Energy Reports 9 (2023) 1887–1895

Contents lists available at ScienceDirect

Energy Reports

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

Research paper

A novel deep learning framework with a COVID-19 adjustment for electricity demand forecasting

Zhesen Cui [a], Jinran Wu [b,*], Wei Lian [a], You-Gan Wang [b]

[a] Department of Computer Science, Changzhi University, Changzhi, Shanxi 046011, PR China
[b] Institute for Learning Sciences & Teacher Education, Australian Catholic University, Brisbane 4001, Australia

ARTICLE INFO

Article history:
Received 2 August 2022
Received in revised form 24 December 2022
Accepted 3 January 2023
Available online xxxx

Keywords:
Time series modeling
Pandemic
Deep learning
Load forecasting

ABSTRACT

Electricity demand forecasting is crucial for practical power system management. However, during the COVID-19 pandemic, the electricity demand system deviated from normal system, which has detrimental bias effect in future forecasts. To overcome this problem, we propose a deep learning framework with a COVID-19 adjustment for electricity demand forecasting. More specifically, we first designed COVID-19 related variables and applied a multiple linear regression model. After eliminating the impact of COVID-19, we employed an efficient deep learning algorithm, long short-term memory multiseasonal net deseasonalized approach, to model residuals from the linear model aforementioned. Finally, we demonstrated the merits of the proposed framework using the electricity demand in Taixing, Jiangsu, China, from May 13, 2018 to August 2, 2021.

© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

COVID-19, named by The World Health Organization (WHO) in February 11, 2020, is a novel threat to global health (Li et al., 2020). This virus can cause acute respiratory distress syndrome complications, which contribute to the high mortality rate (Peng et al., 2020; Meng et al., 2020; Mahmud et al., 2020). COVID-19 has caused great damage to the health systems of over 200 countries and WHO declared that this COVID-19 epidemic could be called a global pandemic in March 11, 2020 (Wu et al., 2021).

In addition to methods helping to diagnose COVID-19, many researchers carried out various analysis reports on different fields affected by the outbreak, such as the electricity industry (Zhong et al., 2020; Elavarasan et al., 2020; Wang et al., 2021; Wu et al., 2023), air pollution (Isaifan, 2020; Zoran et al., 2022), and stock prices (Wei et al., 2021; Huang and Liu, 2021). These reports assess and analyze the impact of COVID-19 on these fields after its outbreak. For the electricity industry, demand has dropped sharply with lockdown restrictions executed by governments around the world. Meanwhile, the demand composition and daily demand profile have also changed. Global air quality has been improved because of the reduction of emissions from closed factories, reduced vehicular road traffic, grounded flights, and other activities that consume fossil fuels. As for stock prices, it has been discovered that a cointegration relationship exists between the pandemic and solar stock prices (Wei et al., 2021). In addition, stock price crash risk of energy firms has significantly decreased in the post-COVID-19 period (Huang and Liu, 2021).

While COVID-19 significantly affects economy, society, and daily livelihoods (Zhong et al., 2020; Guan et al., 2020; Haxhimusa and Liebensteiner, 2021; Şahin et al., 2021), it also brings challenges to both demand and generation in cities' electricity systems. Under a series of policies proposed by the Chinese Government, like lockdown, people's living habits have changed. These changes also affect the electricity sectors in each city (Abdeen et al., 2021), as people have to stay at home, consuming more residential demand (Li et al., 2021). On the other hand, a multitude of factories (Baniasad et al., 2021), commercial centres, and entertainment places have to close, reducing industrial demand (Liu et al., 2022; Cihan, 2022). However, with the control policy and COVID-19 vaccine (Sharma et al., 2021; Sun et al., 2021), the epidemic in China has been gradually brought under control. Daily life and economic activities have gradually recovered, and electricity demand has recovered as well. Therefore, electricity demand has changed considerably compared with the pre-COVID-19 era, in many respects.

Although the impact of the COVID-19 epidemic is gradually coming under control, there is still validity in studying what kind of impact from the COVID-19 epidemic is still remaining on demand compared with the normal demand before the epidemic. In addition, the general demand trend in light of COVID-19 should be studied and a long-term analysis of its influence on the future demand should be conducted to provide a reference for the power sector.

* Corresponding author.
E-mail addresses: cuizhesen@gmail.com (Z. Cui), ryan.wu@acu.edu.au (J. Wu), lianwei3@gmail.com (W. Lian), you-gan.wang@acu.edu.au (Y.-G. Wang).

https://doi.org/10.1016/j.egyr.2023.01.019
2352-4847/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

