Forecasting intraday particle number size distribution: A functional time series approach

Han Lin Shang
Department of Actuarial Studies and Business Analytics
Macquarie University

Israel Martinez Hernandez
School of Mathematical Sciences
Lancaster University

Abstract

Particulate matter data now include various particle sizes, which often manifest as a collection of curves observed sequentially over time. When considering 51 distinct particle sizes, these curves form a high-dimensional functional time series observed over equally spaced and densely sampled grids. While high dimensionality poses statistical challenges due to the curse of dimensionality, it also offers a rich source of information that enables detailed analysis of temporal variation across short time intervals for all particle sizes. To model this complexity, we propose a multilevel functional time series framework incorporating a functional factor model to facilitate one-day-ahead forecasting. To quantify forecast uncertainty, we develop a calibration approach and a split conformal prediction approach to construct prediction intervals. Both approaches are designed to minimise the absolute difference between empirical and nominal coverage probabilities using a validation dataset. Furthermore, to improve forecast accuracy as new intraday data become available, we implement dynamic updating techniques for point and interval forecasts. The proposed methods are validated through an empirical application to hourly measurements of particulate matter in 51 size categories in London.

Keywords: functional principal component analysis; multiple functional time series; particle number size distribution; penalized least squares; ridge estimation; split conformal prediction

1 Introduction

Modelling and forecasting particulate matter is crucial in environmetrics because of its implications for public health and policy making. As hundreds of millions of people worldwide experience the impact of climate change, it takes an incalculable toll on human health and well-being. Several statistical methods have been proposed to forecast particle matters (see Dong et al. 2009, for review). These include neural networks (Paschalidou et al. 2011) and multiple linear modelling (Stadlober et al. 2008), where a comparison between these two methods is presented in Slini et al. (2006). A commonality in these works is that they often employ discrete-time models, which depend on the specific time points at which measurements are taken. Since the concentrations of the particle matter are analysed as a discrete time series, we cannot recover the underlying continuous stochastic process that generates these observations. Instead, we can consider a functional data-analytic approach to analyse the changes in shape over time. Unlike previous studies that focus on a single particle size, our work considers a range of particle sizes.

In Section 2, we describe the daily curves of the hourly concentration of particle number size distribution with various aerodynamic diameters, ranging from 16.55 to 604.4 nanometres (nm), cf. 2500nm = 2.5um, which is the size of $PM_{2.5}$. By treating data as a time series of functions, we resort to functional data analysis of Ramsay & Silverman (2005) to extract information and to represent the intraday dependence. Let $P_t^s(u_j)$, $t = 1,\dots,n$, $j = 1,\dots,p$, $s = 1,\dots,S$ be the concentration of particle matter at the intraday time u_j on day t of size s. To stabilise the variance and better normalise the variance, we take the $\log_{10}$ transformation (see also Gromenko et al. 2017, Baerenbold et al. 2023), i.e., we work with the data

$$\mathcal{X}_t^s(u_j) = \log_{10}[P_t^s(u_j) + 1] \tag{1}$$

$$P_t^s(u_j) = 10^{\mathcal{X}_t^s(u_j)} - 1, \tag{2}$$

where one was added to avoid a zero count. By linear interpolation, we consider intraday curves as continuous curves $\mathcal{X}_t^s(u)$ constructed from $\mathcal{X}_t^s(u_j)$, $j = 1,\dots,p$, where p denotes the total number of intraday measurements of particle counts of size s. Since our data are equally spaced and densely observed, a linear interpolation of Hyndman et al. (2025) is computationally efficient for the data (see also Zhang & Wang 2016, Ramos-Carreño et al. 2024). In our example, there are $p = 24$ hourly time intervals representing 24 hours of intraday measurements. Once we have constructed a time series of functions, we work directly with a set of functional time series, denoted

by $[\mathcal{X}_1^s(u), \ldots, \mathcal{X}_n^s(u)]$, for $u \in \mathcal{I}$ where $\mathcal{I}$ denotes a function support range.

In the statistical literature, there has been a surge of interest in the development of (univariate) functional time series forecasting methods (see Horváth & Kokoszka 2012, Kokoszka & Reimherr 2017, for a general background). From a parametric perspective, Bosq (2000) propose the functional autoregressive of order one and derive one-step-ahead forecasts that are based on a regularised form of the Yule-Walker equations. Kokoszka & Reimherr (2013) extend functional autoregressive of order one to a higher order and propose a selection criterion to determine the optimal order. Klepsch & Klüppelberg (2017) propose the functional moving average model and introduce an innovations algorithm to obtain the best linear predictor. Klepsch et al. (2017) propose the functional autoregressive moving average process to predict highway traffic flows. From a nonparametric perspective, Hyndman & Shang (2009) apply functional principal component analysis to decompose a time series of functions into a set of functional principal components and their associated scores. While Hyndman & Shang (2009) use a univariate time-series forecasting method to forecast each set of principal component scores, Aue et al. (2015) consider a multivariate time-series forecasting method. Although the multivariate time-series forecasting method can capture the cross-correlation among scores, the univariate time-series forecasting method can model nonstationarity within each set of scores. Using historical curves and estimated functional principal components, the point forecasts are obtained by multiplying the forecast scores with estimated functional principal components.

The presence of multiple functional time series is not uncommon, such as subnational mortality rates in Jiménez-Varón et al. (2024, 2025), firm-specific cumulative intraday return in Leng et al. (2025), half-hourly energy consumption readings for thousands of London households (Chang et al. 2025), or hourly wind speed (m/s) over 5-km resolution in Saudi Arabia (Martinez-Hernandez & Genton 2023). In Section 3, we implement a multilevel functional time series method of Shang (2016) to produce *one-day-ahead* forecasts for a range of particle sizes. As a joint modelling technique, the multilevel functional time series method resembles a two-way functional analysis of variance studied by Morris et al. (2003), and it is a special case of the general "functional mixed model" in Morris & Carroll (2006).

When a functional time series is constructed from segments of a longer univariate time series, the most recent curve is observed sequentially and thus may not represent a complete curve. When we observe the first m_0 time periods of $\mathcal{X}_{n+1}^s(u)$, denoted by $[\mathcal{X}_{n+1}^s(u_e) = [\mathcal{X}_{n+1}^s(u_1), \ldots, \mathcal{X}_{n+1}^s(u_{m_0})]^\top$, we are interested in forecasting the data in the remaining period of day $n+1$, denoted by $\mathcal{X}_{n+1}^s(u)$,

where $u \in \mathcal{J}_l$ denotes the function support range for the updating period and $s = 1, \ldots, S$ represents different sizes of particles. The one-day-ahead forecasts do not utilise the newly arrived intraday observations, so it is desirable to dynamically update the point forecasts for the remaining time period of the most recent day $n + 1$. In Section 4, we consider several dynamic updating methods to improve point and interval forecast accuracies for different sizes of particles (see also Chiou 2012, Jiao et al. 2023, for function-on-function regression approaches). In function-on-function regression, the predictor is a block corresponding to newly observed data, while the response is another block corresponding to the data in the update period.

In Section 5, we present two ways to construct pointwise prediction intervals. Both approaches are distribution-free, model-agnostic, and computationally fast, since they only rely on a set of validation data. When newly arrived data are observed, they can also be used to update interval forecasts. Using the point and interval forecast error metrics in Section 6, we evaluate and compare the point and interval forecast accuracy in Section 7. The conclusion is presented in Section 8, along with some ideas on how the methodology can be further extended.

2 Intraday particle number size distribution (PNSD)

We consider particle number size distribution data, which provide information on particulate matter sizes and are widely used due to the richness of information they contain. PNSD involves considering particle sizes that range from 10 to 2,500 nm, encompassing the fine and ultrafine domains. Furthermore, observations can be recorded at a fine time resolution, resulting in high-dimensional time series, where the dimension is the number of different sizes in the PNSD data. The PNSD data to be analysed in this article were collected at a single urban background monitoring station located in a residential area of London (North Kensington, 51.52 N, 0.21 W; 27m above sea level). The station is part of the London Air Quality Network and the National Automatic Urban and Rural Network. It is situated on the grounds of Sion Manning School in St Charles Square. The raw data consist of hourly measurements from January 2011 to August 2018, obtained with an SMPS TSI 3080 + CPC TSI 3775, fitted with a long DMA instrument and a diffusion dryer, following the recommendations of the EUSAAR/ACTRIS protocol (Wiedensohler et al. 2012). At each time point u_j (hr), 51 different particle sizes are measured [1], ranging from 16.55 to 604.4 nm; $P_t^1(u_j), \ldots, P_t^{51}(u_j)$. Thus, for each day, the raw data is a 24×51 matrix. Data can be

[1] Strictly speaking, 51 is the number of size categories considered.

downloaded from the Department for Environment, Food and Rural Affairs webpage. Also, see Martinez-Hernandez et al. (2025), where the same data were analysed but to estimate the source profiles rather than forecasting.

We treat PNSD data as multivariate time series of functions (51-dimensional multivariate functional data). For each day t, we assume that we have 51 continuous functions instead of a matrix. Then, we apply the $\log_{10}$ transformation, as described in the introduction, to obtain $\mathcal{X}_t^s(u_j) = \log_{10}[P_t^s(u_j) + 1]$, where one was added to avoid zero count. Finally, the continuous function is obtained using interpolation: $\mathcal{X}_t^1(u), \mathcal{X}_t^2(u), \ldots, \mathcal{X}_t^{51}(u)$, where $\mathcal{X}_t^s(u)$ represents the continuous intraday log PNSD measurement for size s. Figure 1 shows two functional time series corresponding to particle sizes 29.45 nm and 100 nm. Each plot presents only 50 consecutive days. Initially, we can observe that each functional time series exhibits a distinct dynamic and trend. In our model, we will consider the different variabilities between different sizes and obtain the one-day-ahead point and interval forecasts.

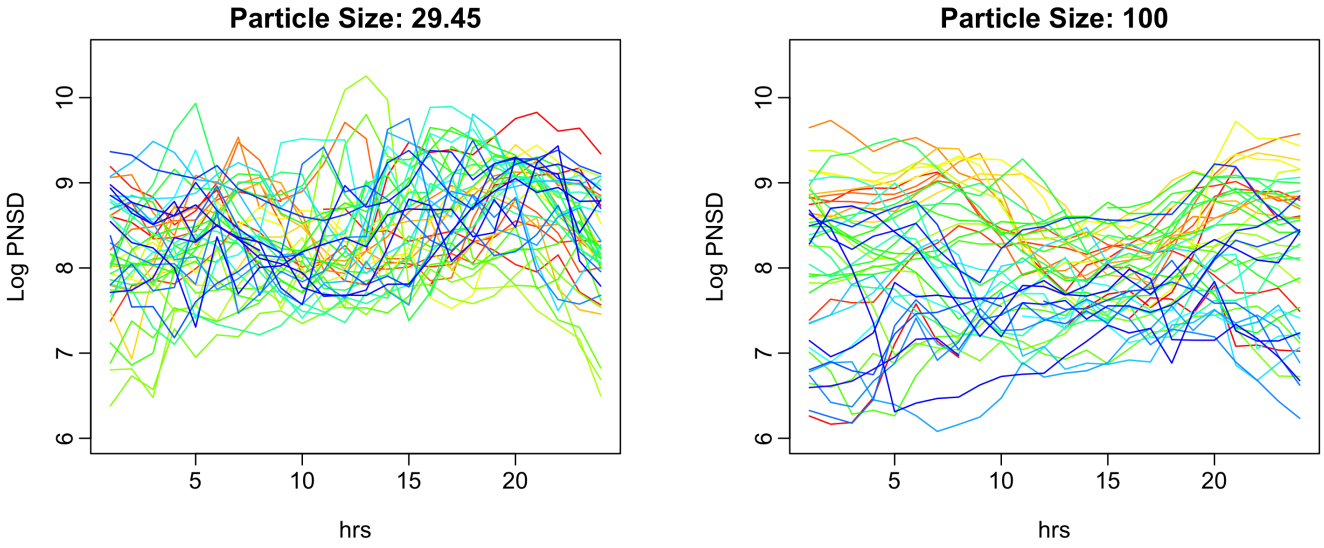


Figure 1: Functional time series of the PNSD for two sizes. Each plot illustrates 50 consecutive days of data, where each function represents one day. We can observe that the evolution of the time series of functions depends on the size.

