*Journal of the Chinese
Statistical Association
Vol. 63, (2025) 276–296*

Multi-year forecasting of annual load curves in Taiwan

**Hsiang-Ling Hsu[1†], Chih-Hao Chang[2],
Chi-Hsiang Chu[1], Chun-Feng Zhang[1],
Zi-Wen You[1], Hong-Ting Chen[3] and Guan-Jhang Chen[3]**

[1] Institute of Statistics, National University of Kaohsiung, Kaohsiung, Taiwan
[2] Department of Statistics, National Chengchi University, Taipei, Taiwan
[3] Industrial Technology Research Institute, Green Energy and Environment Research Laboratories, Hsinchu, Taiwan

ABSTRACT

In the power system, ensuring supply security and stability is crucial. Dispatch units conduct long-term power development planning considering factors like load growth. According to current regulations in Taiwan, a 15% reserve capacity rate is mandated. Presently, domestic capacity assessment models rely on historical data for load simulation. While simple and expedient, this approach overlooks uncertainties and future climate changes. This paper proposes a two-stage method for estimating future annual load curves under statistical frameworks, drawing from existing methods in the literature. This will aid government agencies in conducting comprehensive long-term power supply-demand assessments, considering uncertainties and enhancing estimation accuracy.

Key words and phrases: Functional time series analysis, Load prediction, Regression models.

JEL classification: L94, C32, E27.

[†]Corresponding to: Hsiang-Ling Hsu
E-mail: hsuhl@nuk.edu.tw

1. Introduction

Load forecasting plays a crucial role in modern power systems. With the development of society and the economy, the demand for electricity continues to increase, making accurate predictions of future load trends essential. Load forecasting helps electricity companies to plan production and supply schedules effectively, ensuring the needs of users are met while avoiding energy waste and supply shortages. Furthermore, load forecasting guides the construction and expansion of power infrastructure to better meet future demands; see Lindberg *et al.* (2019) for a comprehensive review.

In recent years, machine learning techniques have been widely applied in the field of load forecasting. Particularly, the emergence of nonlinear models such as neural networks has provided more possibilities for load prediction. Imani (2021) proposes a convolutional neural network method to extract the nonlinear relationships among the load-temperature data. Bian *et al.* (2022) construct the relationship between the load and temperature data via convolution neural network method for short-term load forecasting. These models can better capture the complex relationships between load and factors such as time, weather, and holidays, thereby improving prediction accuracy. In practical applications, many power companies have adopted these new technologies and achieved significant results, making power supply more stable and efficient; see Mamun *et al.* (2020) for more details.

Despite the impressive progress of artificial intelligence techniques in load forecasting, their interpretability remains a challenge. In contrast, advanced statistical methods offer superior explanatory power for load variations. Elías *et al.* (2022) introduce a nonparametric analysis approach for functional time series data, known as the projection method, which predicts future functional time data through weighted averaging of historical functional time data. Shang (2013) develop an R package for the application of a functional time series analysis (FTSA) method to analyze time series data where each observation is a curve or function, allowing for flexible and comprehensive exploration of temporal patterns and trends. Additionally, Chang *et al.* (2022) propose a high-resolution time series method in forecasting high-resolution electricity consumption. Statistical methods, grounded in physical models and statistical theory, provide clear explanations of load variations, enhancing understanding and confidence in load forecasting results. Consequently, in load forecasting, statistical methods complement artificial intelligence techniques, collectively advancing the development and progress of power systems.

In this article, we address the multi-year prediction of annual electricity load in Taiwan. Leveraging the FTSA method introduced by Shang (2013), we propose a two-stage approach for predicting functional time series curves with exogenous covariates. Firstly, we employ a power transformation of the residuals from the FTSA model and subsequently regress these transformed residuals with the exogenous covariates. This process establishes a nonlinear relationship between the response and covariate curves, enhancing prediction performance. Refer to Algorithm regFTSA in Section 3 for a detailed explanation. We apply this method to a dataset comprising annual temperature-load curves collected in Taiwan from 2010 to 2023. As demonstrated in Section 4, our approach yields a significant improvement in multi-year load curve prediction compared to solely utilizing the FTSA method.

The paper is organized as follows. Section 2 introduces the notations for the functional time series model. Algorithms for the estimating the functional time series models are given in Section 3. A comprehensive study for the annual load curve prediction based on the proposed method is given in Section 4. Finally, we offer a brief discussion in Section 5.

2. Models and notations

Let $\mathcal{D} = \{(y_t(x_n), \boldsymbol{z}_t(x_n))\}_{n=1,\ldots,N,\, t=1,\ldots,T}$ be the load dataset, where the load response and the q-dimensional exogenous covariate vector, denoted by $y_t(x_n)$ and $\boldsymbol{z}_t(x_n)$, respectively, are collected in the n-th time index, denoted by x_n, of the t-th year for $n = 1, \ldots, N$ and $t = 1, \ldots, T$. We consider the functional time series model:

$$y_t(x_n) = f_t(x_n) + g\left(\epsilon_t(x_n)\right), \tag{1}$$

where $f_t(x_n)$ is the smooth trend function of x_n for load curve in the t-th year, and $g(\epsilon_t(x_n))$ is the measurement error of $y_t(x_n)$ with $\epsilon_t(x_n)$ being white noise and $g(\epsilon_t(x_n))$

is unknown transformation function of $\epsilon_t(x_n)$ that could dependent on $\boldsymbol{z}_t(x_n)$. Note that the exogenous covariates in $\boldsymbol{z}_t(x_n)$ could be key factors for modeling $f_t(x_n)$ or $g(\epsilon_t(x_n))$. For example, the extremely high or extremely low temperatures can enlarge the variance of the electricity consumption. In this case, we suggest to model $g(\epsilon_t(x_n))$ as a function of $\boldsymbol{z}_t(x_n)$.

Shang (2013) considers the functional principal component regression (FPCR) method to estimate the underlying mean trend $f_t(x_n)$ of the functional curve $y_t(x_n)$ when $g(\epsilon_t(x_n)) = \epsilon_t(x_n)$ with $\text{var}(\epsilon_t(x_n)) = \sigma_t^2(x_n)$ that is independent of $\boldsymbol{z}_t(x_n)$. Ideally, the population mean trend of Shang (2013) is modeled by

$$f_t(x) = \mu(x) + \sum_{k=1}^{\infty} \beta_{t,k}\phi_k(x), \tag{2}$$

where $\mu(x)$ is the constant mean trend function that is independent of t, $\beta_{t,k}$ is the k-th principal component score of the t-th year, and $\phi_k(x)$ is the k-th population functional principal component. However, in practice, due to only finite T realizations of $y_t(x_n)$ are observed, the sieve method is considered to approximate (2) by only a few function principal components. Technically speaking, the functional principal component decomposition of (2) is approximated by

$$\hat{f}_t(x_n) = \hat{\mu}(x_n) + \sum_{k=1}^{K} \hat{\beta}_{t,k}\hat{\phi}_k(x_n), \tag{3}$$