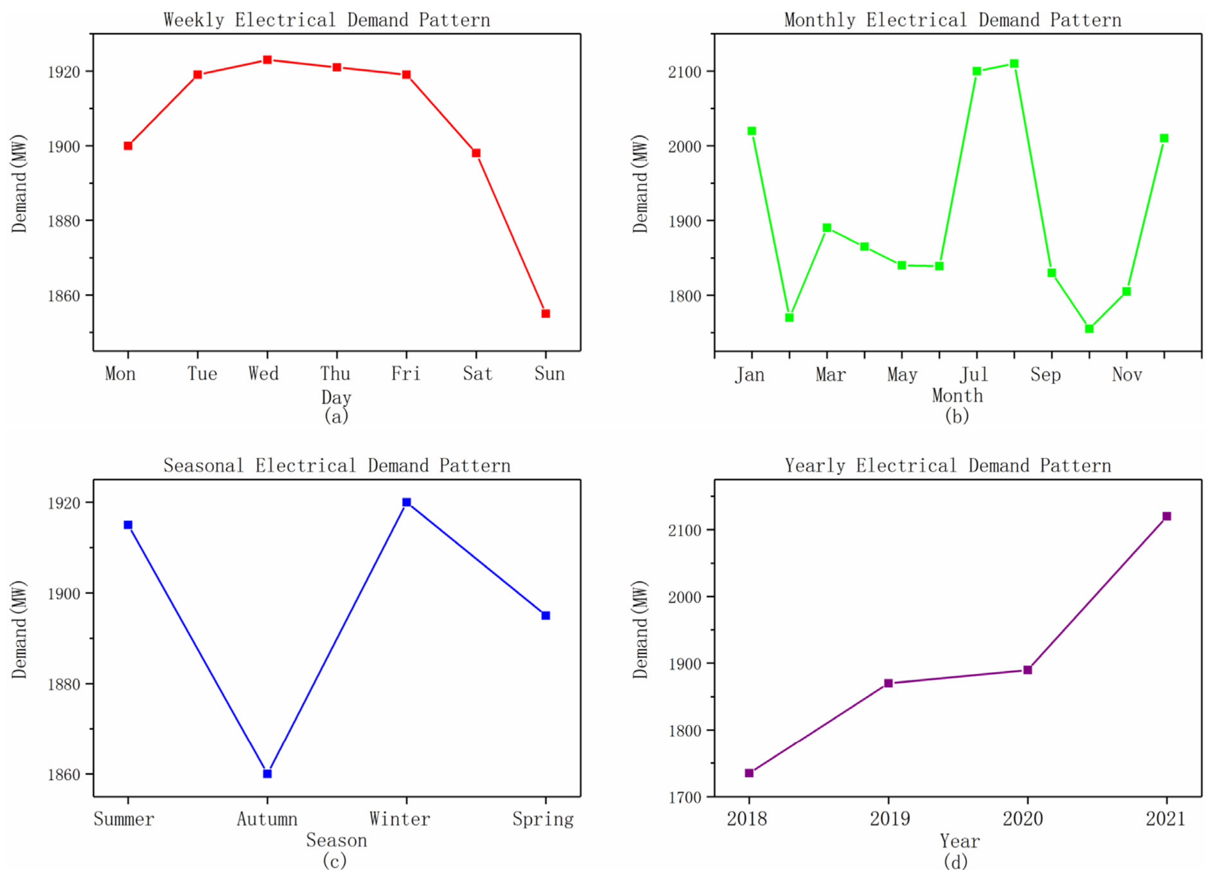

Fig. 1. Patterns for electricity demand. Figures (a), (b), (c), and (d) provide weekly patterns, monthly patterns, seasonal patterns, and annual patterns, respectively.

Due to the changes in government policies and people's living habits, the demand pattern has displayed significant levels of randomness, which may affect the safe operation of the power system and grid. Therefore, an accurate demand forecasting model needs to be able to accurately depict the influences of these changing patterns, especially the lockdown restrictions and the changes in people's daily habits. Since the changes brought on by the outbreak of COVID-19, we have realized that accurately depicting the impact of the epidemic is an important part of achieving accurate demand forecasting. Traditional demand forecasting methods have been largely unable to accurately describe the impact of the epidemic and generate satisfactory predictions, especially the demand during the COVID-19 period (Qin and Li, 2020; Obst et al., 2021). Therefore, we aimed to construct a framework that can describe the effect of COVID-19 and generate accurate demand predictions.

The contributions of this paper are listed as follows:

a. We analyzed and evaluated the demand data set from Taixing, Jiangsu, China and provide an insight of its weekly, monthly, seasonal, and annual patterns. According to these patterns, especially before and after the COVID-19 outbreak, comparisons, and evaluations were conducted to underline how the demand has changed in Taixing.
b. We expressly designed a forecasting framework composed of two stages for the demand data set. We firstly applied a number of variables and a multiple linear regression model to extract the main trend, the seasonality, and the impact of lockdown. In the second stage, we applied a long short-term memory multiseasonal net deseasonalized (LSTM-MSNet-DS) model to improve the prediction with the residuals from the first regression method.

The remaining contents of this paper are organized as follows. Section 2 introduces the basic information of the demand data set in Taixing and gives a preliminary analysis for this data set. Section 3 details the proposed demand forecasting framework. Section 4 displays the experiment results and empirical analysis. Section 5 concludes this paper.

2. Data and preliminary analysis

In this section, we look at our unique electricity demand data set, and introduce some basic knowledge drawn from it. These preliminary findings will guide our research and support our new model in later sections.