The data exhibit an inherent weekly periodicity. For example, consecutive Sundays tend to be more similar to each other than two consecutive days, such as Saturday and Sunday. This similarity arises from people's routines; for example, there is typically less traffic on weekends compared to weekdays. As a result, fewer particles are emitted from vehicle emissions during this time. Therefore, we define our time series of continuous functions as a sequence based on the days of the week, such as the sequence of all Mondays. With this approach, the sample size of the multivariate

functional time series is 408 (corresponding to the total number of weeks in the data). Although it is worth noting that this approach can easily be modified to accommodate weekly curves or even monthly curves. In the forecasting experiment, we also consider weekly curves from 168 data points. Using the functional time-series method, we produce one-week-ahead forecasts, and these forecasts are then segmented according to each day of the week.

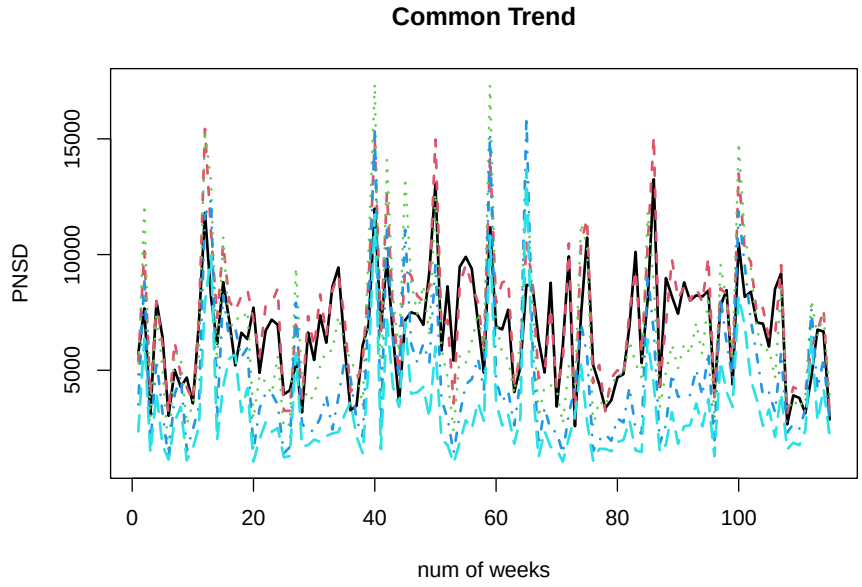


Figure 2: Time series of average daily counts for a given day (sequence of Mondays). Each time series represents different sizes (in nm): $29.45, 39.25, 64.95, 100.00, 133.40$. We can see that there is a common trend.

Given the strong correlation among different particle sizes, it is expected that a common trend will be observed - denoted as $R_t(u)$ in Section 3. One way to visualise this common trend is as follows. Given the sequence of functional data, for each size s, we can average the counts for each functional data, $(1/24)\sum_{j=1}^{24} P_t^s(u_j)$, for all t, resulting in a univariate time series. This time series allows us to visualise the trend over time. Figure 2 shows five time series corresponding to particle sizes (nm) $s = 29.45, 39.25, 64.95, 100.00, 133.40$. We can see that there is evidence of a common trend since all time series have a similar pattern. Therefore, we will consider a common trend component in our model, defined in Section 3.

To examine stationarity, we implement a functional KPSS test of Horvath et al. (2014) for each of the 51 series. From the p-value image plot in Figure 3, we conclude that all series are stationary since the p-value > 0.1.

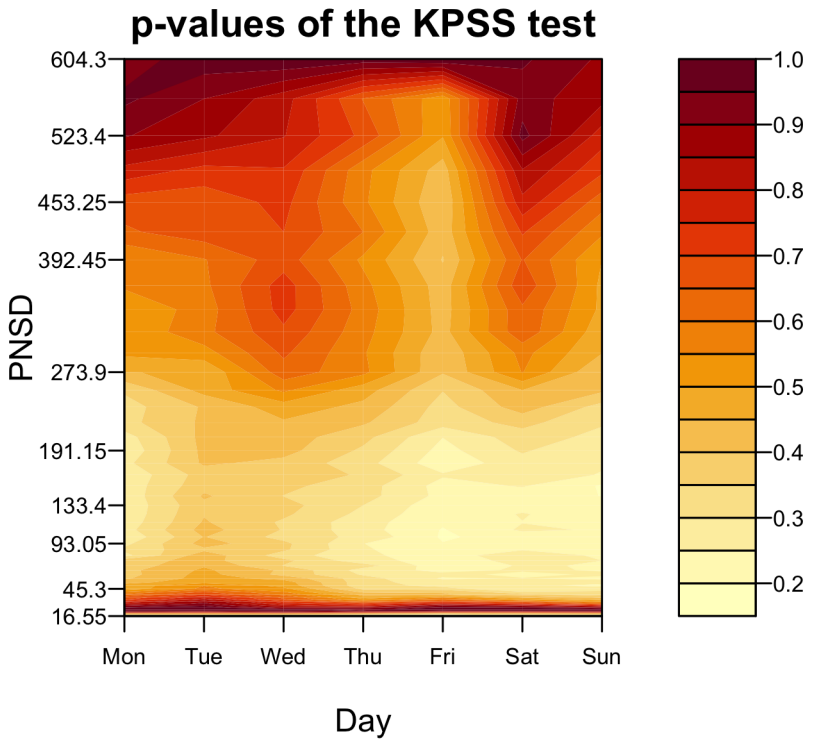


Figure 3: An image plot containing the p-values of the functional KPSS test of Horvath et al. (2014) for each series.

3 Multiple functional time series forecasting method

3.1 Multilevel functional time-series (MLFTS) method

The essence is to decompose curves among different sizes into a size-specific mean $\mu^s(u)$, a common trend among sizes $R_t(u)$, a size-specific residual trend $U_t^s(u)$, and measurement error $e_t^s(u)$. The common trend and size-specific residual trends are modelled by projecting them onto the eigenvectors of the covariance operators of the aggregate and size-specific centred stochastic processes, respectively. The aggregated series can be formed by averaging functional time series in various sizes. To be specific, $X_t^s(u)$ can be written as

$$X_t^s(u) = \mu^s(u) + R_t(u) + U_t^s(u) + e_t^s(u), \qquad u \in \mathcal{I}.$$

The size-specific mean is estimated by $\widehat{\mu}^s(u) = \frac{1}{n}\sum_{t=1}^{n} X_t^s(u)$, the common and size-specific trends can be estimated by

$$\widehat{R}_t(u) = \sum_{k=1}^{\infty} \widehat{\beta}_{t,k}\widehat{\phi}_k(u) \approx \sum_{k=1}^{K} \widehat{\beta}_{t,k}\widehat{\phi}_k(u)$$

$$\widehat{U}_t^s(u) = \sum_{l=1}^{\infty} \widehat{\gamma}_{t,l}^s \widehat{\psi}_l^s(u) \approx \sum_{l=1}^{L} \widehat{\gamma}_{t,l}^s \widehat{\psi}_l^s(u),$$

where $\widehat{\boldsymbol{\beta}}_k = (\widehat{\beta}_{1,k}, \dots, \widehat{\beta}_{n,k})$ represents the k^{th} principal component scores of $\boldsymbol{R}(u) = [R_1(u), \dots, R_n(u)]$, and $\widehat{\boldsymbol{\Phi}}(u) = [\widehat{\phi}_1(u), \dots, \widehat{\phi}_K(u)]$ are the corresponding functional principal components. Similarly, $\widehat{\gamma}_l^s = (\widehat{\gamma}_{1,l}^s, \dots, \widehat{\gamma}_{n,l}^s)$ represents the l^{th} principal component scores of $\boldsymbol{U}^s(u) = [U_1^s(u), \dots, U_n^s(u)]$, and $\widehat{\boldsymbol{\Psi}}^s(u) = [\widehat{\psi}_1^s(u), \dots, \widehat{\psi}_n^s(u)]$ are the corresponding functional principal components. Averaged over 51 series, we compute the common trend as $R_t(u) = \frac{1}{51}\sum_{s=1}^{51} X_t^s(u)$. The series-specific trend is captured by $U_t^s(u) = X_t^s(u) - R_t(u)$. To model $R_t(u)$ and $U_t^s(u)$, we resort to functional principal component analysis of their covariance functions, respectively. So, $X_t^s(u)$ can be estimated by

$$X_t^s(u) = \widehat{\mu}^s(u) + \sum_{k=1}^{K} \widehat{\beta}_{t,k} \widehat{\phi}_k(u) + \sum_{l=1}^{L} \widehat{\gamma}_{t,l}^s \widehat{\psi}_l^s(u) + e_t^s(u). \tag{3}$$

Notice that the intraday dependence is described with the covariance operators obtained with $\boldsymbol{R}(u)$ and $\boldsymbol{U}^s(u)$.

From the estimated eigenvalues, we can estimate the proportion of variability explained by the aggregated data; a measure of within-cluster variability is given by

$$\frac{\int_{u\in\mathcal{I}} \text{Var}[\boldsymbol{R}(u)]du}{\int_{u\in\mathcal{I}} \text{Var}[\boldsymbol{R}(u)]du + \int_{u\in\mathcal{I}} \text{Var}[\boldsymbol{U}^s(u)]du} \approx \frac{\sum_{l=1}^{L} \lambda_l}{\sum_{l=1}^{L} \lambda_l + \sum_{m=1}^{M} \lambda_m}, \tag{4}$$

where λ_l and λ_m denote the population eigenvalues of the common trend and the size-specific residual trend. Equation (4) not only presents a modelling and forecasting approach but also allows us to better understand and interpret the shared variability.

The selection of retained functional principal components has garnered considerable attention. Some popular estimation methods include: (1) scree plots or the fraction of variance explained by the first few functional principal components (Chiou 2012); (2) Akaike information criterion (Aue et al. 2015) and Bayesian information criterion (Otto & Salish 2025); (3) predictive cross-validation with one-curve-leave-out (Rice & Silverman 1991); (4) the bootstrap technique (Hall & Vial 2006); (5) eigenvalue ratio criterion (Li et al. 2020); (6) set $K = L = 6$ (Hyndman et al. 2013). Here, we consider the simplest strategy by setting $K = L = 6$. From a forecast accuracy viewpoint, selecting more than the optimal number of components does not harm the forecast accuracy. Since the eigenvalues are ordered in a decreasing order, the forecasts of the additional principal component scores tend to close to zero.

