99	0.95	0.95	0.1657	0.1418	0.1623	0.1399
	0.95	0.99	0.95	0.2021	0.1769	0.2122	0.1868
	0.95	0.95	0.99	0.1686	0.1435	0.1654	0.1387
TVP/DM	1	0.95	1	0.1786	0.1562	0.1786	0.1562
	1	0.99	1	0.2223	0.1972	0.2215	0.1965
	0.95	1	1	0.2300	0.2139	0.2303	0.2124
	0.99	1	1	0.2360	0.2190	0.2368	0.2196
BMA/BMS	1	1	1	0.2386	0.2215	0.2388	0.2216

Note: α , λ , κ are the forgetting factors for model switching, parameter varying and stochastic volatility.

TABLE VII
THE RATIOS OF FORECAST INDICATORS BETWEEN MODELS WITH KNOWN AND UNKNOWN OBSERVATIONAL VARIANCES IN SPECIAL DATA CASES

	Skewness	Kurtosis	MAPE(A)	RMSE(A)	MAPE(S)	RMSE(S)
Fall	-0.34	4.70	1.0981	1.0926	1.1015	1.0876
Winter	0.78	6.35	1.0047	1.0059	0.9728	0.9960
Spring	-0.50	3.06	0.9915	1.0151	1.0181	1.0222
Summer	0.32	6.77	0.9911	0.9931	1.0182	0.9968

Note: The first two columns denote the skewness and kurtosis of the prediction error $\hat{\varepsilon}_t^{DMA}$. The skewness and kurtosis of standard normal distribution are 0 and 3, respectively. The size of each sample dataset is narrowed down to three months (one season). The one-month data ahead of each season is used to initialize the model.

and parameter varying does bring a better model performance, and the reason behind will be discussed in Section V-D.

Table VII presents the ratios of forecast indicators between dynamic models and those with unknown observational variances. Seasonal datasets, with smaller sample sizes and various distributed features, are utilized for comparison. Datasets in Fall and Spring, whose error distribution approximates to normal distribution, show a superior forecast result in unknown variances setting. The forecasts are comparable for Winter and Summer datasets, who have evidently fat-tailed trends. Thus, dynamic models with relaxing hypotheses perform better in small-sampled and approximately-normal data.

D. Variable Correlation and Model Validity

In this study, the predictors, mainly weather factors, will interact with each other. Few studies are conducted from this perspective only (also refer to Section III-B). The main reason is that the full model with strongly correlated variables may lead to model invalidity. In classic linear models, the predictors are assumed to be mutually independent [40]. If not, the variance of the estimated coefficients $Var(\beta_t)$ increases, and it leads to multicollinearity and an unreliable or invalid model.

Fig. 7 presents bivariate correlation coefficients heatmap based on the predictors data in Section V-A. Those strongly correlated bivariates maintain the correlation in some time periods rather than throughout the whole process, such as temperature-dew point from September to February, dew point-pressure from February to June. Also, they neither coexist in the best predictive models nor present high posterior

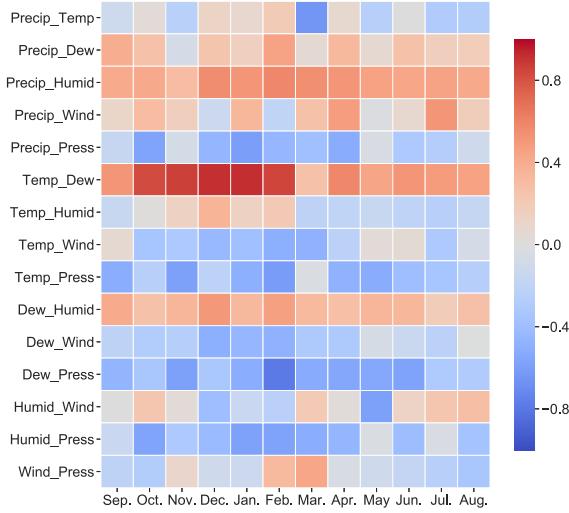


Fig. 7. Time-varying correlation coefficients of the weather factors. The vertical axis represents the abbreviation of different bivariate pairs.

inclusion probabilities simultaneously (refer to Section V-A and V-B). One possible reason is that those models have already failed due to multicollinearity. Their forecast precisions are poor and so do their posterior inclusion probabilities.