where $\hat{\mu}(x_n) = \sum_{t=1}^{T} y_t(x_n)/T$ is the estimated constant mean function, the coefficient $\hat{\beta}_{t,k}$ is the k-th principal component score for t-th year, $\hat{\phi}_k(x_n)$ is the k-th estimated orthogonal eigen-function of the empirical covariance operator:

$$\hat{\Gamma}(x_n, x_m) \equiv \frac{1}{T}\sum_{t=1}^{T} \big(y_t(x_n) - \hat{\mu}(x_n)\big)\big(y_t(x_m) - \hat{\mu}(x_m)\big) \tag{4}$$

with $n, m \in \{1, \ldots, N\}$, and K represents the number of selected principal components, which is a positive integer. The selection of K can be determined through cross-validation methods, such as rolling or expanding windows, to evaluate the predictive performance for different values of K. In our study, we selected K using the eigenvalue ratio and growth ratio methods, aiming to balance model complexity and accuracy by identifying the minimum number of components that capture substantial variance in the data. In this article, we used the `ER_GR()` function from the FTSA package in R to select K for data analysis. Note that $\hat{\beta}_{t,k}$ is given by the least-squared method for the t-th year:

$$\hat{\boldsymbol{\beta}}_t = (\boldsymbol{F}_K'\boldsymbol{F}_K)^{-1}\boldsymbol{F}_K'(\boldsymbol{y}_t - \hat{\boldsymbol{\mu}}), \tag{5}$$

where $\hat{\boldsymbol{\beta}}_t = (\hat{\beta}_{t,1}, \cdots, \hat{\beta}_{t,K})'$, $\boldsymbol{F}_K$ denotes the $N \times K$ matrix, whose (n,k)-th element is $\hat{\phi}_k(x_n)$, $\boldsymbol{y}_t = (y_t(x_1), \cdots, y_t(x_N))'$ and $\hat{\boldsymbol{\mu}} = (\hat{\mu}(x_1), \cdots, \hat{\mu}(x_N))'$. In addition, based on the fixed functional principal components $\hat{\phi}_1(x_n), \dots, \hat{\phi}_K(x_n)$ and $\boldsymbol{y}(x_n) = (y_1(x_n), \dots, y_T(x_n))'$, the h-step-ahead forecasts of $y_{T+h}(x_n)$ can be obtained as:

$$\hat{y}_{T+h}(x_n) = \hat{\mu}(x_n) + \sum_{k=1}^{K} \hat{\beta}_{T+h|T,k}\hat{\phi}_k(x_n), \tag{6}$$

where $\hat{\beta}_{T+h|T,k}$ is the h-step-ahead forecasts of $\beta_{T+h|T,k}$ and obtained by fitting a univariate time series to $\hat{\beta}_{1,k}, \dots, \hat{\beta}_{T,k}$. In this article, we utilize the autoregressive integrated moving average (ARIMA) model to estimate $\hat{\beta}_{T+h|T,k}$. The ARIMA model combines three components: autoregression (AR), differencing (I) to make the data stationary, and moving average (MA) of past errors. It is commonly used for predicting future points in time series data by capturing its trends and patterns. In this article, the order parameters are driven by the information criterion that is provided by `auto.arima()` `function in the FORECAST package of R for the data analysis.`

Later in Section 3, we will introduce our proposed method for regressing the covariate vector $\boldsymbol{z}_t(x_n)$ onto a transformation of the residuals, derived from the previously mentioned functional time series analysis (FTSA) method proposed by Shang (2013).

3. The proposed method

In this article, we assume the model of $g(\epsilon_t(x_n))$ given in (1) to be dependent on $\boldsymbol{z}_t(x_n)$. To address potential heteroscedasticity in the data, we initially apply the Box-Cox transformation to $g(\epsilon_t(x_n))$. This transformation serves to approximate a normal distribution for the residuals, thereby stabilizing the variance and enhancing the accuracy of the subsequent regression analysis. Without loss of generality, we assume the existence of a constant $c > 0$ such that $g(\epsilon_t(x_n)) + c > 0$ ensuring the applicability of the Box-Cox transformation. That is

$$\tilde{e}_{t,\lambda}(x_n) = \begin{cases} \left\{(g(\epsilon_t(x_n)) + c)^\lambda - 1\right\}/\lambda, & \text{if } \lambda \neq 0, \\ \log(g(\epsilon_t(x_n)) + c), & \text{if } \lambda = 0. \end{cases} \tag{7}$$

We then model $\tilde{e}_{t,\lambda}(x_n)$ as

$$\mathrm{E}\big(\tilde{e}_{t,\lambda}(x_n)|\boldsymbol{z}_t(x_n)\big) \equiv b_0 + \boldsymbol{b}'\boldsymbol{z}_t(x_n), \tag{8}$$

where b_0 and $\boldsymbol{b}$ denotes the intercept and regression coefficient vector of the variance model. Note that for an estimator of $f_t(x_n)$, say $\hat{f}_t(x_n)$, the residual $e_t(x_n) = y_t(x_n) - \hat{f}_t(x_n)$ can be an estimator of $g(\epsilon_t(x_n))$; hence b_0 and $\boldsymbol{b}$ of (8) can be estimated by regressing the power transformation of the residuals over $\boldsymbol{z}_t(x_n)$. Note that for the case when $c = 0$ and $\lambda = 2$, we are close to fitting a regression model of $\boldsymbol{z}_t(x_n)$ to estimate $\mathrm{Var}(g(\epsilon_t(x_n)))$. In practice, the power of the transformation, λ, is estimated by maximum likelihood method for the prediction. Moreover, c can be arbitrarily selected to ensure that $e_t(x_n) + c$, $n = 1, \ldots, N$, are all positive, and the impact of its magnitude can be integrated into the estimation of b_0 and $\boldsymbol{b}$.

We consider a two-stage method to estimate $f_t(x_t)$ of (1) and to predict $g(\epsilon_t(x_n))$ of

Algorithm FTSA

Require: $\mathcal{D}_0 = \{y_t(x_n)\}_{n=1,\ldots,N,\, t=1,\ldots,T}$ with $x_n \in (0, \tau]$ for $\tau > 0$, K as the number of estimated eigen-function, and h as the prefixed steps for the prediction.

Ensure: $\{(\hat{y}_{T+h}(x_n), \hat{f}_t(x_n))\}$ for each $x_n \in (0, \tau]$

function FTSA($\mathcal{D}_0, K, h$)

  Obtain the estimated eigen-function $\hat{\phi}_k(x_n)$ for $k = 1, \ldots, K$ given by (3).

  Obtain the estimate principal component score $\hat{\beta}_{t,k}$ for $k = 1, \ldots, K$ and $t = 1, \ldots, T$ given by (5).

  Obtain the predictor $\hat{\beta}_{T+h|T,k}$ of $\beta_{T+h,k}$ by fitting an ARIMA model with order parameters that are driven by the information criterion aicc that is provided by auto.arima in forecast package of R, to $\hat{\beta}_{1,k}, \ldots, \hat{\beta}_{T,k}$ for $k = 1, \cdots, K$.