Daily forecasting of electrical demand plays an important role in transmission system operators because its accuracy affects system operations and management. We collected the daily demand data set for Taixing in China, from May 13, 2018, to August 2, 2021. Here, we carried out our preliminary analysis for this data set. We found divergence among the patterns in weekdays or weekends because of the difference in locations where people gather. In addition, monthly patterns and seasonal patterns varied according to time or seasonal differences. Meanwhile, the demand changed along with the overall economic situation every year. Examples of these factors are displayed in Fig. 1.

Overall, we obtained the following findings. As shown in Fig. 1 (a), electrical demand was lower on weekends. From Fig. 1(c), the electricity demand in summer and winter was larger. The distribution of electricity demand in Fig. 1(b) supports the patterns in Fig. 1(c). The demand in summer and winter months was higher. From Fig. 1(d), we find that the overall trend of electricity demand in Taixing was gradually increasing despite the outbreak of COVID-19 in early 2020. Given the suddenness of COVID-19

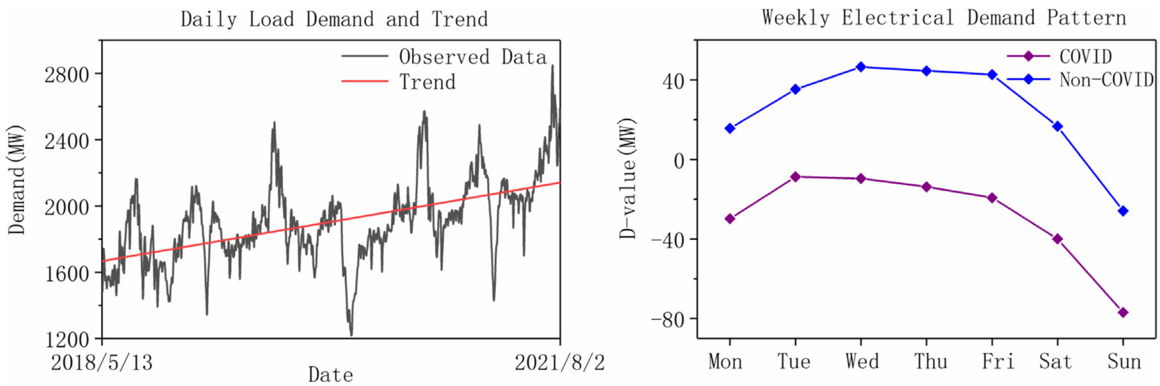

Fig. 2. Patterns for electricity demand. The left figure provides the daily demand and the trend. The right figure provides the comparisons between the COVID period and the non-COVID period where D-value = Demand − 0.4019TimeOrder − 1667.

and the rapidity of transmission, China has adopted working from home and other measures to curb the spread of the pandemic by controlling excessive population movement (Abdeen et al., 2021). With the beginning of the COVID-19 pandemic in Taixing since January 29, 2020, in order to properly describe its impact, we constructed a dummy variable named "COVID" from the data.

To fully illustrate the demand difference between the COVID period and non-COVID period, as shown in the left sub-figure of Fig. 2, we first constructed a linear regression model with the time order TimeOrder to obtain the trend of the demand from May 13, 2018 to August 2, 2021 as 0.4019TimeOrder + 1667. Then, to eliminate the impact of the gradual increase in electricity demand for recent years, we calculated the difference D-value between the actual demand and the trend, i.e., D-value = Demand − 0.4019TimeOrder − 1667, to provide comparisons between the COVID period and the non-COVID period. It is shown that the D-value in the non-COVID period is significantly larger than in the COVID period in Taixing.

3. Methodology

In this section, we construct a certain number of variables as well as a multiple linear regression model. The model mainly aims to capture the impact from COVID along with the outbreak of COVID-19 and illustrate relevant effects on demand. Furthermore, a novel deep learning method (LSTM-MSNet-DS) is employed to model the residuals from the regression model. Finally, our predictions of demand are combined with the predictions from the linear regression model and the improvements from LSTM-MSNet-DS.

3.1. The multiple linear regression model

As discussed in Section 2, we constructed corresponding variables for the patterns in electricity demand in Taixing, i.e., weekly, monthly, seasonal, and yearly patterns. Meanwhile, we built some new variables to mark the following unique features during COVID-19.

a. **Lockdown:** For the lockdown period from January 29, 2020 to February 26, 2020 we constructed a "Lockdown" variable to mark the electricity demand pattern in this period;
b. **COVID:** A "COVID" variable was constructed to depict the general influence on electricity demand of COVID-19 before and after the outbreak;
c. **Weekend:** Since COVID-19 has caused great changes to the operation mode of urban electric power (Li et al., 2021; Baniasad et al., 2021), we constructed a "Weekend" variable to capture the pattern changes of working days and weekends.