By conditioning on the estimated functional principal components $[\widehat{\phi}_1(u),\dots,\widehat{\phi}_K(u)]$ and $[\widehat{\psi}^s_1(u),\dots,\widehat{\psi}^s_L(u)]$ in (3) and observed data, the one-step-ahead point forecast can be obtained as

$$\widehat{X}^s_{n+1|n}(u) = \widehat{\mu}^s(u) + \sum_{k=1}^{K} \widehat{\beta}_{n+1|n,k}\widehat{\phi}_k(u) + \sum_{l=1}^{L} \widehat{\gamma}^s_{n+1|n,l}\widehat{\psi}^s_l(u),$$

where $\widehat{\beta}_{n+1|n,k}$ and $\widehat{\gamma}_{n+1|n,l}$ denote the one-step-ahead forecasts of the k^{th} and l^{th} principal component scores, respectively. These forecasts can be obtained from the exponential smoothing method (see, e.g., Hyndman et al. 2008). Using the inverse transformation in (2), we obtain $\widehat{P}^s_{n+1|n}(u)$.

3.2 Functional factor model

To accommodate strong cross-sectional dependence driven by the latent structure, we may adopt the factor model approach to multiple functional time series. Gao et al. (2019) apply the functional principal component analysis to each series and then model the component scores via the classic factor model. This method may lead to information loss due to two-stage dimension reduction. To avoid this issue, two types of functional factor model have been proposed with different constructions of common components. Tavakoli et al. (2023) introduce a high-dimensional functional factor model with functional loadings and scalar factors, while Guo et al. (2025) propose a different version with scalar loadings and functional factors. Leng et al. (2025) propose a general framework expressed as

$$X^s_t(u) = \sum_{\ell=1}^{L_*} \int_{v\in\mathcal{J}} B_{s\ell}(u,v)F_{t\ell}(v)dv + \varepsilon^s_t(u). \tag{5}$$

For each component of the ℓ-dimensional vector of functional factors, we consider the series approximation:

$$F_{t\ell}(v) = \Phi_\ell(v)^\top G_t + \eta_{t\ell}(v). \tag{6}$$

By plugging (6) into (5), we obtain

$$X^s_t(u) = \sum_{\ell=1}^{L_*} \int_{v\in\mathcal{J}} B_{s\ell}(u,v)[\Phi_\ell(v)^\top G_t + \eta_{t\ell}(v)]dv + \varepsilon^s_t(u) \tag{7}$$

Letting

$$\Lambda_s(u) = \sum_{\ell=1}^{L_*} \int_{v\in\mathcal{J}} B_{s\ell}(u,v)\Phi_\ell(v)^\top dv \tag{8}$$

and

$$y_t^s(u) = \sum_{\ell=1}^{L_*} \int_{v\in\mathcal{J}} B_{s\ell}(u,v)\eta_{t\ell}(v)dv, \tag{9}$$

combining (7)–(9), it leads to

$$\mathcal{X}_t^s(u) = \Lambda_s(u)^\top G_t + \underbrace{y_t^s(u) + \varepsilon_t^s(u)}_{\epsilon_t^s(u)}.$$

To estimate factor loading functions and real-valued factors, we resort to the least-squares objective function:

$$\sum_{t=1}^{n}\sum_{s=1}^{S} \left\|\mathcal{X}_t^s(u) - \Lambda_s(u)^\top G_t\right\|_2^2,$$

where $\|\cdot\|_2$ denotes the norm of elements in the Hilbert space, $\Lambda_s(u)$ is a q-dimensional vector of functions and G_t is a q-dimensional vector of numbers. Minimisation of the least-squares objective function can be solved via eigenanalysis of

$$\boldsymbol{\Delta} = (\Delta_{tt^\dagger})_{n\times n} \quad \text{with} \quad \Delta_{tt^\dagger} = \frac{1}{S}\sum_{s=1}^{S}\int_{u\in\mathcal{J}} \mathcal{X}_t^s(u)\mathcal{X}_{t^\dagger}^s(u)du. \tag{10}$$

Let $\widetilde{\boldsymbol{G}} = (\widetilde{G}_1, \widetilde{G}_2, \ldots, \widetilde{G}_n)^\top$ be a matrix consisting of the eigenvector (multiplied by root-n) corresponding to the q largest eigenvalues of $\boldsymbol{\Delta}$ in (10). The factor loadings are estimated as

$$\widetilde{\Lambda}_s(u) = \frac{1}{n}\sum_{t=1}^{n} \mathcal{X}_t^s(u)\widetilde{G}_t,$$

via least squares, using the normalisation restriction $\frac{1}{n}\sum_{t=1}^{n}\widetilde{G}_t\widetilde{G}_t^\top = \boldsymbol{I}_q$. To select q, we consider an information criterion (see Leng et al. 2025, equation 5.1). Consequently, the model residual function, itself a high-dimensional functional time series, can be expressed as $\widetilde{\epsilon}_t^s(u) = \mathcal{X}_t^s(u) - \widetilde{\Lambda}_s^\top(u)\widetilde{G}_t$. It can be modelled using the multilevel functional time-series method in Section 3.1.

4 Updating point forecasts

Although we present two functional time-series methods to produce one-day-ahead forecasts in Section 3, it is a feature of our problem where intraday measurements are continuously observed. Incorporating newly arrived information may improve forecast accuracy. In this section, we

describe methods for incorporating partially observed curves into the model.

4.1 Block moving (BM) method

The BM method utilises functional time-series methods but redefines the beginning and end time points of curves. Since intraday time is a continuous variable, we can alter its function support from $[u_1, u_p]$ to $[u_1, u_{m_0}] \cup (u_{m_0}, u_p]$. With loss of some data points in the first curve highlighted in blue, a partially observed curve can be completed. In Figure 4, we present a concept diagram.

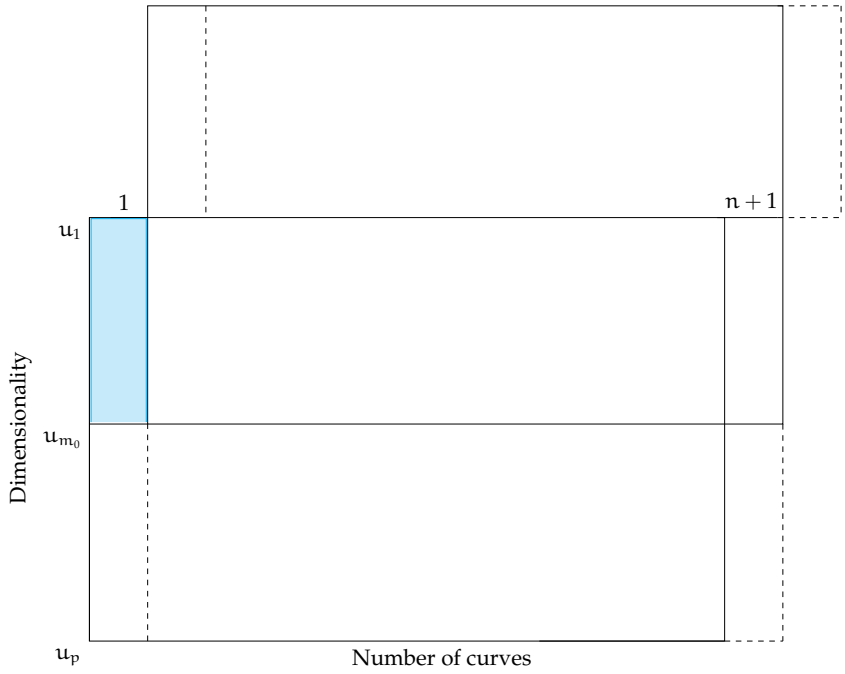


Figure 4: Dynamic update via the block-moving approach. The coloured region shows the data loss on the first day. The forecasts for the rest of the day $(n + 1)$ can be updated using forecasts from a functional time-series method applied to the top block.

Note that the BM method is implemented 23 times to cover different updating periods, from observing one hour of day $n + 1$ to observing 23 hours of that day.

4.2 Regression-based method

4.2.1 Ordinary least squares (OLS)

The remaining part of the most recent curve can be updated using a regression-based approach based on a multivariate functional principal component analysis (see, e.g., Chiou et al. 2014, Shang & Kearney 2022). By stacking s, we can convert a three-dimensional array to a two-dimensional

matrix, in which an eigenanalysis can be performed. Let $\boldsymbol{X}$ be a $(S \times p) \times n$ matrix, where $(S \times p)$ represents the stacked variables. From its empirical covariance of dimension $(S \times p) \times (S \times p)$, we obtain

$$\boldsymbol{X} = \boldsymbol{F}\boldsymbol{B},$$

where $\boldsymbol{F}$ is a $(S \times p) \times N$ matrix, consisting of N retained functional principal components, and $\boldsymbol{B}$ represents principal component scores of dimension $N \times n$. The advantage of multivariate functional principal component analysis is its ability to model the correlation among S particulate matter sizes.

For a given particle size s, let $\mathcal{F}$ be a $p \times N$ matrix and $\mathcal{F}_e$ be the $m_0 \times N$ matrix of the explanatory variable, whose $(i,j)^{\text{th}}$ entry is $\widehat{\phi}_i^s(u_j)$ for $1 \leqslant j \leqslant m_0$ and $1 \leqslant i \leqslant N$. Let $\mathcal{X}_{n+1}^s(u_e)$ be the $m_0 \times 1$ matrix of the response variable, $\boldsymbol{\beta}_{n+1}^s = (\beta_{n+1,1}^s, \ldots, \beta_{n+1,N}^s)^\top$ be the regression coefficients, and $\eta_{n+1}^s(u_e) = [\eta_{n+1}^s(u_1), \ldots, \eta_{n+1}^s(u_{m_0})]^\top$ be the error term. When the mean-adjusted $\widehat{\mathcal{X}}_{n+1}^{*,s}(u_e) = \mathcal{X}_{n+1}^s(u_e) - \widehat{\mu}^s(u_e)$ becomes available, the regression equation can be expressed as

$$\widehat{\mathcal{X}}_{n+1}^{*,s}(u_e) = \mathcal{F}_e^s \boldsymbol{\beta}_{n+1}^s + \eta_{n+1}^s(u_e).$$

The $\boldsymbol{\beta}_{n+1}^s$ can be estimated via an OLS estimator, giving

$$\widehat{\boldsymbol{\beta}}_{n+1}^{s,\text{OLS}} = ((\mathcal{F}_e^s)^\top \mathcal{F}_e^s)^{-1} (\mathcal{F}_e^s)^\top \widehat{\mathcal{X}}_{n+1}^{*,s}(u_e).$$

The OLS forecast of $\mathcal{X}_{n+1}^s(u_l)$ can be given by

$$\widehat{\mathcal{X}}_{n+1}^{s,\text{OLS}}(u_l) = \widehat{\mu}^s(u_l) + \sum_{i=1}^{N} \widehat{\beta}_{n+1,i}^{s,\text{OLS}} \widehat{\phi}_i^s(u_l),$$

where $\{u_{m_0+1}^s, \ldots, u_\tau^s; \tau \leqslant p\}$ represents the discretised time points in the remaining intraday time period.