  Obtain the predictor $\hat{y}_{T+h}(x_n)$ given by (6) and the estimator $\hat{f}_t(x_n)$ given by (3) with $\hat{\mu}(x_n) = \sum_{t=1}^{T} y_t(x_n)/T$.

  return $\hat{y}_{T+h}(x_n)$ and $\hat{f}_t(x_n)$ for $n = 1, \ldots, N$ and $t = 1, \ldots, T$.

end function

Algorithm regFTSA

Require: $\mathcal{D} = \{(y_t(x_n), \boldsymbol{z}_t(x_n)\}_{n=1,\dots,N,\, t=1,\dots,T}$ with $x_n \in (0, \tau]$ for $\tau > 0$, K as the number of estimated eigen-function, and h as the prefixed steps for the prediction.

Ensure: $\tilde{y}_{T+h}(x_n)$ for each $x_n \in (0, \tau]$

1: **function** REGFTSA$(\mathcal{D}, K, h)$

2: Obtain the predictor $\{(\hat{y}_{T+h}(x_n), \hat{f}_t(x_n))\}_{n=1,\dots,N,\, t=1,\dots,T} \leftarrow \text{FTSA}(\mathcal{D}_1, K, h)$ given by Algorithm FTSA with $\mathcal{D}_1 = \{y_t(x_n)\}_{n=1,\dots,N,\, t=1,\dots,T}$.

3: Obtain the predictor $\{\hat{\boldsymbol{z}}_{T+h}(x_n)\}_{n=1,\dots,N}$ partially from $\text{FTSA}(\mathcal{D}_d, K, h)$, where for $d = 1, \dots, q$, $\mathcal{D}_d = \{z_{t,d}(x_n)\}_{n=1,\dots,N,\, t=1,\dots,T}$ with $z_{t,d}(x_n)$ the d-th element of $\boldsymbol{z}_t(x_n)$.

4: Obtain the estimators of $\tilde{e}_{t,\lambda}(x_n)$ given by (7) with $g(\epsilon_t(x_n))$ being estimated by $e_t(x_n) = y_t(x_n) - \hat{f}_t(x_n)$ for each candidate value of λ.

5: Obtain the LS estimators $(\hat{b}_0(\lambda), \hat{\boldsymbol{b}}(\lambda))$ by fitting the linear regression model of (8) onto $\tilde{e}_{t,\lambda}(x_n)$ for each candidate value of λ.

6: Obtain the predictor of $g(\epsilon_{T+h}(x_n))$, denoted by $\hat{e}_{T+h}(x_n)$, which is obtained by

$$\max(0, \hat{b}_0(\hat{\lambda}) + \hat{\boldsymbol{b}}(\hat{\lambda})'\hat{\boldsymbol{z}}_{T+h}(x_n)) = \begin{cases} \{(\tilde{e}_{T+h}(x_n) + \hat{c})^{\hat{\lambda}} - 1\}/\hat{\lambda}, & \text{if } \hat{\lambda} \neq 0, \\ \log(\tilde{e}_{T+h}(x_n) + \hat{c}), & \text{if } \hat{\lambda} = 0, \end{cases}$$

where $\hat{c} \equiv 1 - \min_{t=1,\dots,T} e_t(x_n)$ is set as default and $\hat{\lambda}$ is given by (9).

7: Obtain the predictor of $y_{T+h}(x_n)$, denoted by $\tilde{y}_{T+h}(x_n)$, which is given by

$$\tilde{y}_{T+h}(x_n) = \hat{y}_{T+h}(x_n) + \hat{e}_{T+h}(x_n), \quad n = 1, \dots, N.$$

8: **return** $\tilde{y}_{T+h}(x_n)$ for $n = 1, \dots, N$.

9: **end function**

(8), respectively. In the first stage of the proposed method, we apply Algorithm FTSA proposed by Shang (2013) to the estimation $f_t(x_n)$ and the prediction of $y_{T+h}(x_n)$, for some $h \in \mathbb{N}$. Next in the second stage, we proposed Algorithm regFTSA to regressing the covariate $\boldsymbol{z}_t(x_n)$ onto the $\tilde{e}_{t,\lambda}(x_n)$ defined in (7). Ideally, the realization of $\boldsymbol{z}_{T+h}(x_n)$ is required for the prediction of $y_{T+h}(x_n)$ for $n = 1, \dots, N$, which is unobservable and shall be predicted. Hence we apply Algorithm FTSA to obtain the predictor $\hat{\boldsymbol{z}}_{T+h}(x_n)$ for the prediction of $y_{T+h}(x_n)$; see Algorithm regFTSA for more details. In addition, for the power transformation of residuals given in (7) with c arbitrary chosen to ensure the existence of $\tilde{e}_{t,\lambda}(x_n)$ in (7) for any $\lambda \in \mathbb{R}$, λ is selected as the optimizer

$$\hat{\lambda} = \arg\min_{\lambda \in \mathbb{R}} \frac{1}{T} \sum_{t=1}^{T} \left(\tilde{e}_{t,\lambda}(x_n) - \left(\hat{b}_0(\lambda) + \hat{\boldsymbol{b}}(\lambda)'\boldsymbol{z}_t(x_n)\right)\right)^2, \tag{9}$$

where $\hat{b}_0(\lambda)$ and $\hat{\boldsymbol{b}}(\lambda)$ are the regression coefficient estimators corresponding to $\lambda \in \mathbb{R}$. We used the `boxcox()` function in MASS package of R to select λ of (9), where the default setting of `boxcox()` selects λ from a set ranging from -2 to 2 in steps of 0.5. In this article, we choose λ from a larger set ranging from -5 to 5 in steps of 0.01 for the data analysis.

To assess the predictive performance of our proposed method, we employ three commonly used metrics: (i) the mean absolute percentage error (MAPE), (ii) the mean absolute error (MAE), and (iii) the root mean square error (RMSE), which are widely utilized in electricity demand forecasting models; see, for example, Filik *et al.* (2011), Moroff *et al.* (2021), and Shah *et al.* (2022). These metrics are defined as follows:

$$\begin{aligned}
\mathrm{MAPE}(h) &= \frac{100\%}{N}\sum_{n=1}^{N}\left|\frac{\tilde{y}_{T+h}(x_n) - y_{T+h}(x_n)}{y_{T+h}(x_n)}\right|, \\
\mathrm{MAE}(h) &= \frac{1}{N}\sum_{n=1}^{N}\left|\tilde{y}_{T+h}(x_n) - y_{T+h}(x_n)\right|, \\
\mathrm{RMSE}(h) &= \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\tilde{y}_{T+h}(x_n) - y_{T+h}(x_n)\right)^2}.
\end{aligned} \tag{10}$$

4. Load data analysis

4.1 Electric power load data in Taiwan

This article focuses on analyzing an electric power load dataset collected in Taiwan from 2010 to 2023, provided by the Taiwan Power Company. The dataset comprises nationwide electricity consumption and temperature variables recorded at 15 minute time intervals. Temperature data has long been recognized as highly related to load data, making it a crucial factor considered in load prediction literature; see, for example, Imani (2021) and Bian *et al.* (2022). Temperatures were recorded in degrees Celsius (°C), while load was measured in megawatts per hour (MWh) from 2010 to 2023.