Inspired by Zhao et al. (2022) and Wu et al. (2023), we also applied corresponding harmonic functions as our variables. Specifically, we have $\sin(2\pi t/T)$ and $\cos(2\pi t/T)$, where t is time order in time series and T represents the cycle time. In addition, we applied eight variables, involving the "COVID" variable and eight relevant harmonic functions ($\sin(2\pi t/7)$, $\cos(2\pi t/7)$, $\sin(2\pi t/(365.25/12))$, $\cos(2\pi t/(365.25/12))$, $\sin(2\pi t/(365.25/4))$, $\cos(2\pi t/(365.25/4))$, $\sin(2\pi t/365.25)$, and $\cos(2\pi t/365.25)$). To depict the general trends of electricity demand in Taixing, we also included the time order series t as a variable. The detailed information is shown in Table 1.

Here, we constructed a multiple linear regression model with those variables, according to the general formula,

$$\begin{aligned}\mathrm{Model_{MLR}} =& \omega_1\mathrm{Weekend} + \omega_2\mathrm{Lockdown} + \omega_3\mathrm{Weekend\_COVID}\\ &+ \omega_4\mathrm{WeeklySin} + \omega_5\mathrm{WeeklyCos}\\ &+ \omega_6\mathrm{MonthlySin} + \omega_7\mathrm{MonthlyCos} + \omega_8\mathrm{SeasonlySin}\\ &+ \omega_9\mathrm{SeasonlyCos} + \omega_{10}\mathrm{YearlySin}\\ &+ \omega_{11}\mathrm{YearlyCos} + \omega_{12}\mathrm{WeeklySin\_COVID}\\ &+ \omega_{13}\mathrm{WeeklyCos\_COVID}\\ &+ \omega_{14}\mathrm{MonthlySin\_COVID} + \omega_{15}\mathrm{MonthlyCos\_COVID}\\ &+ \omega_{16}\mathrm{SeasonlySin\_COVID}\\ &+ \omega_{17}\mathrm{SeasonlyCos\_COVID} + \omega_{18}\mathrm{YearlySin\_COVID}\\ &+ \omega_{19}\mathrm{YearlyCos\_COVID} + \omega_{20}\mathrm{COVID} + \omega_{21}\mathrm{Time} + \mathrm{b},\end{aligned} \tag{1}$$

where $\omega = [\omega_1, \omega_2, \ldots, \omega_{21}]^T$ is the coefficient vector for the variables, and b is an intercept item. We denote the $\mathrm{Model_{MLR}}$ as $\hat{\mathrm{F}}$ in the following contents.

3.2. The LSTM-MSNet-DS model

Since the variables mainly referred to the effect of COVID-19, the predictions obtained by our multiple linear regression model only represents partial patterns. Thus, further adjustment was necessary to improve the overall accuracy of the model. Accordingly, we built an adjustment scheme for the demand series in terms of effects from the COVID-19 pandemic. The adjustment was realized by obtaining residuals between the original series and the output of the multiple linear regression model, which is the key part of our approach. With the adjusted series, we applied a novel time series forecasting method, named LSTM-MSNet-DS in Bandara et al. (2020) to generate accurate predictions.

LSTM-MSNeT-DS is a forecasting framework to predict time series with multiple seasonality based on decomposition techniques. MSNeT refers to multiseasonal net, which uses a deseasonalization strategy to detach the multiseasonal components

Table 1
21 variables constructed for multiple linear regression model.

Variable	Definition	Variable	Definition
Weekend	$\begin{cases} 1 & \text{if weekend} \\ 0 & \text{otherwise} \end{cases}$	Time	$t = 1, 2, \ldots$
Lockdown	$\begin{cases} 1 & \text{if lockdown} \\ 0 & \text{otherwise} \end{cases}$	COVID	$\begin{cases} 1 & \text{if COVID} \\ 0 & \text{otherwise} \end{cases}$
WeeklySin	$\sin(2\pi t/T), T = 7$	WeeklyCos	$\cos(2\pi t/T), T = 7$
MonthlySin	$\sin(2\pi t/T), T = 365.25/12$	MonthlyCos	$\cos(2\pi t/T), T = 365.25/12$
SeasonlySin	$\sin(2\pi t/T), T = 365.25/4$	SeasonlyCos	$\cos(2\pi t/T), T = 365.25/4$
YearlySin	$\sin(2\pi t/T), T = 365.25$	YearlyCos	$\cos(2\pi t/T), T = 365.25$
WeeklySin_COVID	WeeklySin×COVID	WeeklyCos_COVID	WeeklyCos×COVID
MonthlySin_COVID	MonthlySin×COVID	MonthlyCos_COVID	MonthlyCos×COVID
SeasonlySin_COVID	SeasonlySin×COVID	SeasonlyCos_COVID	SeasonlyCos×COVID
YearlySin_COVID	YearlySin×COVID	YearlyCos_COVID	YearlyCos×COVID
Weekend_COVID	Weekend×COVID		

from a time series. The model above has three layers, namely: (1) a pre-processing layer, consisting of the normalization, the variance stabilization, and the seasonal decomposition phase; (2) a recursion layer, consisting of an LSTM-based stacking structure for training the network; (3) a post-processing layer, which denormalizes and re-seasonalizes the time series to derive the final prediction results.