With the advantage of forgetting factors, the failure submodels in DMA and DMS can maintain a minor inclusion probability (refer to Eq. (4)). The recursive forecast process also enables to alleviate the effect of the weakened predictions. Besides, we cannot directly eliminate those correlated but not completely replaceable predictors. It can not only spoil the meaning of all weather impacts, but also hinder a comprehensive understanding for decision-makers. In sum, dynamic models achieve a balance in the trade-off between forecast precision and model validity, through enabling both model structures and parameters to vary over time.

VI. CONCLUSION

The shared-bike gains its popularity in the near decades and becomes the main transit alternative in many urban cities. However, too much exposure to the outdoor environment makes cyclists vulnerable to environmental changes, especially weather conditions. This paper proposes various dynamic models to reveal the time-varying weather-cycling relationship. Transportation managers and policy makers can optimize cycling condition and enhance the travel experience of shared-bike users with the help of our work. Even, they can develop response measures and propose travel alternatives for acute weather changes.

Taking NYC CitiBike as an instance, we illustrate why and how to utilize these novel models through several measurement metrics. We discuss dynamic models with known and unknown observational variances respectively and make a comparison of forecast performance with classic models for verification. Besides, we analyze how dynamic models deal with a common problem of linear models: multicollinearity.

The possible directions for further research are listed. First, our work introduces the dynamic model strategy into ITS

research through a basic issue, the weather-cycling relationship. Numerous studies in ITS involve temporal or even time-varying characteristics, like [28]. The proposed models may inspire related research to avoid some model limitations. Second, we explore cycling patterns mainly from the temporal dimension in this work, without including the spatial features. A thoroughly-considered multivariate dynamic model may help discover the spatial-temporal transit patterns. Third, a real-time monitoring system can be developed since the proposed models can start to measure, update and predict the cycling production at any time. Also, our work can be integrated into the existing models for a better transit forecast. Fourth, referring to Table I and Section III-A, cyclists with different demographic attributes may hinder various transit characteristics. More individual information is provided since 2018, thus the underlying patterns are worth exploring.

APPENDIX A A BRIEF FLOW OF DMA AND DMS

1) Initialization

Given the initial values $\beta_0^{(k)}, \Sigma_0^{(k)}, \pi_{0|0,k}$ for $k = 1, 2, \dots, K$

$$\beta_{t-1}^{(k)}|L_{t-1} = k, Y_{t-1} \sim N(\hat{\beta}_{t-1}^{(k)}, \Sigma_{t-1}^{(k)})$$

2) Prediction equations

For model switching

$$\pi_{t|t-1,k} = \frac{\pi_{t-1|t-1,k}^a + c}{\sum_{l=1}^K \pi_{t-1|t-1,l}^a + c}$$

for $k = 1, 2, \dots, K$. c , for adjustment, is slightly larger than 0.

For parameter varying

$$W_t^{(k)} = \lambda^{-1} \Sigma_{t-1}^{(k)} \\ \beta_t^{(k)}|L_t = k, Y_{t-1} \sim N(\hat{\beta}_{t-1}^{(k)}, W_t^{(k)})$$

3) One-step ahead forecasting

$$\begin{aligned} \hat{y}_t^{DMA} &= \sum_{l=1}^K \pi_{t|t-1,l} \hat{y}_t^{(l)} \\ &= \sum_{l=1}^K \pi_{t|t-1,l} x_t^{(l)T} \hat{\beta}_{t|t-1}^{(l)} \\ \hat{y}_t^{DMS} &= x_t^{(k^*)T} \hat{\beta}_{t|t-1}^{(k^*)} \\ k^* &= \arg \max \pi_{t|t-1,k} \end{aligned}$$

4) Updating equations

$$\begin{aligned} \hat{e}_t^{(k)} &= y_t - x_t^{(k)T} \hat{\beta}_{t|t-1}^{(k)} \\ V_t^{(k)} &= \kappa V_{t-1}^{(k)} + (1 - \kappa) \hat{e}_t^{(k)} \hat{e}_t^{(k)T} \end{aligned}$$

For model switching

$$\begin{aligned} y_t|Y_{t-1} &\sim N(x_t^{(k)T} \hat{\beta}_{t|t-1}^{(k)}, V_t^{(k)} + x_t^{(k)T} W_t^{(k)} x_t^{(k)}) \\ \pi_{t|t,k} &= \frac{\pi_{t|t-1,k} f_k(y_t|Y_{t-1})}{\sum_{l=1}^K \pi_{t|t-1,l} f_l(y_t|Y_{t-1})} \end{aligned}$$

where $f_k(y_t|Y_{t-1})$ is the probability density function of $y_t|Y_{t-1}$.