The data is preprocessed to check for missing values, aggregate the data, and perform necessary transformations. To fill in missing values in the data, both linear and nonlinear interpolation methods are commonly used. In this paper, based on the stability, simplicity, and fast computation of linear interpolation, we adopt linear

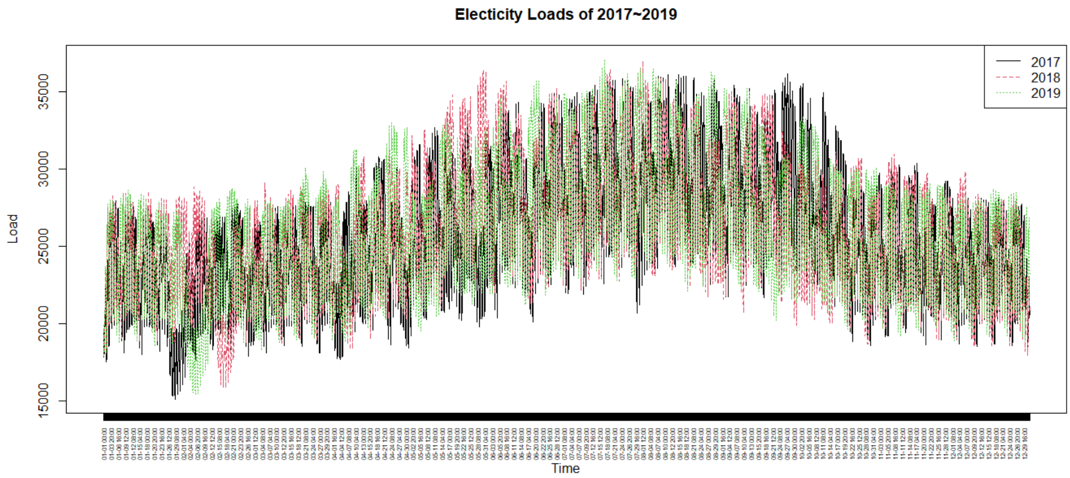
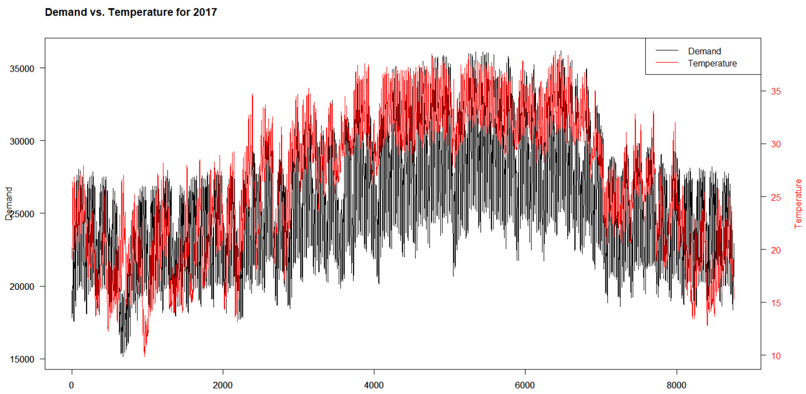
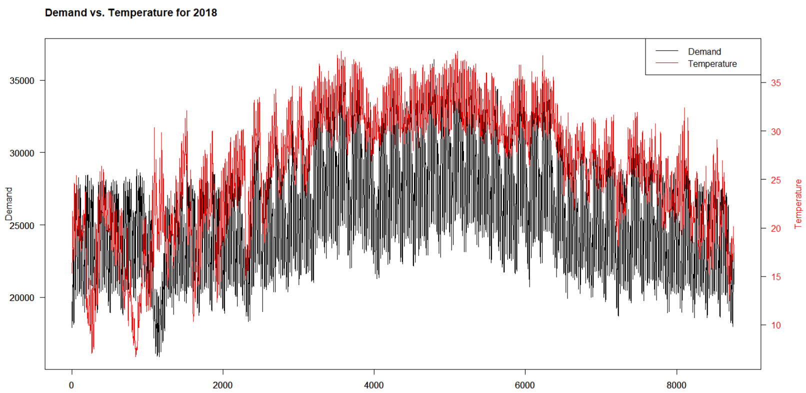


Figure 1: The annual load curves in Taiwan for 2017 to 2019.

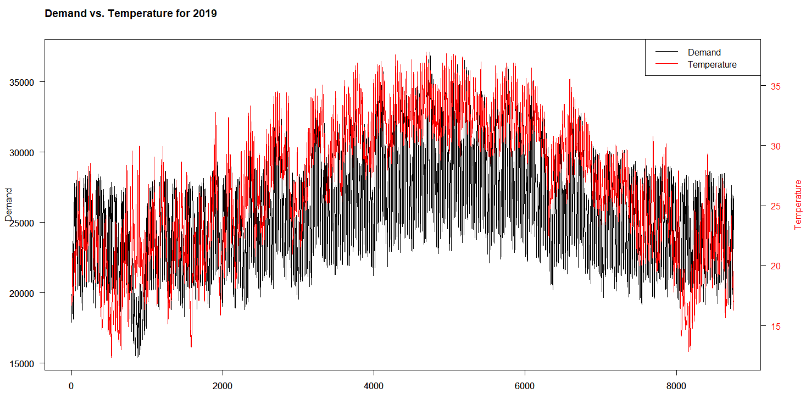
interpolation for missing data in each variable of the dataset. Specifically, the missing value for a given time period is replaced by the weighted average of the observed values from adjacent periods. If there are consecutive missing periods within a day, the missing values are filled by averaging the data from the same periods of the previous and following days. Additionally, since this study focuses on the performance of the annual hourly demand profiles that is 8760 hours per year, we aggregate all variables by taking the hourly averages. For prediction stability consideration, we standardize the logarithm of the hourly demand for each segmented dataset. The annual hourly demand profiles from 2017 to 2019 are displayed in Figure 1, showing that the power load curves across different years follow similar trends. Load tends to be lower during the Lunar New Year, rises significantly in the summer, and drops during the winter, indicating a seasonal or monthly periodicity in the power load curves. The time series and scatter plots of hourly temperature and demand is given in Figures 2 and Figures 3. The temperature curve and the load curve exhibit similar periodic variations as shown in Figure 2, with higher values in the summer and relatively lower values in the winter. Therefore, we consider temperature as the exogenous variable in Algorithm regFTSA for the load curve prediction.