We firstly generated an adjusted residual series r between original demand series and the series obtained by Eq. (1). This residual series removes the patterns of our variables and reduces the complexity of the original series. The residual series is defined as:

$$\mathrm{r} = \mathrm{y} - \mathrm{F},$$

where F and y represent the output of multiple linear regression model and real demand observations, respectively. The residual series was composed of two parts: regression modeling residual $\mathrm{r_{mod}}$ and regression prediction residual $\mathrm{r_{pre}}$. They were obtained in our training set and test set, respectively.

For series r, seasonal decomposition method was applied to make a proper seasonal adjustment. Since the seasonal variables in Table 1 only capture the linear seasonal pattern, we applied a de-seasonalization of the seasonal decomposition method, which forms part of complexity of the demand series and improves the learning process. The corresponding formula can be defined as follows:

$$\mathrm{r} = \hat{\mathrm{S}}_\mathrm{t} + \hat{\mathrm{T}}_\mathrm{t} + \hat{\mathrm{R}}_\mathrm{t},$$

where $\hat{\mathrm{S}}_\mathrm{t}$, $\hat{\mathrm{T}}_\mathrm{t}$ and $\hat{\mathrm{R}}_\mathrm{t}$ are the seasonal, trend, and remaining components, respectively. Then we combine $\hat{\mathrm{T}}_\mathrm{t}$ and $\hat{\mathrm{R}}_\mathrm{t}$ into a new time series r_DES.

LSTM unit (Bandara et al., 2020) is used in this model for its advantages in handling the time series. In LSTM, its gating mechanism and memory cell in each unit enable a neural network to capture non-linear long-term temporal correlation in a time series. The updated formulas of LSTM are shown as follows:

$$\mathrm{i_t} = \sigma(\mathrm{W_i}[\mathrm{h_{t-1}}, \mathrm{r\_DES_t}] + \mathrm{b_i}),$$

$$\mathrm{f_t} = \sigma(\mathrm{W_f}[\mathrm{h_{t-1}}, \mathrm{r\_DES_t}] + \mathrm{b_f}),$$

$$\tilde{\mathrm{C}}_\mathrm{t} = \tanh(\mathrm{W_C}[\mathrm{h_{t-1}}, \mathrm{r\_DES_t}] + \mathrm{b_C}),$$

$$\mathrm{o_t} = \sigma(\mathrm{W_o}[\mathrm{h_{t-1}}, \mathrm{r\_DES_t}] + \mathrm{b_o}),$$

$$\mathrm{C_t} = \mathrm{f_t} * \mathrm{C_{t-1}} + \mathrm{i_t} * \tilde{\mathrm{C}}_\mathrm{t},$$

and

$$\mathrm{h_t} = \mathrm{o_t} * \tanh(\mathrm{C_t}),$$

where $\mathrm{W_i}$, $\mathrm{W_f}$, $\mathrm{W_c}$ and $\mathrm{W_o}$ represent the weight matrices of the input gate, forget gate, memory cell gates, and output gate, respectively. The biases of the gates are represented by $\mathrm{b_i}$, $\mathrm{b_f}$, $\mathrm{b_c}$ and $\mathrm{b_o}$. $\tilde{\mathrm{C}}_\mathrm{t}$ is to update the state of the original memory cell $\mathrm{C_t}$. $\mathrm{h_{t-1}}$ is the hidden state. We have $*$ represent the element-wise multiplication operation, σ representing the sigmoid function, and tanh representing the hyperbolic tangent function.

With the LSTM model, we can train and obtain a prediction sequence $\hat{\mathrm{r}}_\mathrm{pre}$ corresponding to regression prediction residual $\mathrm{r_{pre}}$. Then the final predictions can be obtained by

$$\hat{\mathrm{y}} = \hat{\mathrm{r}}_\mathrm{pre} + \mathrm{F}, \tag{2}$$

where $\hat{\mathrm{y}}$ refers to final demand predictions.

4. The case study

The proposed demand forecasting framework is composed of two parts, i.e., our multiple linear regression model for a certain number of variables and the LSTM-MSNet-DS model. In this section, corresponding analyses are carried out for the two parts mentioned above, and final demand predictions.

4.1. The experimental configuration

For the collected demand data set in Taixing, China, we divided it into two exclusive parts: the training set and test set. The training set consisted of data from May 13, 2018 to June 30, 2021 and the rest formed the test data set i.e., from July 1, 2021 to August 2, 2021 and the segmentation of the residuals obtained by multi-variable linear regression model is the same as the demand data set.

We applied seven popular algorithms to form a comparison pool, i.e., support vector regression (SVR) (Yang et al., 2019; Wu and Wang, 2022), long short-term memory (LSTM) (Selvin et al., 2017), bidirectional long short-term memory (Bi-LSTM) (Shahid et al., 2020), gate recurrent unit (GRU) (Fu et al., 2016), bidirectional gate recurrent unit (Bi-GRU) (Chung et al., 2014), spatio-temporal attention long short-term memory (STA-LSTM) (Yuan et al., 2020) and the introduced LSTM-MSNet-DS. Our deep learning framework with a COVID-19 adjustment is implemented through two steps: (1) forecasting the original demand data set; and (2) forecasting the residuals obtained by the multi-variable linear-regression model.