For parameter varying

$$\begin{aligned}\hat{\beta}_t^{(k)} &= \hat{\beta}_{t-1}^{(k)} + W_t^{(k)} x_t^{(k)T} (V_t^{(k)} + x_t^{(k)T} W_t^{(k)} x_t^{(k)})^{-1} \hat{\varepsilon}_t^{(k)} \\ \Sigma_t^{(k)} &= W_t^{(k)} - W_t^{(k)} x_t^{(k)} (V_t^{(k)} + x_t^{(k)T} W_t^{(k)} x_t^{(k)})^{-1} x_t^{(k)T} W_t^{(k)} \\ \beta_t^{(k)} | L_t = k, Y_t &\sim N(\hat{\beta}_t^{(k)}, \Sigma_t^{(k)})\end{aligned}$$

For $t = 1, 2, \dots, T$, iterate Step 2 to 4 to fetch all $\hat{\beta}_t^{(k)}, \pi_{t|t,k}$, and $x_t^{(k*)}, k = 1, 2, \dots, K$.

APPENDIX B DMA AND DMS WITH UNKNOWN OBSERVATIONAL VARIANCES

1) Initialization

Given the initial values $\beta_0^{(k)}, \Sigma_0^{(k)}, \pi_{0|0,k}$ for $k = 1, 2, \dots, K$

$$\begin{aligned}\beta_{t-1}^{(k)} | L_{t-1} = k, Y_{t-1} &\sim T_{n_{t-1}}(\hat{\beta}_{t-1}^{(k)}, \Sigma_{t-1}^{(k)}) \\ V_{t-1}^{(k)} | L_{t-1} = k, Y_{t-1} &\sim Ga\left(\frac{n_{t-1}}{2}, \frac{n_{t-1} \hat{V}_{t-1}^{(k)}}{2}\right)\end{aligned}$$

2) Prediction equations

For model switching

$$\pi_{t|t-1,k} = \frac{\pi_{t-1|t-1,k}^\alpha + c}{\sum_{l=1}^K \pi_{t-1|t-1,l}^\alpha + c}$$

for $k = 1, 2, \dots, K$. c , for adjustment, is slightly larger than 0.

For parameter varying

$$\begin{aligned}W_t^{(k)} &= \lambda^{-1} \Sigma_{t-1}^{(k)} \\ \beta_t^{(k)} | L_t = k, Y_{t-1} &\sim T_{n_{t-1}}(\hat{\beta}_{t-1}^{(k)}, W_t^{(k)})\end{aligned}$$

3) One-step ahead forecasting

$$\begin{aligned}\hat{y}_t^{DMA} &= \sum_{l=1}^K \pi_{t|t-1,l} \hat{y}_t^{(l)} \\ &= \sum_{l=1}^K \pi_{t|t-1,l} x_t^{(l)T} \hat{\beta}_{t|t-1}^{(l)} \\ \hat{y}_t^{DMS} &= x_t^{(k*)T} \hat{\beta}_{t|t-1}^{(k*)} \\ k^* &= \arg \max \pi_{t|t-1,k}\end{aligned}$$

4) Updating equations

$$\begin{aligned}\hat{\varepsilon}_t^{(k)} &= y_t - x_t^{(k)T} \hat{\beta}_{t|t-1}^{(k)} \\ V_t^{-1(k)} | L_t = k, Y_t &\sim Ga\left(\frac{n_t}{2}, \frac{n_t \hat{V}_t^{(k)}}{2}\right)\end{aligned}$$

where

$$\begin{aligned}n_t &= n_{t-1} + 1 \\ M_t &= x_t^{(k)T} W_t^{(k)} x_t^{(k)} \\ \hat{V}_t^{(k)} &= \hat{V}_{t-1}^{(k)} + \frac{\hat{V}_{t-1}^{(k)}}{n_t} \left(\frac{\hat{\varepsilon}_t^{(k)} \hat{\varepsilon}_t^{(k)T}}{\hat{V}_{t-1}^{(k)} + M_t} - 1 \right)\end{aligned}$$

For model switching

$$\begin{aligned}y_t | Y_{t-1} &\sim T_{n_{t-1}}(x_t^{(k)T} \hat{\beta}_{t|t-1}^{(k)}, V_t^{(k)} + M_t) \\ \pi_{t|t,k} &= \frac{\pi_{t|t-1,k} f_k(y_t | Y_{t-1})}{\sum_{l=1}^K \pi_{t|t-1,l} f_l(y_t | Y_{t-1})}\end{aligned}$$

where $f_k(y_t | Y_{t-1})$ is the probability density function of $y_t | Y_{t-1}$.