4.2.2 Ridge regression (RR)

The OLS estimator can be numerically unstable due to the use of the pseudo-inverse. To solve this problem, we employ the ridge regression estimator proposed by Hoerl & Kennard (1970) by introducing a penalty. The RR method shrinks the regression coefficient estimates towards zero. In statistics, it is also known as Tikhonov regularisation (see, e.g., Benatia et al. 2017). It aims to

minimise a penalised residual sum of squares:

$$\underset{\boldsymbol{\beta}^{s}_{n+1}}{\arg\min}\left\{(\widehat{\mathcal{X}}^{*,s}_{n+1}(u_e)-\boldsymbol{\mathcal{F}}^{s}_{e}\boldsymbol{\beta}^{s}_{n+1})^{\top}(\widehat{\mathcal{X}}^{*,s}_{n+1}(u_e)-\boldsymbol{\mathcal{F}}^{s}_{e}\boldsymbol{\beta}^{s}_{n+1})+\lambda(\boldsymbol{\beta}^{s}_{n+1})^{\top}\boldsymbol{\beta}^{s}_{n+1}\right\},$$

where $\lambda > 0$ is a shrinkage parameter that controls the amount of shrinkage. Taking the first-order derivative with respect to $\boldsymbol{\beta}_{n+1}$, the RR estimate is

$$\widehat{\boldsymbol{\beta}}^{s,RR}_{n+1}=((\boldsymbol{\mathcal{F}}^{s}_{e})^{\top}\boldsymbol{\mathcal{F}}_{e}+\lambda\boldsymbol{I}_{N})^{-1}(\boldsymbol{\mathcal{F}}^{s}_{e})^{\top}\widehat{\mathcal{X}}^{*,s}_{n+1}(u_e),$$

where $\boldsymbol{I}_N$ is an $(N \times N)$ identity matrix. When $\lambda \to 0$, $\widehat{\boldsymbol{\beta}}^{s,RR}_{n+1}$ approaches $\widehat{\boldsymbol{\beta}}^{s,OLS}_{n+1}$; when $\lambda \to \infty$, $\widehat{\boldsymbol{\beta}}^{s,RR}_{n+1} \to 0$; when $0 < \lambda < \infty$, $\widehat{\boldsymbol{\beta}}^{s,RR}_{n+1}$ is a weighted average between 0 and $\widehat{\boldsymbol{\beta}}^{s,OLS}_{n+1}$. With an optimally selected λ value, the RR forecast of $\mathcal{X}^{s}_{n+1}(u_l)$ can be given by

$$\widehat{\mathcal{X}}^{s,RR}_{n+1}(u_l)=\widehat{\mu}^{s}(u_l)+\sum_{i=1}^{N}\widehat{\beta}^{s,RR}_{n+1,i}\widehat{\phi}^{s}_{i}(u_l).$$

4.2.3 Penalized least squares (PLS)

As an alternative shrinkage estimator, the PLS coefficient estimate minimises a penalised residual sum of squares:

$$\underset{\boldsymbol{\beta}^{s}_{n+1}}{\arg\min}\left\{(\widehat{\mathcal{X}}^{*,s}_{n+1}(u_e)-\boldsymbol{\mathcal{F}}^{s}_{e}\boldsymbol{\beta}^{s}_{n+1})^{\top}(\widehat{\mathcal{X}}^{*,s}_{n+1}(u_e)-\boldsymbol{\mathcal{F}}^{s}_{e}\boldsymbol{\beta}^{s}_{n+1})+\lambda(\boldsymbol{\beta}^{s}_{n+1}-\widehat{\boldsymbol{\beta}}^{s,TS}_{n+1|n})^{\top}(\boldsymbol{\beta}^{s}_{n+1}-\widehat{\boldsymbol{\beta}}^{s,TS}_{n+1|n})\right\},$$

where $\widehat{\boldsymbol{\beta}}^{s,TS}_{n+1|n}$ denotes the one-step-ahead point forecasts of the principal component scores using a univariate time-series forecasting method, such as exponential smoothing of Hyndman et al. (2008). Instead of shrinking towards zero, the PLS regression coefficient estimates shrink towards the TS coefficient estimates. By taking the first-order derivative with respect to $\boldsymbol{\beta}_{n+1}$, we obtain

$$\widehat{\boldsymbol{\beta}}^{s,PLS}_{n+1}=\left[(\boldsymbol{\mathcal{F}}^{s}_{e})^{\top}\boldsymbol{\mathcal{F}}^{s}_{e}+\lambda\boldsymbol{I}_{N}\right]^{-1}\left[(\boldsymbol{\mathcal{F}}^{s}_{e})^{\top}\widehat{\mathcal{X}}^{*,s}_{n+1}(u_e)+\lambda\widehat{\boldsymbol{\beta}}^{s,TS}_{n+1|n}\right].$$

When the shrinkage parameter $\lambda \to 0$, $\widehat{\boldsymbol{\beta}}^{s,PLS}_{n+1}$ approaches $\widehat{\boldsymbol{\beta}}^{s,OLS}_{n+1}$; when $\lambda \to \infty$, $\widehat{\boldsymbol{\beta}}^{s,PLS}_{n+1}$ approaches $\widehat{\boldsymbol{\beta}}^{s,TS}_{n+1}$; when $0 < \lambda < \infty$, $\widehat{\boldsymbol{\beta}}^{s,PLS}_{n+1}$ is a weighted average between $\widehat{\boldsymbol{\beta}}^{s,OLS}_{n+1}$ and $\widehat{\boldsymbol{\beta}}^{s,TS}_{n+1}$. With an optimally

selected λ value, the PLS forecast of $\mathcal{X}^s_{n+1}(u_l)$ can be given by

$$\widehat{\mathcal{X}}^{s,\text{PLS}}_{n+1}(u_l) = \widehat{\mu}^s(u_l) + \sum_{i=1}^{N} \widehat{\beta}^{s,\text{PLS}}_{n+1,i} \widehat{\phi}^s_i(u_l).$$

4.2.4 Selection of shrinkage parameters

The superior performance of the RR and PLS estimators is based on an accurate selection of the λ parameter. Depending on the updating period, the value of λ should also change, particularly as more data points are observed throughout the day. We split our data sample, consisting of each day of a week, into a training sample comprising curves 1 to 306 and a testing sample comprising curves 307 to 408. We further divide the training sample into the training set, consisting of curves 1 to 204, and the validation set, consisting of curves 205 to 306. Within the validation set, the optimal values of λ for different updating periods were determined by minimising a forecast error. In Figure 5, we present a concept diagram.

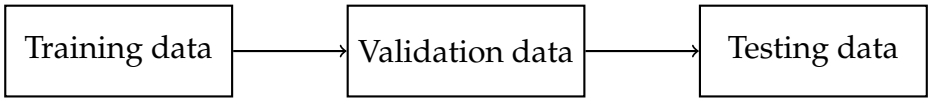


Figure 5: Illustration of the sample splitting. To select the optimal tuning parameter, a model is constructed using data from the training set to forecast data in the validation set. To evaluate forecast accuracy, a model is constructed using data from the training and validation sets to forecast data in the testing set. The validation dataset is used to select the optimal tuning parameters.

5 Interval forecast methods

Prediction intervals are a valuable tool for measuring the probabilistic uncertainty associated with point forecasts. As emphasised in Chatfield (1993, 2000), it is important to provide interval forecasts as well as point forecasts to (i) assess future uncertainty; (ii) enable different strategies to be planned for a range of possible outcomes indicated by the interval forecasts; (iii) compare forecasts from different methods more thoroughly; and (iv) explore different scenarios based on different assumptions.

5.1 Prediction interval calibration

We split the data sample, consisting of 408 curves, into training, validation, and testing sets. Using the data in the training sample, we implement an expanding window forecast scheme to obtain the one-step-ahead forecast in the validation set. We have 102 forecasts compared to the holdout sample in the validation set. From these residual functions, we compute the pointwise standard deviation, denoted by

$$\delta^s(u) = \text{sd}\left[\widehat{\xi}^s_1(u), \widehat{\xi}^s_2(u), \dots, \widehat{\xi}^s_{102}(u)\right], \tag{11}$$

where $\widehat{\xi}^s_\omega(u) = P^s_\omega(u) - \widehat{P}^s_\omega(u)$, for $\omega = 1, 2, \dots, 102$. For a level of significance α, customarily $\alpha = 0.2$ or 0.05, we aim to determine $(\underline{\theta}^s_\alpha, \overline{\theta}^s_\alpha)$ so that $(1-\alpha)\times 100\%$ of the residuals satisfy

$$\underline{\theta}^s_\alpha \delta^s(u) \leqslant \widehat{\xi}^s_\omega(u) \leqslant \overline{\theta}^s_\alpha \delta^s(u).$$

For ease of optimisation, the constants $\underline{\theta}^s_\alpha$ and $\overline{\theta}^s_\alpha$ are chosen equal. By the law of large numbers, one may achieve

$$\text{pr}[-\theta^s_\alpha \delta^s(u) \leqslant P^s_{n+1}(u) - \widehat{P}^s_{n+1}(u) \leqslant \theta^s_\alpha \delta^s(u)] \approx \frac{1}{102}\sum_{\omega=1}^{102} \mathbb{1}\left[-\theta^s_\alpha \delta^s(u) \leqslant \widehat{\xi}^s_\omega(u) \leqslant \theta^s_\alpha \delta^s(u)\right],$$

where $\mathbb{1}[\cdot]$ denotes the binary indicator function. For a given level of significance α, the value of θ_α is selected by achieving the smallest absolute difference between the empirical and nominal coverage probabilities (see also Shang & Haberman 2025).

5.2 Conformal prediction interval

As an alternative, a popular methodology known as conformal prediction (Shafer & Vovk 2008) can be used to construct probabilistic forecasts calibrated on out-of-sample interval forecast errors. Conformal prediction is model-agnostic and provides a distribution-free method to construct prediction sets with a finite-sample coverage guarantee. From the absolute value of $\widehat{\xi}^s_\omega(u)$, we compute its $(1-\alpha)\times 100\%$ quantile, denoted by $q^s_\alpha(u)$. The prediction interval can be obtained as

$$\left[\widehat{P}^s_{n+1}(u) - q^s_\alpha(u), \widehat{P}^s_{n+1}(u) + q^s_\alpha(u)\right], \tag{12}$$

where $\widehat{P}^s_{n+1}(u)$ denotes the one-day-ahead point forecasts for the data in the testing set. A limitation is that the conformal prediction approach has a fixed width of $2 \times q^s_\alpha(u)$, which is not dependent on $\widehat{\mathcal{X}}^s_{n+1}(u)$ (Romano et al. 2019). However, without the selection of any tuning parameter, the approach of the conformal prediction interval is computationally advantageous and works particularly well when the sample sizes in the validation and testing sets are reasonably large (Dhillon et al. 2024).

5.3 Updating interval forecasts

As new intraday data become available sequentially, prediction intervals can be dynamically updated to improve the accuracy of interval forecasts. Within the validation set, the optimal ridge penalty parameter λ is selected based on point forecast errors. In addition, the associated residual functions are computed. Using newly observed intraday data, the residuals corresponding to the remaining forecast period generally tend to diminish in magnitude, facilitating the refinement of the prediction intervals. This refinement can be performed using either the sd approach in (11) or the conformal prediction in (12).

6 Measures of point and interval forecast accuracy

We evaluate the performance of the forecast methods using different measures: mean absolute percentage error (MAPE), coverage probability difference (CPD) and mean interval scores.