The parameters for the considered models were set as follows: (1) SVR is a branch of the support vector machine (SVM). We chose "rbf" function as the kernel function. The penalty parameter C and epsilon were selected as 10 and 0.1. (2) For LSTM, Bi-LSTM, GRU, and Bi-GRU, the number of hidden layer nodes was selected as 200 after trial and error. (3) The STA-LSTM contained

Table 2
The coefficients for the 21 variables in multiple linear regression model.

Variable	Coefficients	Variable	Coefficients
Weekend	−23.50	Time	0.52***
Lockdown	−188.12***	Weekend_COVID	36.71
WeeklySin	−24.33...	WeeklyCos	−5.93
MonthlySin	20.13*	MonthlyCos	2.28
SeasonlySin	−19.00...	SeasonlyCos	−3.08
YearlySin	41.01***	YearlyCos	30.82**
WeeklySin_COVID	5.43	WeeklyCos_COVID	−6.79
MonthlySin_COVID	−28.14...	MonthlyCos_COVID	−7.55
SeasonlySin_COVID	31.16*	SeasonlyCos_COVID	−18.74
YearlySin_COVID	47.32**	YearlyCos_COVID	−60.60***
COVID	−115.64***	Intercept	1656.54***

Significant codes: *** : <0.001; ** : <0.01; * : <0.05; ... : <0.1.

spatial-LSTM and temporal-LSTM, and the number of hidden layer nodes was also selected as 200. (4) The number of hidden layer nodes in LSTM-MSNet-DS model was 200. The loss function was selected as Huber loss, named SmoothL1Loss in Pytorch, to obtain more robust predictions.

In the experiments, each model in the approach pool was applied to generate one-day ahead prediction. To eliminate accidental factors, each experiment was repeated 30 times. All the experiments were conducted by Pycharm integrating Python 3.7 and Pytorch 1.9.0.

4.2. Evaluation criterion

To evaluate the forecasting performance of all benchmark models, mean absolute error (MAE), root mean square error (RMSE), mean relative error (MRE), and R-squared (R^2) were applied to evaluate the predictions,

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|,$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2},$$

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|,$$

and

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y}_i - y_i)^2},$$

where n is for the sample size, $\hat{y}_i$ for the prediction, $\bar{y}_i$ for the average value, and y_i for the observation.

4.3. The experimental result

In this subsection, we will analyze the experimental results from three aspects: (1) How do the constructed variables influence the demand? (2) Is it using the COVID-19 adjustment or not? and (3) How does the proposed approach perform compared with other approaches? The predictions are plotted in Fig. 3. The results of relative error (RE) between LSTM-MSNet-DS and the comparison models are presented in Fig. 4, which indicates that the proposed framework can obtain minimum RE values for most predicted values.

4.3.1. The COVID impact modeling

Our first step was to construct a regression model with twenty-one variables in Table 1. The coefficient estimates and the corresponding significant test for our twenty-one variables are shown in Table 2. The p-values of Time, Lockdown, MonthlySin, YearlySin, YearlyCos, SeasonlySin_COVID, YearlySin_COVID, YearlyCos_COVID, COVID, and Intercept are less than 0.05, which indicates that these coefficient estimates are significant, and this means these constructed variables are valid. In addition, a multicollinearity test was conducted for our constructed variables, and we find there is no significantly multicollinearity that is detailed in Appendix. Therefore, we retrained a multiple linear regression model with these nine variables.

Table 3
The coefficients for the 9 variables in multiple linear regression model.

Variable	Coefficients	Variable	Coefficients
Time	0.53***	Lockdown	−409.45***
MonthlySin	7.72	YearlySin	40.38***
YearlyCos	30.17**	SeasonlySin_COVID	12.26
YearlySin_COVID	47.02**	YearlyCos_COVID	−59.97***
COVID	−114.55***	Intercept	1649.08***

Significant codes: *** : <0.001; ** : <0.01; * : <0.05.

The coefficient estimates and the corresponding significant test for 9 variables are shown in Table 3 where only two coefficient estimates are insignificant. Our target is to construct a forecasting model considering COVID impacts. Given the small number of predictors, it is unclear whether the removal of the non-significant variables would give rise to higher forecasting accuracy due to potential type I and II errors occurred. Moreover, the explanatory variables are developed meaningfully. Therefore, two insignificant coefficients are not removed. In addition, the R-squared value and *p*-value of this model are 0.431 and 2.21e−132, respectively, which indicates that the regression model with these 9 variables is significant. Therefore, we presented the estimated linear regression model as,

$$\begin{aligned}\hat{F} =& 0.53\mathrm{time} - 409.45\mathrm{lockdown} + 7.72\mathrm{MonthlySin}\\ &+ 40.38\mathrm{YearlySin} + 30.17\mathrm{YearlyCos}\\ &+ 12.26\mathrm{SeasonlySin\_COVID} + 47.02\mathrm{YearlySin\_COVID}\\ &- 59.97\mathrm{YearlyCos\_COVID}\\ &- 114.55\mathrm{COVID} + 1649.08. \end{aligned} \quad (3)$$

Here, we found the coefficient of the lockdown variable is −409.45. This demonstrates that the lockdown policy caused a significant decline in the demand in Taixing. Meanwhile, the coefficient of the COVID variable is −114.55. This indicates that COVID-19 had a generally negative impact on demand compared with the period prior to this pandemic. According to the time variable, which is a positive number 0.53, we can conclude that the general trend of demand in Taixing increased despite the negative effect of strict government policies and the COVID-19 pandemic.