For parameter varying

$$\begin{aligned}\hat{\beta}_t^{(k)} &= \hat{\beta}_{t-1}^{(k)} + W_t^{(k)} x_t^{(k)T} (V_t^{(k)} + M_t)^{-1} \hat{\varepsilon}_t^{(k)} \\ \Sigma_t^{(k)} &= W_t^{(k)} - W_t^{(k)} x_t^{(k)} (V_t^{(k)} + M_t)^{-1} x_t^{(k)T} W_t^{(k)} \\ \beta_t^{(k)} | L_t = k, Y_t &\sim T_{n_t}(\hat{\beta}_t^{(k)}, \Sigma_t^{(k)})\end{aligned}$$

For $t = 1, 2, \dots, T$, iterate Step 2 to 4 to fetch all $\hat{\beta}_t^{(k)}, \pi_{t|t,k}$, and $x_t^{(k*)}, k = 1, 2, \dots, K$.

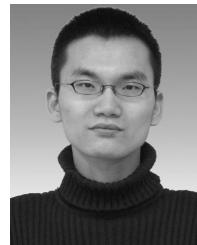
REFERENCES

- [1] J. Zhang, *Life-Oriented Behavioral Research for Urban Policy*. Berlin, Germany: Springer, 2017.
- [2] F. Huang, S. Qiao, J. Peng, and B. Guo, "A bimodal Gaussian inhomogeneous Poisson algorithm for bike number prediction in a bike-sharing system," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 8, pp. 2848–2857, Aug. 2019.
- [3] G. Tang, S. Keshav, L. Golab, and K. Wu, "Bikeshare pool sizing for Bike-and-Ride multimodal transit," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 7, pp. 2279–2289, Jul. 2018.
- [4] R. B. Noland, M. J. Smart, and Z. Guo, "Bikeshare trip generation in New York city," *Transp. Res. A, Policy Pract.*, vol. 94, pp. 164–181, Dec. 2016.
- [5] A. Faghih-Imani and N. Eluru, "Incorporating the impact of spatio-temporal interactions on bicycle sharing system demand: A case study of New York CitiBike system," *J. Transp. Geography*, vol. 54, pp. 218–227, Jun. 2016.
- [6] A. Faghih-Imani, S. Anowar, E. J. Miller, and N. Eluru, "Hail a cab or ride a bike? A travel time comparison of taxi and bicycle-sharing systems in New York city," *Transp. Res. A, Policy Pract.*, vol. 101, pp. 11–21, Jul. 2017.
- [7] K. Wang, G. Akar, and Y.-J. Chen, "Bike sharing differences among millennials, Gen Xers, and baby boomers: Lessons learnt from New York City's bike share," *Transp. Res. A, Policy Pract.*, vol. 116, pp. 1–14, Oct. 2018.
- [8] J. Zhao, J. Wang, Z. Xing, X. Luan, and Y. Jiang, "Weather and cycling: Mining big data to have an in-depth understanding of the association of weather variability with cycling on an off-road trail and an on-road bike lane," *Transp. Res. A, Policy Pract.*, vol. 111, pp. 119–135, May 2018.
- [9] G. Koop and D. Korobilis, "Forecasting inflation using dynamic model averaging," *Int. Econ. Rev.*, vol. 53, no. 3, pp. 867–886, 2012.
- [10] E. E. Leamer, *Specification Searches: Ad Hoc Inference With Nonexperimental Data*, vol. 53. Hoboken, NJ, USA: Wiley, 1978.
- [11] T. M. Fragoso, W. Bertoli, and F. Louzada, "Bayesian model averaging: A systematic review and conceptual classification," *Int. Stat. Rev.*, vol. 86, no. 1, pp. 1–28, Dec. 2017.
- [12] A. E. Raftery, M. Kárný, and P. Ettler, "Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill," *Technometrics*, vol. 52, no. 1, pp. 52–66, Feb. 2010.
- [13] G. Koop and D. Korobilis, "Large time-varying parameter VARs," *J. Econometrics*, vol. 177, no. 2, pp. 185–198, Dec. 2013.
- [14] L. Onorante and A. E. Raftery, "Dynamic model averaging in large model spaces using dynamic Occam's window," *Eur. Econ. Rev.*, vol. 81, pp. 2–14, Jan. 2016.
- [15] C. Médard de Chardon, G. Caruso, and I. Thomas, "Bicycle sharing system 'success' determinants," *Transp. Res. A, Policy Pract.*, vol. 100, pp. 202–214, Jun. 2017.
- [16] N. Fournier, E. Christofa, and M. A. Knodler, "A sinusoidal model for seasonal bicycle demand estimation," *Transp. Res. D, Transp. Environ.*, vol. 50, pp. 154–169, Jan. 2017.
- [17] A. Ermagun, G. Lindsey, and T. H. Loh, "Urban trails and demand response to weather variations," *Transp. Res. D, Transp. Environ.*, vol. 63, pp. 404–420, Aug. 2018.
- [18] K. Gebhart and R. B. Noland, "The impact of weather conditions on bikeshare trips in washington, DC," *Transportation*, vol. 41, no. 6, pp. 1205–1225, Aug. 2014.
- [19] W. El-Assi, M. S. Mahmoud, and K. N. Habib, "Effects of built environment and weather on bike sharing demand: A station level analysis of commercial bike sharing in toronto," *Transportation*, vol. 44, no. 3, pp. 589–613, Dec. 2015.
- [20] F. Reynaud, A. Faghih-Imani, and N. Eluru, "Modelling bicycle availability in bicycle sharing systems: A case study from montreal," *Sustain. Cities Soc.*, vol. 43, pp. 32–40, Nov. 2018.