6.1 Mean absolute percentage error

We compute the point forecasts for the data in the testing set and evaluate the forecast accuracy by MAPE. MAPE measures the closeness of the forecasts compared to the actual values, regardless of the sign. It can be expressed as

$$\text{MAPE}^s(u_j) = \frac{1}{102}\sum_{\omega=1}^{102}\frac{\left|P^s_\omega(u_j) - \widehat{P}^s_\omega(u_j)\right|}{P^s_\omega(u_j)} \times 100\%, \qquad j = 1, \ldots, p.$$

where $P^s_\omega(u_j)$ denotes the actual holdout sample for the j^{th} intraday period on day ω for a particle size s, and $\widehat{P}^s_\omega(u_j)$ denotes the one-day-ahead point forecasts for the holdout sample in the

validation or testing set, in which each of them contains 102 curves.

6.2 Coverage probability difference and mean interval scores

To evaluate the interval forecast accuracy, we utilise the CPD, which is the absolute difference between the empirical and nominal coverage probabilities. For each year in the validation or testing set, the one-step-ahead prediction intervals were computed at the $(1-\alpha)\times 100\%$ nominal coverage probability. Let $[\widehat{P}_{\omega}^{s,lb}(u_j), \widehat{P}_{\omega}^{s,ub}(u_j)]$ denote the lower and upper bounds, respectively, for ω^{th} in the testing set. For a given particle size s at a particular hour of the day, the empirical coverage probability is defined as

$$\text{ECP}^{s}(u_j) = 1 - \frac{1}{102}\sum_{\omega=1}^{102}\left[\mathbb{1}(P_{\omega}^{s}(u_j) > \widehat{P}_{\omega}^{s,ub}(u_j)) + \mathbb{1}(P_{\omega}^{s}(u_j) < \widehat{P}_{\omega}^{s,lb}(u_j))\right],$$

where $\mathbb{1}(\cdot)$ denotes the binary indicator function. The CPD is the absolute difference between the empirical and nominal coverage probabilities.

As defined by Gneiting & Raftery (2007) and Gneiting & Katzfuss (2014), a scoring rule for the interval forecast at the time point u_j is

$$\begin{aligned} S_{\alpha}[\widehat{P}_{\omega}^{s,lb}(u_j), \widehat{P}_{\omega}^{s,ub}(u_j), P_{\omega}^{s}(u_j)] = [\widehat{P}_{\omega}^{s,ub}(u_j) - \widehat{P}_{\omega}^{s,lb}(u_j)] + \frac{2}{\alpha}[\widehat{P}_{\omega}^{s,lb}(u_j) - P_{\omega}^{s}(u_j)]\mathbb{1}[P_{\omega}^{s}(u_j) < \widehat{P}_{\omega}^{s,lb}(u_j)] \\ + \frac{2}{\alpha}[P_{\omega}^{s}(u_j) - \widehat{P}_{\omega}^{s,ub}(u_j)]\mathbb{1}[P_{\omega}^{s}(u_j) > \widehat{P}_{\omega}^{s,ub}(u_j)], \end{aligned}$$

where α denotes a level of significance. Averaging over 102 days, the mean interval score is given by

$$\overline{S}_{\alpha}^{s}(u_j) = \frac{1}{102}\sum_{\omega=1}^{102} S_{\alpha}\left[\widehat{P}_{\omega}^{s,lb}(u_j), \widehat{P}_{\omega}^{s,ub}(u_j), P_{\omega}^{s}(u_j)\right].$$

The mean interval score rewards a narrow prediction interval, provided that the actual observations lie within the prediction interval.

7 Results

7.1 Expanding window scheme

An expanding window scheme is implemented to evaluate model and parameter stability over time. The expanding window analysis determines the variability of the model's parameters by iteratively computing parameter estimates and their resultant forecasts over an expanding window. Using the first 306 curves, we produce an one-day-ahead forecast for the 307^{th} curve. Then, we increase the training sample by one day. Using the first 307 curves, we produce an one-day-ahead forecast for the 308^{th} curve. This procedure continues until the training sample encompasses the entire dataset. In doing so, there are 102 days in the testing sample for evaluation of the forecast accuracy. Similarly, we can divide the 306 curves further into a training set and a 102-day validation set for parameter selection. In Figure 6, we present a concept diagram.

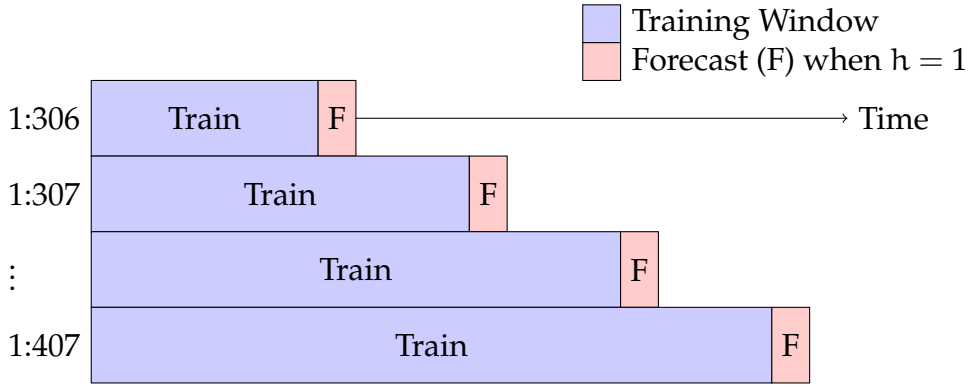


Figure 6: A diagram of the expanding-window one-step-ahead forecast scheme.

7.2 Evaluation of point forecast accuracy

For producing one-day-ahead forecasts, we consider two approaches: the MLFTS method and a combination of the functional factor model and the MLFTS method, as described in Section 3. In the second approach, the functional factor model can remove the common trend and residuals can then be modelled via the MLFTS method. Via the expanding-window forecast scheme, we sort the data according to the day of the week, and there are 408 observations for each day (number of weeks in the dataset). For example, we consider historical data on Mondays to produce the one-day-ahead forecast for Monday. In doing so, we may have a homogeneous group. The temporal dependence

is generally stronger for the same day between two successive weeks rather than two successive weekdays when producing one-step-ahead point forecasts.

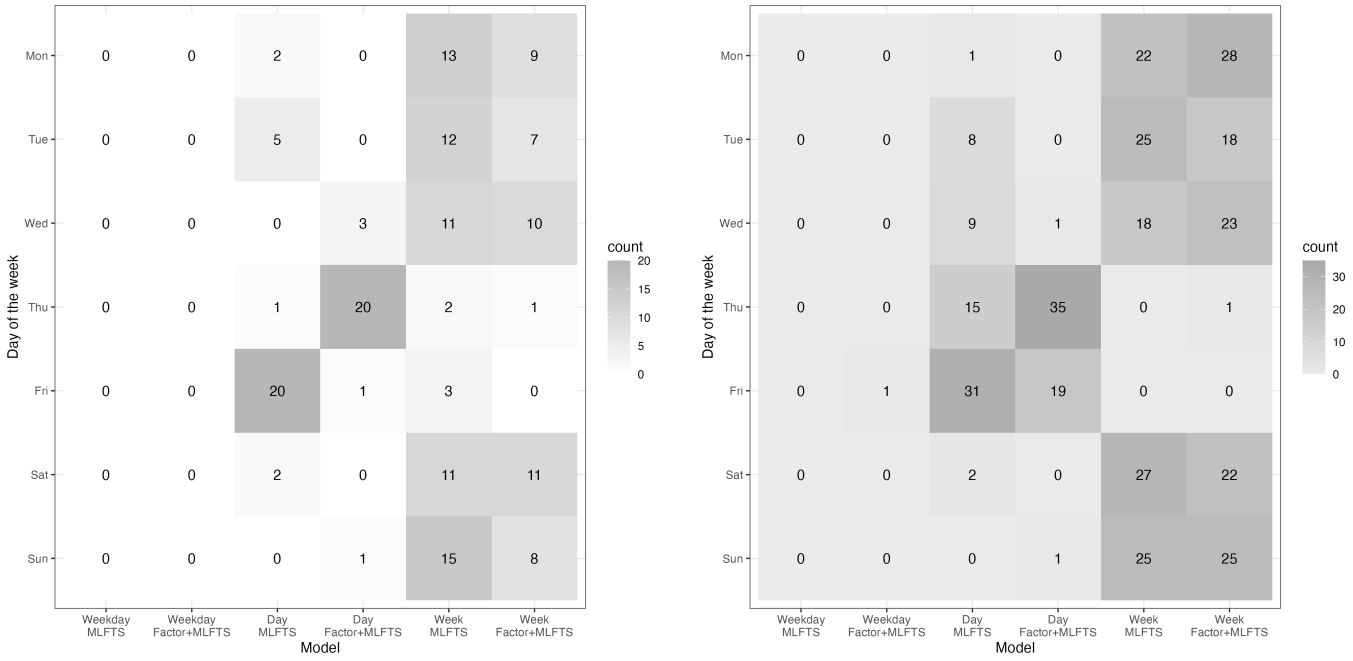
Table 1: As measured by the MAPE, a comparison of point forecast accuracy of the functional time-series methods, averaged over 24 hours and 51 particle sizes. We consider three ways of segmenting a univariate time series, namely, each weekday from historical weeks (Weekday), every consecutive day (Day), and every consecutive week (Week).

Data	Method	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Weekday	MLFTS	100.89	86.84	99.55	99.68	92.52	83.72	109.25
	Factor + MLFTS	99.39	85.76	97.49	98.47	92.59	84.76	109.06
Day	MLFTS	88.79	75.95	89.96	62.05	67.09	80.00	113.81
	Factor + MLFTS	89.33	76.31	89.52	61.60	67.86	80.95	113.83
Week	MLFTS	77.91	70.12	81.64	81.82	81.61	68.01	87.23
	Factor + MLFTS	78.22	70.21	81.54	82.18	82.90	68.08	87.27

In Table 1, we compare the one-step-ahead point forecast accuracy of two functional time series methods. The first method follows the MLFTS framework, which decomposes multiple functional time series into a common trend and series-specific residual trends. The second method builds upon the MLFTS approach, further incorporating a functional factor model to remove the trend component. There are three ways to form our functional data: 1) We segment our data into each day of the week for 408 weeks (i.e., Weekday); 2) We consider $408 \times 7 = 2856$ consecutive days, which consist of weekdays and weekends, as our functional data; 3) We consider the weekly curve as our functional data for 408 weeks. Among the three ways to form our functional data, the weekly curves provide the smallest point forecast errors for most days, except for Thursday and Friday. The second way is computationally time-consuming for training models using the expanding-window approach, and thus is not feasible for the purpose of dynamic updating. Using the `dm.test` function in the forecast package (Hyndman et al. 2025), with equal performance superiority, we compute the p-value for the functional data constructed from each weekday and week and obtain the p-value of zero. For the weekly curves, the p-value of the Diebold & Mariano's (2002) test is given as 0.1481 between the MLFTS and Factor+MLFTS methods, and this indicates equal forecast superiority.

In Figure 7a, we present a heatmap with a grey colour scaling. For the weekly curves, we segment its forecasts according to each day of the week, to facilitate a fair comparison. For each day of the week, we count the number of times a model provides the smallest MAPE averaged over 51 particle sizes. Each row of the heatmap sums up to 24 hours a day. Similarly, we also

display another heatmap in Figure 7b where we count the number of times a model provides the smallest MAPE averaged over 24 hours. Each row of the heatmap sums up to 51 particle sizes. Between the two models, it is advantageous to consider the combination of the factor model and MLFTS for most days, except Saturday.