According to our regression model, the residual sequence is adjusted by subtracting the output of regression model from the

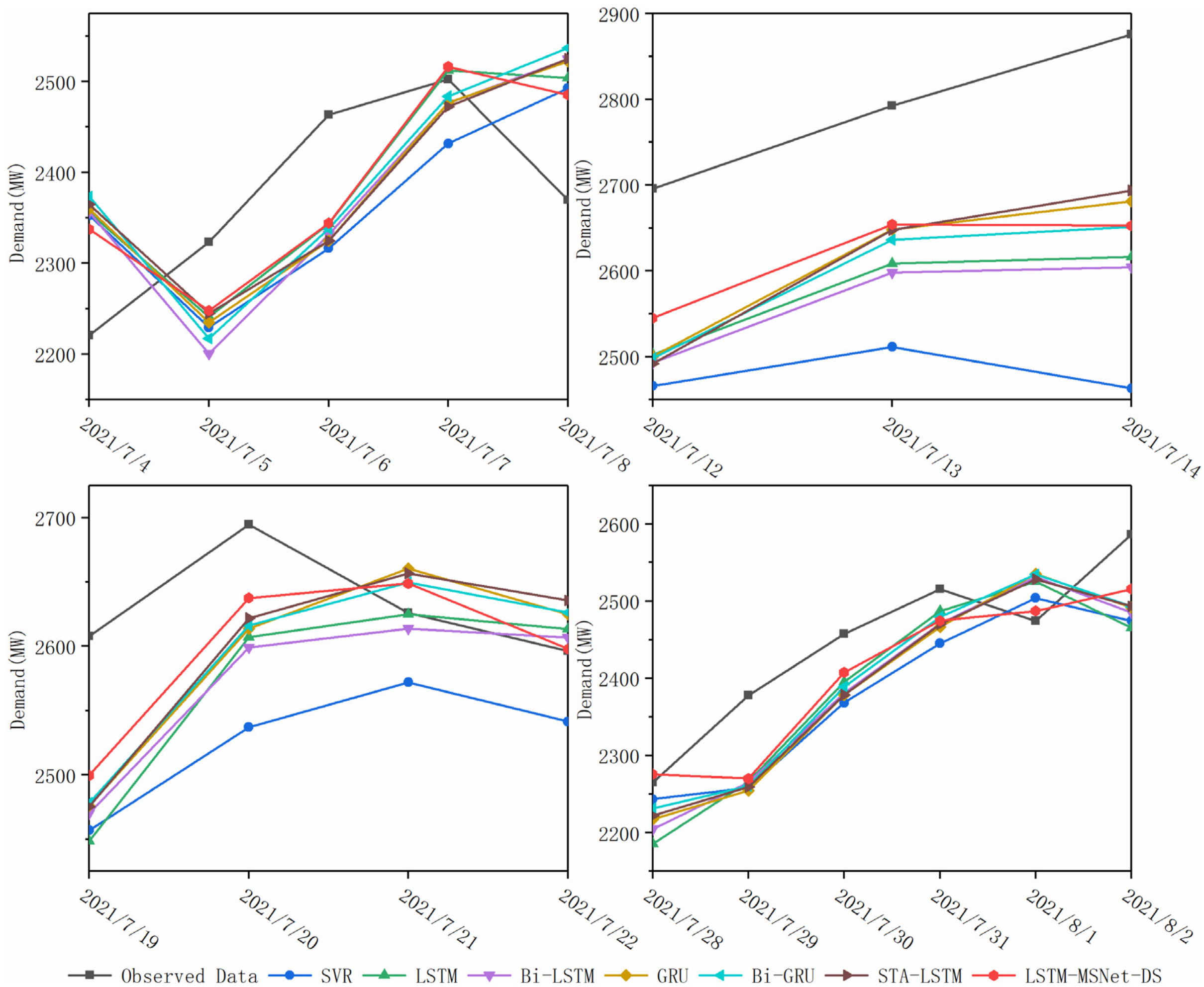

Fig. 3. The predictions for the seven models with adjustment.