- [21] K. C. Dey, A. Mishra, and M. Chowdhury, "Potential of intelligent transportation systems in mitigating adverse weather impacts on road mobility: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1107–1119, Jun. 2015.
- [22] F. Sun, P. Chen, and J. Jiao, "Promoting public bike-sharing: A lesson from the unsuccessful pronto system," *Transp. Res. D, Transp. Environ.*, vol. 63, pp. 533–547, Aug. 2018.
- [23] R. Giot and R. Cherrier, "Predicting bikeshare system usage up to one day ahead," in *Proc. IEEE Symp. Comput. Intell. Vehicles Transp. Syst. (CIVTS)*, Dec. 2014, pp. 22–29.
- [24] S. Tao, J. Corcoran, F. Rowe, and M. Hickman, "To travel or not to travel: 'Weather' is the question. Modelling the effect of local weather conditions on bus ridership," *Transp. Res. C, Emerg. Technol.*, vol. 86, pp. 147–167, Jan. 2018.
- [25] M. Zhou, D. Wang, Q. Li, Y. Yue, W. Tu, and R. Cao, "Impacts of weather on public transport ridership: Results from mining data from different sources," *Transp. Res. C, Emerg. Technol.*, vol. 75, pp. 17–29, Feb. 2017.
- [26] M. Mesbah, J. Lin, and G. Currie, "'Weather' transit is reliable? Using AVL data to explore tram performance in melbourne, australia," *J. Traffic Transp. Eng. (English Ed.)*, vol. 2, no. 3, pp. 125–135, Jun. 2015.
- [27] S. Dunne and B. Ghosh, "Weather adaptive traffic prediction using neurowavelet models," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 1, pp. 370–379, Mar. 2013.
- [28] X. Ma, S. Sun, X. C. Liu, C. Ding, Z. Chen, and Y. Wang, "A time-varying parameters vector auto-regression model to disentangle the time varying effects between drivers' responses and tolling on high occupancy toll facilities," *Transp. Res. C, Emerg. Technol.*, vol. 88, pp. 208–226, Mar. 2018.
- [29] P. Cai, Y. Wang, G. Lu, P. Chen, C. Ding, and J. Sun, "A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting," *Transp. Res. C, Emerg. Technol.*, vol. 62, pp. 21–34, Jan. 2016.
- [30] M. Lippi, M. Bertini, and P. Frasconi, "Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 871–882, Jun. 2013.
- [31] C. Ding, J. Duan, Y. Zhang, X. Wu, and G. Yu, "Using an ARIMA-GARCH modeling approach to improve subway short-term ridership forecasting accounting for dynamic volatility," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 4, pp. 1054–1064, Apr. 2018.
- [32] P. Duan, G. Mao, W. Liang, and D. Zhang, "A unified spatio-temporal model for short-term traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 9, pp. 3212–3223, Sep. 2019.
- [33] Q. Ge, B. Ciuffo, and M. Menendez, "Comprehensive approach for the sensitivity analysis of high-dimensional and computationally expensive traffic simulation models," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2422, no. 1, pp. 121–130, Jan. 2014.
- [34] Y. Wang, P. Coppola, A. Tzimitis, A. Messmer, M. Papageorgiou, and A. Nuzzolo, "Real-time freeway network traffic surveillance: Large-scale field-testing results in southern Italy," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 548–562, Jun. 2011.
- [35] C. Liu, Y. O. Susilo, and A. Karlström, "Measuring the impacts of weather variability on home-based trip chaining behaviour: A focus on spatial heterogeneity," *Transportation*, vol. 43, no. 5, pp. 843–867, May 2015.
- [36] C. McBain and B. Caulfield, "An analysis of the factors influencing journey time variation in the cork public bike system," *Sustain. Cities Soc.*, vol. 42, pp. 641–649, Oct. 2018.
- [37] S. Saneinejad, M. J. Roorda, and C. Kennedy, "Modelling the impact of weather conditions on active transportation travel behaviour," *Transp. Res. D, Transp. Environ.*, vol. 17, no. 2, pp. 129–137, Mar. 2012.
- [38] M. Mesbah, M. Luy, and G. Currie, "Investigating the lagged effect of weather parameters on travel time reliability," in *Proc. 15th Sustain. City*, Sep. 2014, pp. 795–801.
- [39] M. West and J. Harrison, *Bayesian Forecasting and Dynamic Models*. Berlin, Germany: Springer, 2006.
- [40] J. F. Monahan, *A Primer Linear Models*. Boca Raton, FL, USA: CRC Press, 2008.