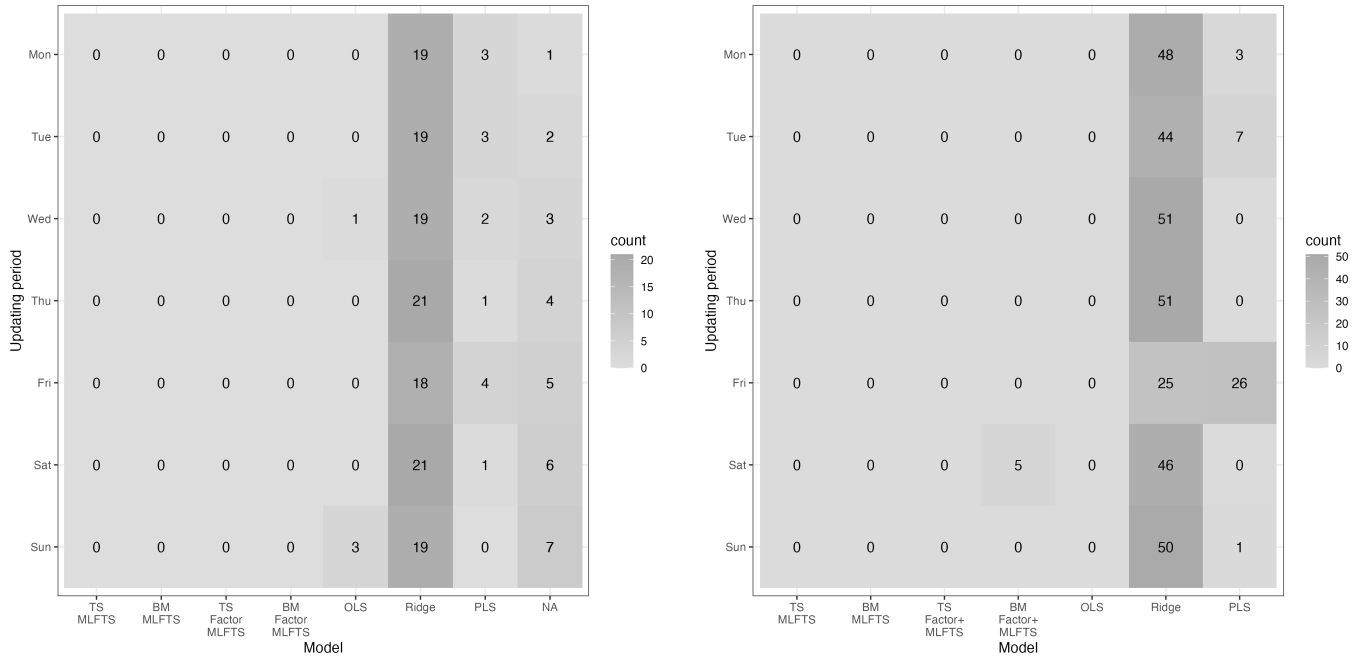
(a) Averaged over 51 sizes

(b) Averaged over 24 hours

Figure 7: Frequency of the most accurate functional time-series forecasting method for each day of the week. In (a), each row sums to 24 hours in a day. In (b), each row sums to 51 particulate matter sizes. We consider three ways of segmenting a univariate time series, namely each weekday from historical weeks (Weekday), every consecutive day (Day) and every consecutive week (Week).

In Figure 8, we consider the block-moving approach and regression-based approaches based on the OLS, ridge, and PLS estimators. In contrast to the functional time-series forecasting method, implementing a dynamic updating approach is advantageous in achieving better forecast accuracy. Between the two dynamic updating approaches, the regression-based approach shows superior forecast accuracy. Among the regression-based approach, the ridge estimator provides the most accurate forecasts.

In Figure 9, we plot the selected optimal (non-negative) tuning parameters in the ridge and PLS estimators throughout the intraday periods. These tuning parameters are chosen on the basis of the smallest overall MAPE in the validation data. Although a general pattern appears to exist, it also exhibits daily variation. For the ridge estimator, the estimated regression coefficient gradually shrinks towards zero; for the PLS estimator, the estimated regression coefficient gradually



(a) Averaged over 51 sizes

(b) Averaged over 24 hours

Figure 8: Frequency of the most accurate dynamic updating method for each day of the week. In (a), each row sums to 24 hours in a day. In (b), each row sums to 51 particulate matter sizes.

shrinks towards the OLS estimator throughout the day. Regarding the PLS estimator, the tuning parameters selected for Fridays tend to be larger than other days. Although it is difficult to pinpoint an affirmative reasoning, we believe that it might be due to the fact that Fridays often act as a transition day, where people living near the measurement station in London spend weekends elsewhere, thus influencing the PM measurement (Beevers et al. 2009).

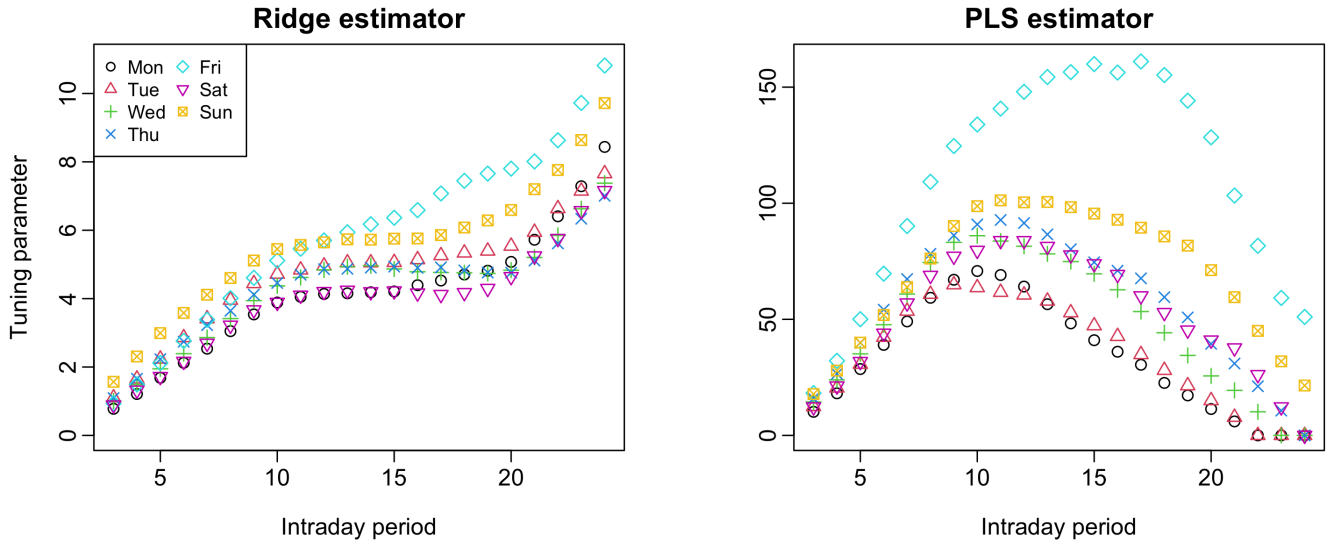


Figure 9: Optimally selected tuning parameter λ based on the MAPE for different updating periods for each day of the week.

7.3 Evaluation of interval forecast accuracy

In Table 2, we compare the interval forecast accuracy based on the CPD and the mean interval scores at the nominal coverage probabilities of 80% and 95%. At the 80% nominal coverage probability, the differences between the two prediction interval construction methods are marginal. However, at the 95% nominal coverage probability, the standard deviation approach shows a slight advantage over the conformal prediction. Based on the CPD, the *factor + MLFTS* method provides a smaller difference than the ones obtained from the MLFTS method. Oftentimes, the superiority in CPD is attributed to the larger interval scores.

Table 2: A comparison of the interval forecast accuracy, as measured by the CPD and mean interval scores, at the nominal coverage probabilities of 80% and 95% between the MLFTS and factor + MLFTS methods. Further, we consider two ways of constructing our functional data, constructed either from 24 or 168 hourly data points.

			CPD				Mean interval score			
			MLFTS		Factor + MLFTS		MLFTS		Factor + MLFTS	
Method	α	Day	sd	conformal	sd	conformal	sd	conformal	sd	conformal
Weekday	0.2	Mon	0.0191	0.0162	0.0183	0.0163	8982	8905	9106	9028
		Tue	0.0324	0.0319	0.0272	0.0251	8754	8715	8804	8769
		Wed	0.0308	0.0360	0.0306	0.0377	10049	9957	10204	10118
		Thu	0.0241	0.0284	0.0218	0.0260	9115	9063	9249	9213
		Fri	0.0202	0.0180	0.0149	0.0139	8645	8374	8806	8572
		Sat	0.0398	0.0387	0.0369	0.0367	7604	7511	7668	7596
		Sun	0.0233	0.0211	0.0169	0.0137	7929	7835	8021	7919
		Mean	0.0271	0.0272	0.0238	0.0242	8726	8623	8837	8745
	0.05	Mon	0.0201	0.0250	0.0194	0.0246	18241	18414	18411	18618
		Tue	0.0085	0.0059	0.0081	0.0054	17552	17623	17634	17734
		Wed	0.0177	0.0215	0.0182	0.0230	21046	21314	21358	21656
		Thu	0.0301	0.0348	0.0291	0.0346	18319	18519	18523	18730
		Fri	0.0179	0.0223	0.0163	0.0204	16610	16289	16673	16324
		Sat	0.0115	0.0090	0.0118	0.0084	14225	14338	14312	14427
		Sun	0.0057	0.0038	0.0047	0.0041	15637	15615	15673	15693
		Mean	0.0159	0.0175	0.0154	0.0172	17376	17445	17512	17597