(a) 2017

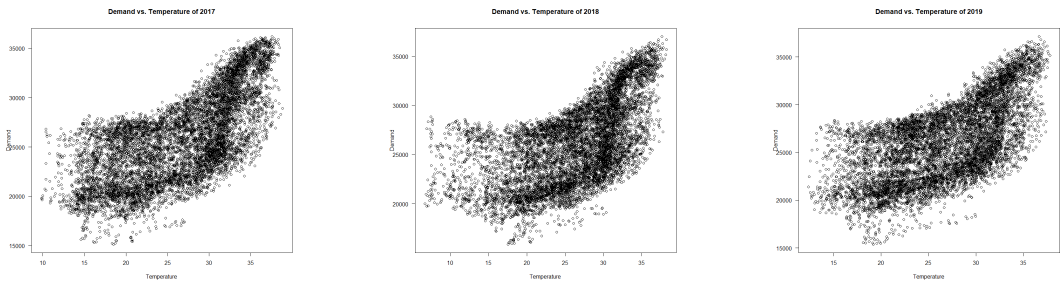


(b) 2018



(c) 2019

Figure 2: Time series plot of the demand and temperature from 2017 to 2019.



(a) 2017 (b) 2018 (c) 2019

Figure 3: The scatter plots of the demand and temperature from 2017 to 2019.

4.2 Segmented forecasting comparison analysis for annual load curves

Due to the high periodicity inherent in load data, the prediction of load curves can be performed on weekly, monthly, quarterly, or yearly scales. Therefore, we split the training dataset into four different time periods of sub-datasets: weekly, monthly, quarterly, and yearly, respectively, for applying to prediction and evaluation procedures. We apply the `ER_GR()` function in FTSA package of R to select K, the number of eigenfunctions, for Algorithms FTSA and Algorithms regFTSA. Among all the scenarios, $K = 1$ is selected for the data analysis.

Consider the dataset $\mathcal{D}^{(s)} = \left\{\left(y^{(s)}_{2009+t}(x_n), z^{(s)}_{2009+t}(x_n)\right)\right\}$, where $t = 1, \ldots, 14$ representing the year index from 2010 to 2023; s ranges from 1 to S, representing the segmented sub-dataset index which maximum index is denoted by S; and x_n ranges from 1 to N, representing the time index within each segmented sub-dataset. Each entry $\left(y^{(s)}_{2009+t}(x_n), z^{(s)}_{2009+t}(x_n)\right)$ denotes a temperature-load sample observed at the n-th hour in the s-th segmented sub-dataset of year $2009+t$, where $y_{2009+t}(x_n)$ represents the hourly load response and $z_{2009+t}(x_n)$ represents the hourly temperature covariate. To evaluate the predictive performance of the proposed method, we split the data into a training set spanning from 2010 to 2019, totaling 10 years, and a test set comprising the years 2020 to 2023, totaling four years of testing data. Let $\hat{y}^{(s)}_{2009+T+h}(x_n)$ and $\tilde{y}^{(s)}_{2009+T+h}(x_n)$ with $T = 10$, $h = 1, 2, 3, 4$ and $x_n = 1, \ldots, N$, denotes the h-step-ahead prediction curves for each k by Algorithm FTSA and Algorithm regFTSA with $K = 1$ for both algorithms We then employ MAPE metric given in (10) to assess the predictive effectiveness.

However, as demonstrated in Algorithm FTSA and Table 1, when predictions are made on an annual basis, despite having a full year of data available for estimating regression coefficients more precisely, it is assumed that the regression coefficients for each hour of data within that year are the same parameter vector. This assumption, particularly over long time scales, carries a relatively high risk of not being met, thereby resulting in model errors. Consequently, in our subsequent research, we segment the annual data into weekly curves and make predictions based on multi-year data for each weekly curve. On the other hand, the better forecasting results in Table 1 are acquired based on the weekly segmented modeling procedure. Subsequently, we will adopt the multi-step weekly predicted approach to get the forecast weekly curves and then aggregate those curves into an annual predicted curve.

The segmented method is commonly used to improve the prediction of time series curves, as exemplified in Spencer *et al.* (2019). The segmented method leverages the FTSA method to fit weekly curves, thereby increasing model complexity and consequently enhancing estimation performance. However, this approach runs the risk of overfitting if the variation in weekly curves can be fully captured by the empirical functional principal components ($\hat{\phi}_k(x_n)$'s). In such cases, the principal component scores ($\hat{\beta}_{t,k}$'s) may become independent of x_n. Achieving a good approximation of the underlying mean trend $f_t(x_n)$ may require a large number of selected principal components (K), necessitating a larger sample size (T). To mitigate the dependence on the precision estimation of $\hat{\phi}_k(x_n)$ and the selection of K when T is small, we propose weekly segmentation for predicting annual load curves in the subsequent analysis process. Although

Table 1: Summary of MAPE(h) of $\{\hat{y}^{(s)}_{2009+T+h}(x_n)\}$ obtained by the Algorithm FTSA for different segmented forecasting procedures.

Segmented basis \ Year (h)	2020 ($h=1$)	2021 ($h=2$)	2022 ($h=2$)	2023 ($h=4$)
Weekly Basis	5.275	7.839	8.604	7.929
Monthly Basis	8.263	10.055	9.332	7.665
Quarterly Basis	8.189	9.807	9.033	7.099
Annually Basis	12.997	13.729	13.178	11.995

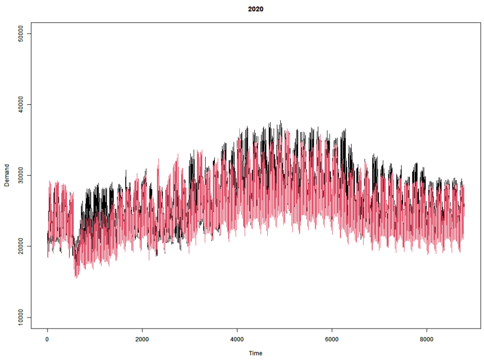
the segmented weekly forecasting method shown here provides satisfactory annual load curve results, there is still room for improvement in the prediction accuracy. The next subsection will concentrate on adjusting the forecasting method based on segmented weekly forecasting considerations and will include the corresponding comparative analysis.

4.3 Prediction performance and adjustments for segmented weekly forecasting

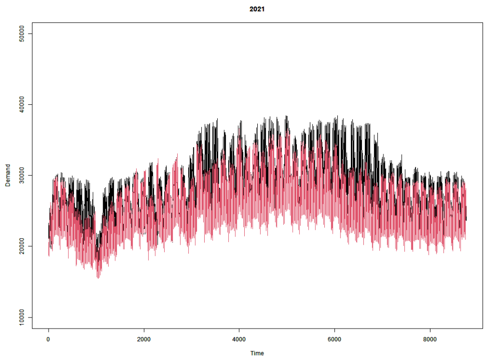
Note that due to the variation in the number of days in each month or year, there are minor differences in the data preprocessing procedures when handling segmented datasets. Since this study primarily focuses on the prediction procedures of weekly sub-datasets, the data preprocessing procedures for other segmented datasets (i.e., monthly, quarterly, or yearly scales) used in the previous subsection will not be elaborated upon. To account for the possibility that the first and last weeks of each year may not span a full seven days, we supplement the first week by retrieving data from the preceding days to complete the seven-day period. Similarly, for the last week of a year, we extend it to a full seven days by retrieving data from subsequent days. However, if there is insufficient data available in the final week, we supplement the missing days by using the data from the corresponding days of the preceding week. With these adjustments, we can acquire weekly individual curves for and then adopt to yearly load prediction.