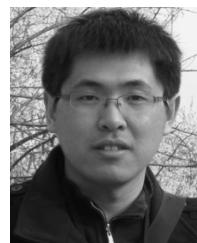
Guanying Jiang received the B.S. degree from Guangzhou University, Guangzhou, China, in 2017. She is currently pursuing the M.Ec. degree with Jinan University, Guangzhou. She is also a Visiting Research Assistant with the Guangdong Key Laboratory of Intelligent Transportation System, School of intelligent systems engineering, Sun Yat-sen University, Guangzhou, China. Her research interests focus on large-scale transportation data and intelligent transportation systems.



Ronghui Zhang received the B.Sc. (Eng.) degree from the Department of Automation Science and Electrical Engineering, Hebei University, Baoding, China, in 2003, the M.S. degree in vehicle application engineering from Jilin University, Changchun, China, in 2006, and the Ph.D. (Eng.) degree in mechanical and electrical engineering from the Changchun Institute of Optics, Fine Mechanics and Physics, The Chinese Academy of Sciences, Changchun, in 2009. After finishing his post-doctoral research work at INRIA, Paris, France, in February 2011, he is currently an Associate Professor with the Guangdong Key Laboratory of Intelligent Transportation System, School of intelligent systems engineering, Sun Yat-sen University, Guangzhou, China. He has published more than 20 articles in international journals. His current research interests include computer vision, intelligent control, and ITS.



Xiaobo Qu received the B.Eng. from Jilin University, China, the M.Eng. degree from Tsinghua University, Beijing, China, and the Ph.D. degree from the National University of Singapore. He is currently a Professor and Chair in urban mobility systems with the Chalmers University of Technology, Sweden. His research is focused on traffic flow theory, intelligent transportation systems, and urban mobility modelling.



Dezong Zhao (Senior Member, IEEE) received the B.S. and M.S. degrees from the School of Control Science and Engineering, Shandong University, Jinan, China, in 2003 and 2006, respectively, and the Ph.D. degree from the Department of Automation, Tsinghua University, Beijing, China, in 2010, all in control science and engineering. Since 2017, he has been a Lecturer in intelligent systems with the Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough, U.K. Since 2018, he has been an EPSRC-UKRI Innovation Fellow in robotics and artificial intelligence systems. His research interests include connected and autonomous vehicles, low carbon vehicles, machine learning, and nonlinear control theory and applications. He is a fellow of the Higher Education Academy and a recipient of the Excellence 100 Campaign of Loughborough University.