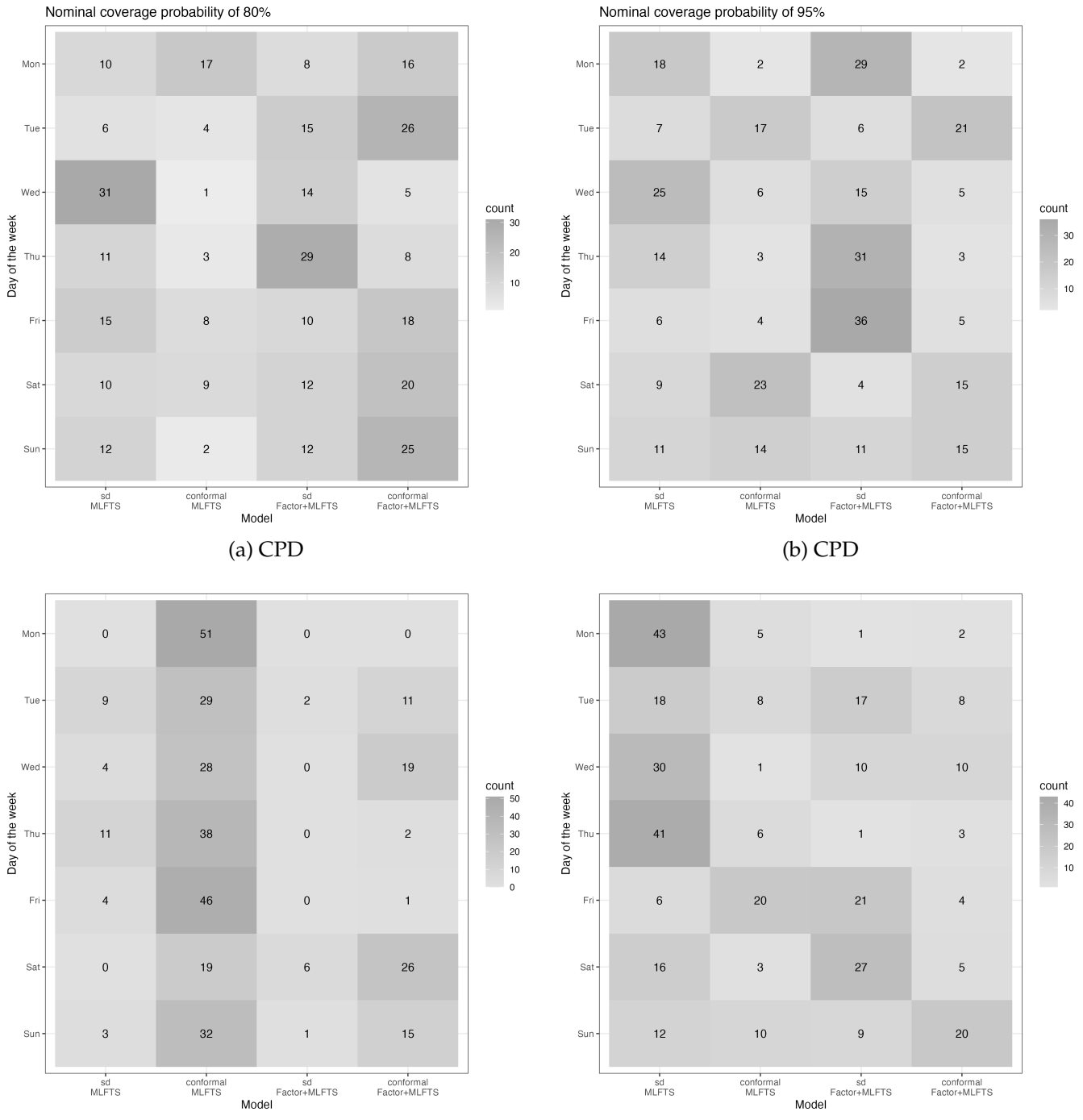
Continued on next page

			CPD				Mean interval score			
			MLFTS		Factor + MLFTS		MLFTS		Factor + MLFTS	
Method	α	Day	sd	conformal	sd	conformal	sd	conformal	sd	conformal
Week	0.2	Mon	0.0154	0.0101	0.0159	0.0106	8722	8650	8805	8721
		Tue	0.0577	0.0243	0.0579	0.0186	8683	8522	8735	8575
		Wed	0.0073	0.0339	0.0047	0.0303	10051	9974	10114	10004
		Thu	0.0339	0.0131	0.0383	0.0156	8818	8771	8926	8877
		Fri	0.0068	0.0102	0.0078	0.0087	8479	8221	8515	8242
		Sat	0.0097	0.0110	0.0102	0.0130	7249	7163	7330	7267
		Sun	0.0421	0.0143	0.0411	0.0177	7709	7601	7782	7663
		Mean	0.0247	0.0167	0.0251	0.0163	8530	8414	8601	8479
	0.05	Mon	0.0071	0.0189	0.0076	0.0190	173523	17614	17468	17822
		Tue	0.0094	0.0039	0.0088	0.0033	17222	17336	17239	17384
		Wed	0.0125	0.0193	0.0121	0.0197	20605	20882	20763	21092
		Thu	0.0245	0.0283	0.0255	0.0303	17568	17736	17682	17846
		Fri	0.0084	0.0088	0.0075	0.0080	16352	15933	16362	15945
		Sat	0.0055	0.0034	0.0058	0.0035	13603	13595	13665	13633
		Sun	0.0117	0.0012	0.0110	0.0011	14981	14863	15051	14955
		Mean	0.0113	0.0120	0.0112	0.0121	16812	16852	16890	16954

In Table 2, we consider our functional data constructed from 24 hourly data points in each day of the week or constructed from 168 hourly data in a week. We observe that the construction of weekly curves provides the smallest CPD and mean interval score at the nominal coverage probabilities of 80% and 95%. Using the `dm.test` function, we compute the p-value between the two methods for two ways of constructing functional curves. Based on the CPD, we find that the p-values are all zero, implying a statistically significant difference in the interval forecast accuracy.

In Figure 10, we show the number of times one method performs the best out of 51 different particle sizes for each day of the week. At the 80% nominal coverage probability, the Factor + MLFTS method, in combination with the conformal prediction, generally provides the smallest CPD. In contrast, the MLFTS method with the conformal prediction generally provides the smallest interval scores. At the 95% nominal coverage probability, the Factor + MLFTS method, in combination with the sd approach, generally provides the smallest CPD. In contrast, the MLFTS method with the sd

approach typically provides the smallest interval scores. Between the sd and conformal prediction approaches, there is no clear recommendation.



(a) CPD

(b) CPD

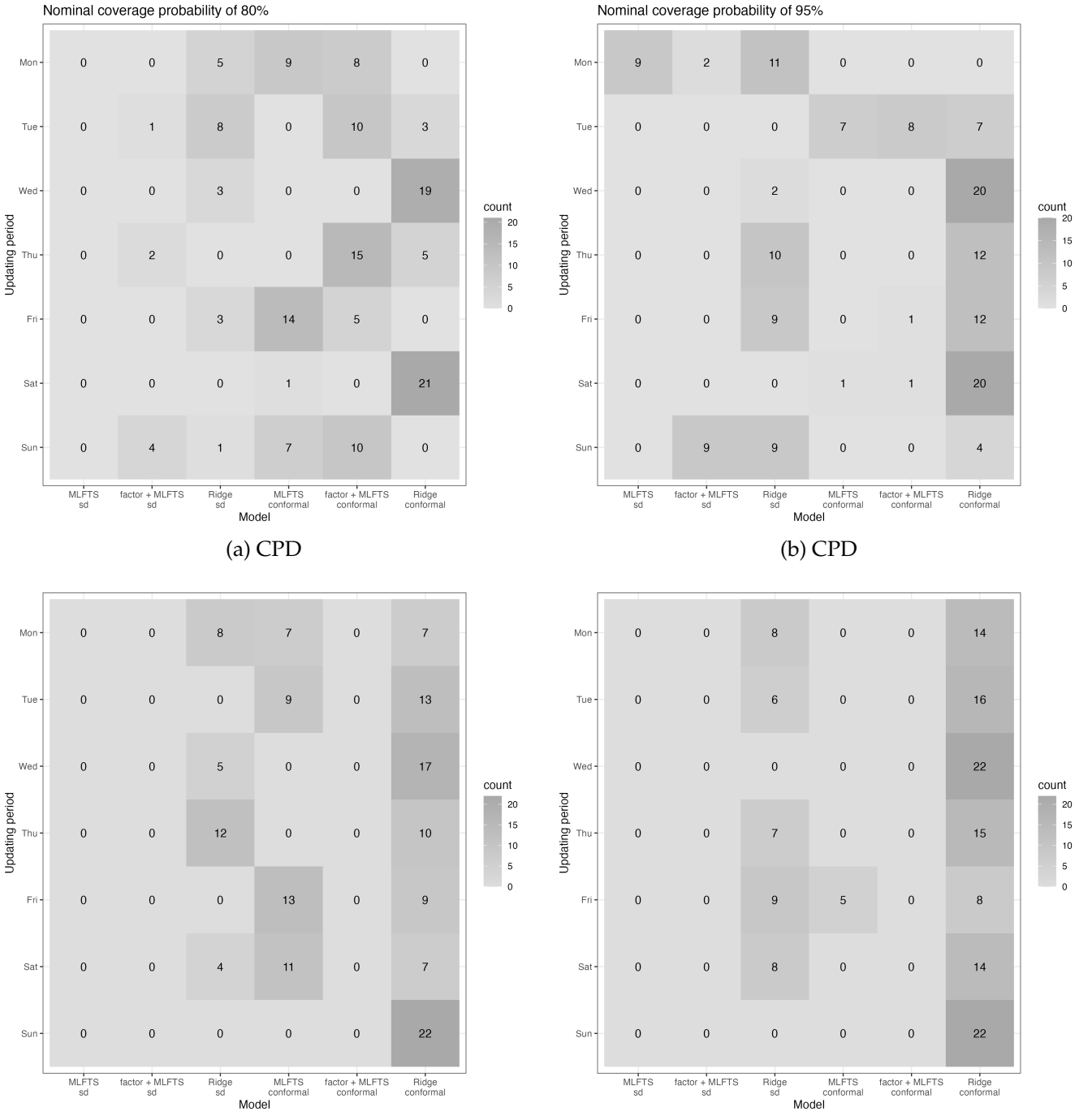
(c) Mean interval score

(d) Mean interval score

Figure 10: The CPD and mean interval scores among the functional time-series forecasting methods, comparing the sd and conformal prediction approaches.

When we have partially observed data in the most recent curve, we can also update our prediction intervals for the remaining period to achieve improved accuracy. Using the estimated

optimal λ values, we implement the sd and conformal prediction interval approaches to construct the prediction intervals for the remaining period. For the former method, it requires re-estimating the θ_{α}^{s} values based on the validation dataset. In Figure 11, the ridge estimator generally provides the smallest CPD and mean interval scores. Between the sd and conformal prediction approaches, the latter seems to be more advantageous.



(a) CPD

(b) CPD

(c) Mean interval score

(d) Mean interval score

Figure 11: For 22 different updating periods, we display the frequency of the most accurate method with the smallest CPD and mean interval score for each day of the week.

8 Conclusion

Understanding the dynamics of particulate matter of various sizes is crucial for public health. For example, it is well known that ultrafine particles (particles with diameter smaller than 100 nm) are more harmful to health, as they cause more severe pulmonary inflammation and remain longer in the lungs. We present functional time-series methods to produce one-day-ahead point and interval forecasts. By evaluating the accuracy of the point and interval forecast, we show the superiority of the factor model in combination with the MLFTS method. While the factor model removes the trend, the MLFTS method studies the patterns in the residuals. Such residual patterns include a common pattern shared by all sizes and a specific pattern for a given size.

When intraday data are observed sequentially, it is advantageous to consider dynamic updating to improve forecast accuracy. Among dynamic updating methods, we recommend the ridge estimator and demonstrate how the estimated regression coefficient of this estimator varies over intraday periods. Furthermore, we present the sd approach and the conformal prediction approach to construct prediction intervals. For producing one-step-ahead prediction intervals, the factor model, in combination with the multilevel functional time-series method, yields the smallest CPD, albeit at the expense of larger mean interval scores compared to the multilevel functional time-series method alone. When updating interval forecasts, the ridge estimator is recommended, along with the conformal prediction approach. For reproducibility, the computer  code is available at https://github.com/hanshang/PNSD.

There are several ways in which the methodology presented here can be further extended, and we briefly mention five below: 1) We study the MLFTS method, but other multiple functional time-series forecasting methods can be utilised. 2) The presence of outliers can affect the estimation of the covariance structure, thereby affecting the accuracy of the prediction. One could consider a robust functional principal component analysis. 3) With a set of validation data, the tuning parameters in the ridge and PLS estimators can be adaptively chosen without re-computing them. 4) In the construction of prediction intervals, we compute pointwise prediction intervals. However, other standard deviations based on functional depth may be considered. 5) With continuous hourly measurements, functional data can be constructed at daily, monthly, quarterly, biannual, and annual levels. This motivates us to consider temporal forecast reconciliation of Athanasopoulos et al. (2024) and Girolimetto et al. (2024) across different data frequencies in the construction of functional data for continuously measured data sets.