Therefore, the weekly dataset $\mathcal{D}^{(s)} = \left\{\left(y^{(s)}_{2009+t}(x_n), z^{(s)}_{2009+t}(x_n)\right)\right\}$, where $t = 1, \ldots, 14$, representing the year index of the data from 2010 to 2023; $s = 1, \ldots, 52$ (or 53), representing the week index; and $x_n = 1, \ldots, 168$, the time index within each week. Let $\hat{y}^{(s)}_{2009+T+h}(x_n)$ and $\tilde{y}^{(s)}_{2009+T+h}(x_n)$ with $T = 10$, $h = 1, 2, 3, 4$ and $x_n = 1, \ldots, 168$, denotes the h-step-ahead weekly prediction curves for $s = 1, \ldots, 52$ (or 53) obtained respectively by Algorithm FTSA and Algorithm regFTSA with $K = 1$ for both algorithms, and $\hat{c} = 7319$ and $\hat{\lambda} = 2$ for regFTSA that is selected among $\lambda \in \{-5.00, -4.99, \ldots, 4.99, 5.00\}$ by (9). Based on these weekly prediction curves, we can aggregate those segmented prediction curves to the annual prediction curve corresponding to years 2020∼2023. However, before assessing the aggregated annual prediction load curves, some adjustments would be considered for the prediction curves based on the domain knowledge of the fields for electrical engineering and energy systems.

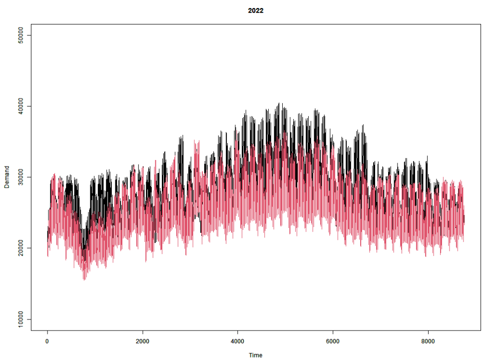
We now introduce the first adjustment for the aggregated annual prediction load curves. Due to significant differences in load curve performance between New Year's Day and the Lunar New Year period compared to regular days, based on this electricity usage characteristic, we reforecast each year's single day of New Year's Day (with 24-hour sample observation points) and the Lunar New Year's Eve to the fifth day of the first lunar month (with 144-hour sample observation points) for the first adjustment. After forecasting for each week within the year, when integrating into the overall annual prediction curve, we then replace the corresponding New Year's Day and Lunar New Year periods with the forecasted values. We then denote the adjusted annual predicted curves by and $\hat{y}_{T+h}(x_n^{(h)})$ and $\tilde{y}_{T+h}(x_n^{(h)})$ as the aggregated annual prediction curve corresponding to 2020∼2023 corresponding to Algorithm FTSA and Algorithm regFTSA, where $x_n^{(h)}$ ranges from the first hour of the New Year's Day to the last hour of December 31 for $2019+h$ year. For the prediction performance comparison with the proposed method, the baseline year involves either extending or truncating the curve of the last year in the training data to match the length of the prediction year, and then directly using it as the prediction curve for the prediction year. For the prediction performance comparison with the proposed method, we use the data from the last year in the training set (i.e., 2019) as the prediction curve for the future years (2020∼2023), referring to this method as simple prediction. Since the length or number of weeks in the last year's data may differ from the future years, we adjust by extending or truncating the data as needed.Then we employ evaluation metrics given in (10) to assess the predictive effectiveness. In order to provide long-term forecast results for reference by government agencies and facilitate the establishment of power generation facilities, we conducted curve predictions for the years 2020∼2023 as depicted in Figures 4 and Figures 5 using the trained model obtained by FTSA and regFTSA procedures. Table 2 reveals that both the FTSA and regFTSA methods outperform the simple prediction method. Particularly, after adjusting for the temperature exogenous variable, the predictive performance of the regFTSA method surpasses that of the FTSA method alone across all evaluation metrics.



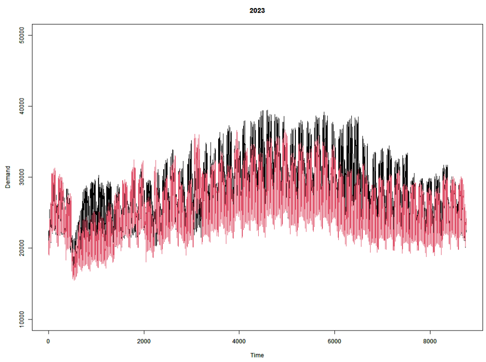
(a) 2020



(b) 2021



(c) 2022

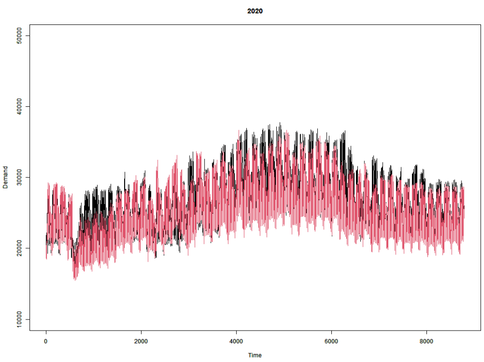


(d) 2023

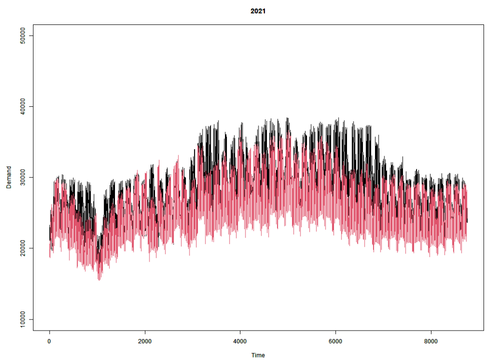
Figure 4: The predicted yearly curves of the electricity demand from 2020 to 2023 obtain by Algorithm FTSA, where the black curve is the true electricity demand and the red curve is the predicted electricity demand.

Table 2: Summaries of evaluation results under different metrics of the predicted curves obtained by the Algorithm FTSA and regFTSA.