References

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N. & Panagiotelis, A. (2024), 'Forecast reconciliation: A review', *International Journal of Forecasting* **40**(2), 430–456.

Aue, A., Norinho, D. D. & Hörmann, S. (2015), 'On the prediction of stationary functional time series', *Journal of the American Statistical Association: Theory and Methods* **110**(509), 378–392.

Baerenbold, O., Meis, M., Martinez-Hernandez, I., Euán, C., Burr, W. S., Tremper, A., Fuller, G., Pirani, M. & Blangiardo, M. (2023), 'A dependent Bayesian Dirichlet process model for source apportionment of particle number size distribution', *Environmetrics* **34**(1), e2763.

Beevers, S., Carslaw, D., Westmoreland, E. & Mittal, H. (2009), Air pollution and emissions trends in London, Technical report, King's College London and the Institute for Transport Studies.
URL: https://uk-air.defra.gov.uk/reports/cat05/1004010934_MeasurementvsEmissionsTrends.pdf

Benatia, D., Carrasco, M. & Florens, J.-P. (2017), 'Functional linear regression with functional response', *Journal of Econometrics* **201**(2), 269–291.

Bosq, D. (2000), *Linear Processes in Function Spaces*, Lecture notes in Statistics, New York.

Chang, J., Fang, Q., Qiao, X. & Yao, Q. (2025), 'On the modeling and prediction of high-dimensional functional time series', *Journal of the American Statistical Association: Theory and Methods* **in press**.

Chatfield, C. (1993), 'Calculating interval forecasts', *Journal of Business and Economic Statistics* **11**(2), 121–135.

Chatfield, C. (2000), *Time-Series Forecasting*, 1st edn, Chapman and Hall/CRC, New York.

Chiou, J.-M. (2012), 'Dynamical functional prediction and classification with application to traffic flow prediction', *Annals of Applied Statistics* **6**(4), 1588–1614.

Chiou, J.-M., Chen, Y.-T. & Yang, Y.-F. (2014), 'Multivariate functional principal component analysis: A normalization approach', *Statistica Sinica* **24**, 1571–1596.

Dhillon, G. S., Deligiannidis, G. & Rainforth, T. (2024), On the expected size of conformal prediction sets, *in* 'Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS)', Vol. 238, Valencia, Spain.
URL: https://proceedings.mlr.press/v238/dhillon24a/dhillon24a.pdf

Diebold, F. X. & Mariano, R. S. (2002), 'Comparing predictive accuracy', *Journal of Business & Economic Statistics* **20**(1), 134–144.

Dong, M., Yang, D., Kuang, Y., He, D., Erdal, S. & Kenski, D. (2009), '$PM_{2.5}$ concentration prediction using hidden semi-Markov model-based time series data mining', *Expert Systems with Applications* **36**(5), 9046–9055.

Gao, Y., Shang, H. L. & Yang, Y. (2019), 'High-dimensional functional time series forecasting: An application to age-specific mortality rates', *Journal of Multivariate Analysis* **170**, 232–243.

Girolimetto, D., Athanasopoulos, G., Di Fonzo, T. & Hyndman, R. J. (2024), 'Cross-temporal probabilistic forecast reconciliation: Methodological and practical issues', *International Journal of Forecasting* **40**(3), 1134–1151.

Gneiting, T. & Katzfuss, M. (2014), 'Probabilistic forecasting', *Annual Review of Statistics and Its Application* **1**, 125–151.

Gneiting, T. & Raftery, A. E. (2007), 'Strictly proper scoring rules, prediction and estimation', *Journal of the American Statistical Association: Review Article* **102**(477), 359–378.

Gromenko, O., Kokoszka, P. & Reimherr, M. (2017), 'Detection of change in the spatiotemporal mean function', *Journal of the Royal Statistical Society: Series B* **79**(1), 29–50.

Guo, S., Qiao, X., Wang, Q. & Wang, Z. (2025), 'Factor modelling for high-dimensional functional time series', *Journal of Business & Economic Statistics* **in press**.

Hall, P. & Vial, C. (2006), 'Assessing the finite dimensionality of functional data', *Journal of the Royal Statistical Society (Series B)* **68**(4), 689–705.

Hoerl, A. & Kennard, R. (1970), 'Ridge regression: Biased estimation for nonorthogonal problems', *Technometrics* **12**(1), 55–67.

Horváth, L. & Kokoszka, P. (2012), *Inference for Functional Data with Applications*, Springer, New York.

Horvath, L., Kokoszka, P. & Rice, G. (2014), 'Testing stationarity of functional time series', *Journal of Econometrics* **179**(1), 66–82.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E. & Yasmeen, F. (2025), *forecast: Forecasting functions for time series and linear models*. R package version 8.24.0.
URL: https://cran.r-project.org/web/packages/forecast/index.html

Hyndman, R. J., Booth, H. & Yasmeen, F. (2013), 'Coherent mortality forecasting: the product-ratio method with functional time series models', *Demography* **50**(1), 261–283.

Hyndman, R. J., Koehler, A., Ord, K. & Snyder, R. (2008), *Forecasting with Exponential Smoothing: the State Space Approach*, Springer, Berlin ; London.

Hyndman, R. J. & Shang, H. L. (2009), 'Forecasting functional time series (with discussions)', *Journal of the Korean Statistical Society* **38**(3), 199–221.

Jiao, S., Aue, A. & Ombao, H. (2023), 'Functional time series prediction under partial observation of the future curve', *Journal of the American Statistical Association: Theory and Methods* **118**(541), 315–326.

Jiménez-Varón, C. F., Sun, Y. & Shang, H. L. (2024), 'Forecasting high-dimensional functional time series: Application to sub-national age-specific mortality', *Journal of Computational and Graphical Statistics* **33**(4), 1160–1174.

Jiménez-Varón, C. F., Sun, Y. & Shang, H. L. (2025), 'Forecasting density-valued functional panel data', *Australian & New Zealand Journal of Statistics* **in press**.

Klepsch, J. & Klüppelberg, C. (2017), 'An innovations algorithm for the prediction of functional linear processes', *Journal of Multivariate Analysis* **155**, 252–271.

Klepsch, J., Klüppelberg, C. & Wei, T. (2017), 'Prediction of functional ARMA processes with an application to traffic data', *Econometrics and Statistics* **1**, 128–149.

Kokoszka, P. & Reimherr, M. (2013), 'Determining the order of the functional autoregressive model', *Journal of Time Series Analysis* **34**(1), 116–129.

Kokoszka, P. & Reimherr, M. (2017), *Introduction to Functional Data Analysis*, Chapman and Hall/CRC, New York.

Leng, C., Li, D., Shang, H. L. & Xia, Y. (2025), Covariance function estimation for high-dimensional functional time series with dual factor structures, Working paper, arXiv.
URL: https://arxiv.org/abs/2401.05784

Li, D., Robinson, P. M. & Shang, H. L. (2020), 'Long-range dependent curve time series', *Journal of the American Statistical Association: Theory and Methods* **115**(530), 957–971.

Martinez-Hernandez, I., Euán, C., Burr, W. S., Meis, M., Blangiardo, M. & Pirani, M. (2025), 'Modelling particle number size distribution: A continuous approach', *Journal of the Royal Statistical Society Series C: Applied Statistics* **74**(1), 229–248.

Martinez-Hernandez, I. & Genton, M. G. (2023), 'Surface time series models for large spatio-temporal datasets', *Spatial Statistics* **53**, 100718.

Morris, J. S. & Carroll, R. J. (2006), 'Wavelet-based functional mixed models', *Journal of the Royal Statistical Society: Series B* **68**, 179–199.

Morris, J. S., Vannucci, M., Brown, P. J. & Carroll, R. J. (2003), 'Wavelet-based non-parametric modeling of hierarchical functions in colon carcinogenesis', *Journal of the American Statistical Association: Applications and Case Studies* **98**(463), 573–583.

Otto, S. & Salish, N. (2025), Approximate factor models for functional time series, Working paper, arXiv.
URL: https://arxiv.org/abs/2201.02532

Paschalidou, A. K., Karakitsios, S., Kleanthous, S. & Kassomenos, P. A. (2011), 'Forecasting hourly PM_{10} concentration in Cyprus through artificial neural networks and multiple regression model: Implications to local environmental management', *Environmental Science and Pollution Research* **18**(2), 316–327.

Ramos-Carreño, C., Torrecilla, J. L., Carbajo-Berrocal, M., Marcos, P. & Suárez, A. (2024), 'scikit-fda: A Python package for functional data analysis', *Journal of Statistical Software* **109**(2).

Ramsay, J. O. & Silverman, B. W. (2005), *Functional Data Analysis*, 2nd edn, Springer, New York.

Rice, J. & Silverman, B. W. (1991), 'Estimating the mean and covariance structure nonparametrically when the data are curves', *Journal of the Royal Statistical Society (Series B)* **53**(1), 233–243.

Romano, Y., Patterson, E. & Candès, E. J. (2019), Conformal quantile regression, *in* '33rd Conference on Neural Information Processing System (NeurIPS 2019)', Vancouver, Canada.
URL: https://proceedings.neurips.cc/paper_files/paper/2019/file/5103c3584b063c431bd1268e9b5e76fb-Paper.pdf

Shafer, G. & Vovk, V. (2008), 'A tutorial on conformal prediction', *Journal of Machine Learning Research* **9**, 371–421.

Shang, H. L. (2016), 'Mortality and life expectancy forecasting for a group of populations in developed countries: A multilevel functional data method', *Annals of Applied Statistics* **10**(3), 1639–1672.

Shang, H. L. & Haberman, S. (2025), 'Constructing prediction intervals for the age distribution of deaths', *Scandinavian Actuarial Journal* **in press**.

Shang, H. L. & Kearney, F. (2022), 'Dynamic functional time-series forecasts of foreign exchange implied volatility surfaces', *International Journal of Forecasting* **38**(3), 1025–1049.

Slini, T., Kaprara, A., Karatzas, K. & Moussiopoulos, N. (2006), 'PM_{10} forecasting for Thessaloniki, Greece', *Environmental Modelling & Software* **21**(4), 559–565.

Stadlober, E., Hörmann, S. & Pfeiler, B. (2008), 'Quality and performance of a PM10 daily forecasting model', *Atmospheric Environment* **42**(6), 1098–1109.

Tavakoli, S., Nisol, G. & Hallin, M. (2023), 'Factor models for high-dimensional functional time series II: Estimation and forecasting', *Journal of Time Series Analysis* **44**, 601–621.

Wiedensohler, A., Birmili, W., Nowak, A., Sonntag, A., Weinhold, K., Merkel, M., Wehner, B., Tuch, T., Pfeifer, S., Fiebig, M. et al. (2012), 'Mobility particle size spectrometers: Harmonization of technical standards and data structure to facilitate high quality long-term observations of atmospheric particle number size distributions', *Atmospheric Measurement Techniques* **5**(3), 657–685.

Zhang, X. & Wang, J.-L. (2016), 'From sparse to dense functional data and beyond', *The Annals of Statistics* **44**(5), 2281–2321.