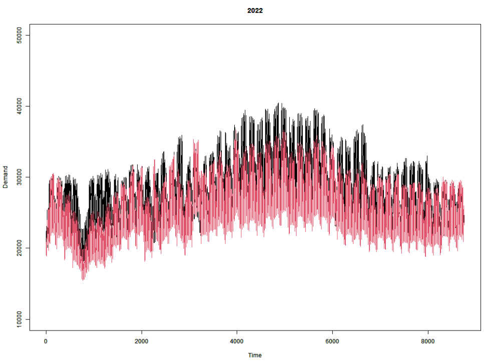
Evaluation Metric	Year / Algorithm	2020 ($h=1$)	2021 ($h=2$)	2022 ($h=2$)	2023 ($h=4$)
MAPE(h)(%)	Simple Prediction	8.669	10.597	10.338	9.894
	FTSA	5.275	7.839	8.604	7.929
	regFTSA	5.209	7.730	8.497	7.838
MAE(h)	Simple Prediction	2332.332	3022.643	2996.042	2809.695
	FTSA	1419.170	2223.619	2450.019	2241.954
	regFTSA	1400.114	2191.778	2418.516	2215.336
RMSE(h)	Simple Prediction	3095.794	3797.709	3758.075	3671.347
	FTSA	1835.906	2729.355	2878.523	2765.966
	regFTSA	1816.464	2398.794	2847.177	2738.816



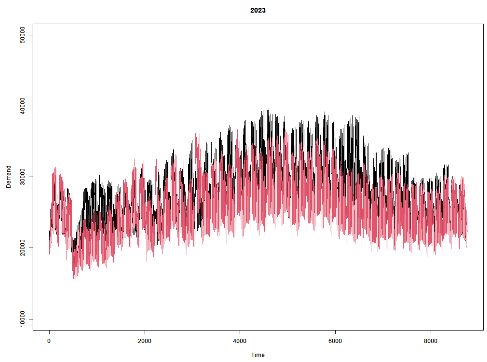
(a) 2020



(b) 2021



(c) 2022



(d) 2023

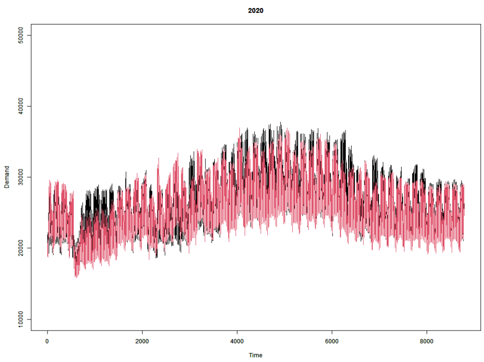
Figure 5: The predicted yearly curves of the electricity demand from 2020 to 2023 obtain by Algorithm regFTSA, where the black curve is the true electricity demand and the red curve is the predicted electricity demand.

In recent years, in the post-pandemic era, with the revival of technology and related industries, we have observed a year-on-year increase in average load in Taiwan. However, in (3), treating the average curve of all years as a constant value (i.e, $\hat{\mu}(x_n)$) is evidently contrary to the growth trend of Taiwan's electricity average load. Additionally, Figures 4 and Figures 5 clearly show a general underestimation trend as h increases. Therefore, we make the final adjustment by correcting the bias in the underestimation of the load prediction curves. Let $\bar{y}_{2009+t}$ denote the average of the annual load in year $2009 + t$. In addition, let $d^{(h)}_{2009+t} = \bar{y}_{2009+t+h} - \bar{y}_{2009+t}$, $t = 1, \dots, T - h$. Hence the h-step-ahead predicted bias, denoted by $\hat{d}^{(h)}_{2009+T+h}$, is given as the one-step-ahead predictor of an ARIMA model for $d^{(h)}_{2009+t} = \bar{y}_{2009+t+h} - \bar{y}_{2009+t}$, $t = 1, \dots, T - h$ with order parameters that are determined by the information criterion aicc that is provided by auto.arima in forecast package of R. We then obtain the debiased annual prediction load curves by

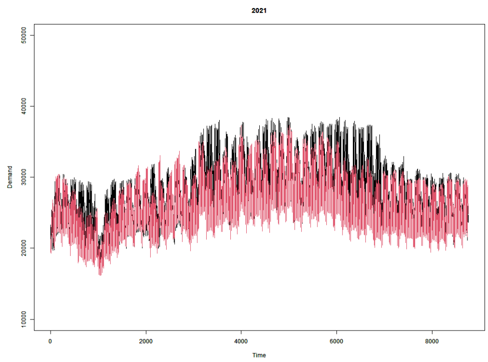
$$\hat{y}_t^{\dagger}(x_n^{(h)}) = \hat{y}_t(x_n^{(h)}) + \hat{d}_{2009+T+h}, \quad \text{and} \quad \tilde{y}_t^{\dagger}(x_n^{(h)}) = \tilde{y}_t(x_n^{(h)}) + \hat{d}_{2009+T+h}.$$

Table 3: Summaries of evaluation results under different metrics of the debiased predicted curves obtained by the Algorithm FTSA and regFTSA.

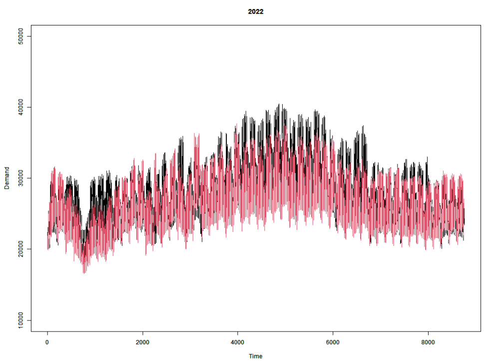
Evaluation Metric	Year / Algorithm	2020 ($h=1$)	2021 ($h=2$)	2022 ($h=2$)	2023 ($h=4$)
MAPE(h)(%)	FTSA	4.766	6.070	5.980	5.826
	regFTSA	4.721	5.988	5.909	5.801
MAE(h)	FTSA	1279.962	1736.838	1714.378	1629.975
	regFTSA	1266.314	1712.426	1692.857	1621.158
RMSE(h)	FTSA	1694.648	2275.849	2142.252	2137.999
	regFTSA	1680.757	2250.057	2119.178	2129.458



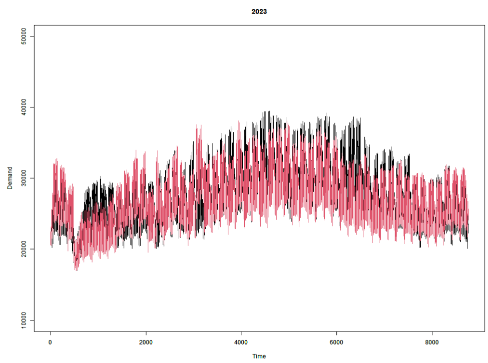
(a) 2020



(b) 2021

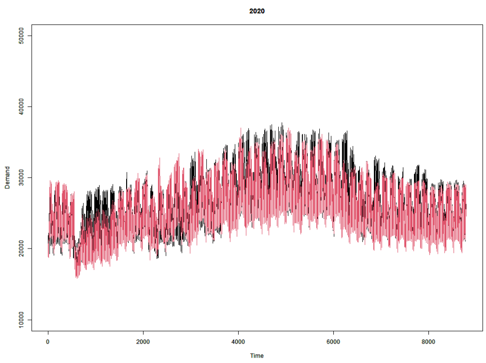


(c) 2022

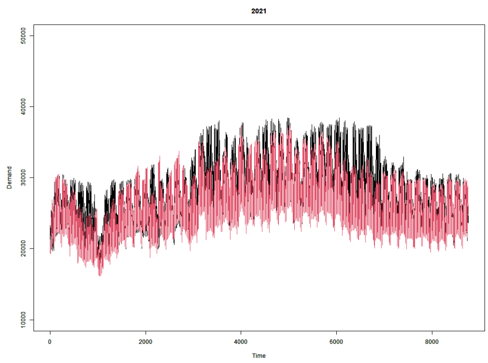


(d) 2023

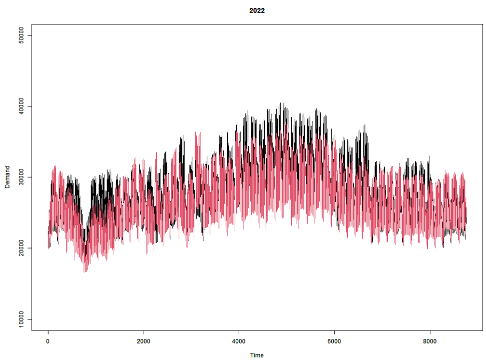
Figure 6: The debiased predicted yearly curves of the electricity demand from 2020 to 2023 obtain by Algorithm FTSA, where the black curve is the true electricity demand and the red curve is the predicted electricity demand.



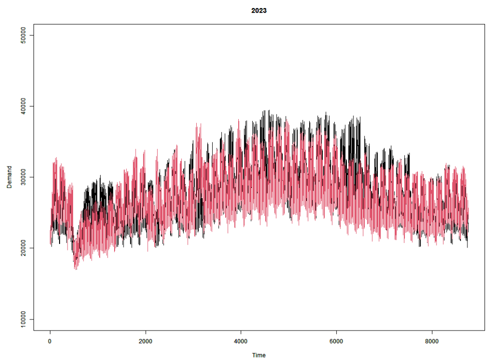
(a) 2020



(b) 2021



(c) 2022



(d) 2023

Figure 7: The debiased predicted yearly curves of the electricity demand from 2020 to 2023 obtain by Algorithm regFTSA, where the black curve is the true electricity demand and the red curve is the predicted electricity demand

As shown in Table 3, the debiased FTSA and regFTSA prediction curves, shown in Figures 6 and Figures 7, both have significant improvement across all evaluation metrics after the adjustment from the underestimation.

5. Discussion

This paper presents a multi-step forecasting method for future annual load curves using system load data. The data analysis is based on the FTSA method (Shang 2013), with adjustments made by incorporating exogenous variables into regression models to improve predictive accuracy. Due to the correlation between temperature and load data, temperature is treated as an exogenous variable, and regression models are applied to the FTSA framework to refine predictive performance.

FTSA allows for the analysis of curve data for any chosen time period, followed by multi-step forecasting. In this study, we found that segmenting the annual load curve into weekly datasets and then using FTSA for weekly multi-step forecasting, before integrating these into annual predictions, improves forecasting performance.

Overall, the approach of system load prediction based on the FTSA method supplemented by exogenous variable regression modeling for future annual load curve predictions performs well, as demonstrated by evaluation metrics including MAPE, MAE, and RMSE. However, it is essential to note that any system load curve prediction method involving FTSA relies on data quality and quantity. Therefore, augmenting data resources in the future would not only enable better capture of variations in load curves but also enhance predictive accuracy.

6. Acknowledgement

The financial support provided by Energy Administration is gratefully acknowledged.

References

[1] Bian, H., Wang, Q., Xu, G., and Zhao, X. (2022). Research on short-term load forecasting based on accumulated temperature effect and improved temporal convolutional network. *Energy Reports*, 8, pages 1482-1491.

[2] Chang, C. H., Chen, Z. B., and Huang, S. F. (2022). Forecasting of high-resolution electricity consumption with stochastic climatic covariates via a functional time series approach. *Applied Energy*, 309, 118418.

[3] Elías, A., Jiménes, R., and Shang, H. L. (2022). On projection methods for functional time series forecasting. *Journal of Multivariate Analysis*, 189, 104890.

[4] Filik, Ü. B., Gerek, Ö. N., and Kurban, M. (2011). A novel modeling approach for hourly forecasting of long-term electric energy demand. *Energy Conversion and Management*, 52(1), pages 199-211.

[5] Imani, M. (2021). Electrical load-temperature CNN for residential load forecast-

ing. *Energy*, 227, 120480.

[6] Lindberg, K. B., Seljom, P., Madsen, H., Fischer, D., and Korpås, M. (2019). Long-term electricity load forecasting: Current and future trends. *Utilities Policy*, 58, pages 102-119.

[7] Mamun, A. A., Sohel, M., Mohammad, N., Haque Sunny, M. S., Dipta, D. R., and Hossain, E. (2020). A comprehensive review of the load forecasting techniques using single and hybrid predictive models. *IEEE Access*, 8, pages 134911-134939.

[8] Moroff, N. U., Kurt, E., and Kamphues, J. (2021). Machine learning and statistics: A study for assessing innovative demand forecasting models. *Procedia Computer Science*, 180, pages 40-49.

[9] Shah, I., Iftikhar, H., and Ali, S. (2022). Modeling and forecasting electricity demand and prices: A comparison of alternative approaches. *Journal of Mathematics*, 2022(1), 3581037.

[10] Shang, H. L. (2013). ftsa: An R package for analyzing functional time series. *The R Journal*, 5, pages 64-72.

[11] Spencer, B., Alfandi, O., and Al-Obeidat, F. (2019). Forecasting temperature in a smart home with segmented linear regression. *Procedia Computer Science*, 155, pages 511-518.

[Received July 2024; accepted December 2024.]

台灣多年年負載曲線預測

許湘伶 [1†] **張志浩** [2] **朱基祥** [1] **張君鳳** [1] **尤姿文** [1] **陳竑廷** [3] **陳冠彰** [3]

[1] 國立高雄大學統計學研究所

[2] 國立政治大學統計學系

[3] 工業技術研究院綠能與環境研究所

摘　要

在電力系統中，確保供應安全和穩定至關重要。調度單位在進行長期電力發展規劃時需考量負載變化及成長等因素。根據台灣現行法規，備用容量率必須達到 15%。目前，國內的容量評估模型依賴歷史數據進行負載模擬，雖然此方法簡單且快捷，但忽略了不確定性和未來氣候變遷的影響。基於文獻中的現有方法，本研究提出一種基於統計框架的兩階段方法，用於預測未來年度負載曲線。此方法將有助於政府機構進行全面的長期電力供需評估，考慮不確定性並提高估算準確性。

關鍵詞：函數型時間序列分析、負載預測、迴歸模型。

JEL classification: L94, C32, E27.

[†]通訊作者：許湘伶
E-mail: hsuhl@nuk.edu.tw