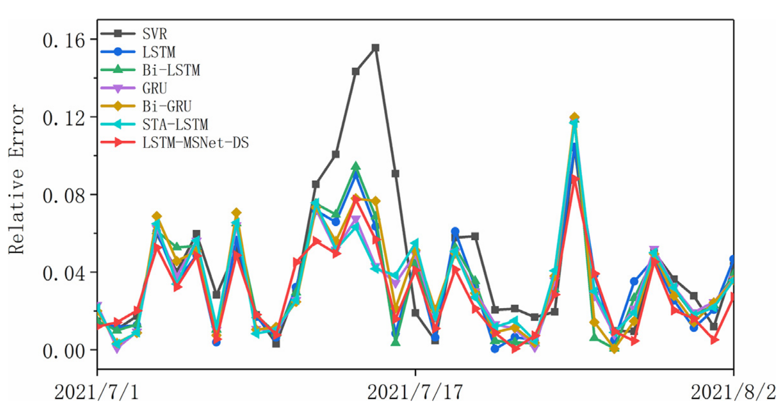

Fig. 4. The relative error comparison for different models with adjustment.

original demand sequence. With this process, the patterns of COVID-19 and lockdown restrictions can be removed. Additionally, the linear monthly and yearly patterns can also be removed from the demand series. Consequently, the computing complexity of demand is reduced, with improved accuracy of predictions.

4.3.2. *The impact from COVID-19 adjustment*

Here we compare the results from seven forecasting models with or without the COVID-19 adjustment. Comparing the results of the second to fifth columns and sixth to ninth columns in Table 4, all models in the proposed framework can consistently

Table 4
The forecasting performance of all benchmark models.

Model	Forecasting original demand				With adjustment			
	MAE	RMSE	MRE	R^2	MAE	RMSE	MRE	R^2
SVR	139.12	160.70	0.0052	0.099	112.73	154.53	0.0439	0.167
LSTM	134.08	167.95	0.0530	0.014	85.93	110.19	0.0342	0.576
Bi-LSTM	109.89	134.02	0.0439	0.318	88.20	115.12	0.0351	0.537
GRU	124.60	150.48	0.0498	0.137	86.34	107.19	0.0346	0.599
Bi-GRU	119.39	143.80	0.0476	0.219	87.96	113.10	0.0351	0.553
STA-LSTM	137.28	162.87	0.0546	0.004	86.57	107.32	0.0347	0.598
LSTM-MSNet-DS	91.34	112.21	0.0364	0.548	74.99	94.01	0.0299	0.691

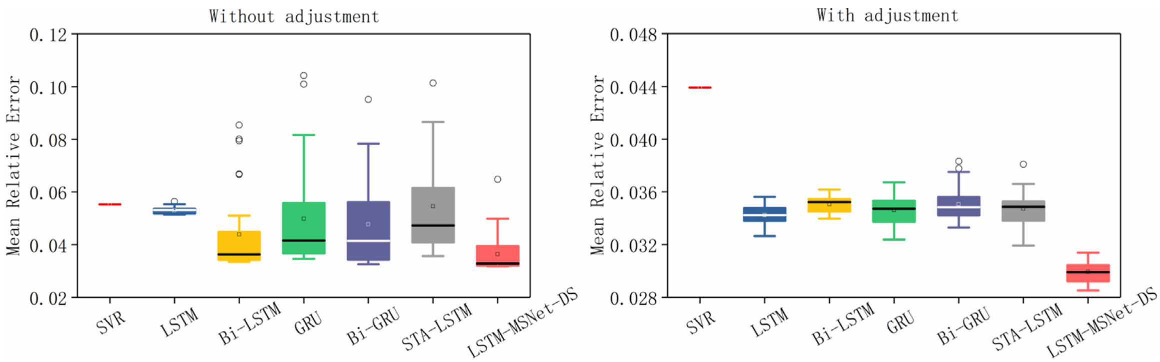

Fig. 5. The boxplot figure of experimental results in the test set.

perform better than when forecasting the original demand series. Specifically, the models using our framework perform over 17% more effectively compared with forecasting the original demand series in terms of MAE. Moreover, in terms of RMSE and MRE, the differences were over 4% and 17.1%, respectively. Furthermore, the proposed model had the best performance in R-squared.

For MRE criterion, we drew a boxplot figure for 30 experiments of each model in Fig. 5. We can see that the proposed demand forecasting framework can improve the forecasting performance for any general model with the emergence of COVID-19. The main reason for the better performance of our framework may be that removing the patterns of the constructed nine variables from the original series can improve the efficiency of demand modeling.

4.3.3. The forecasting comparison

As for the seven models in the approach pool, we compared the forecasting performance of the LSTM-MSNet-DS model with other approaches. According to the two sub-figures in Fig. 5, we can see that our LSTM-MSNet-DS model achieves the best performance among these models.

Since combining the forecasting model into the proposed framework can achieve a better performance, we can focus on the results of the models with the COVID-19 adjustment. Among the seven models, LSTM, Bi-LSTM, GRU, Bi-GRU, and STA-LSTM showed similar performances in terms of MAE, RMSE, and MRE criterion, but our LSTM-MSNet-DS approach could achieve a better prediction performance. Our LSTM-MSNet-DS model achieved the best performance, as it was over 12.7% superior to the other models.

As shown in Fig. 6, we plotted the correlation coefficient graphs for the real observations and predictions of the test set. According to the Pearson coefficient in this figure, we can also conclude that the LSTM-MSNet-DS can achieve the best forecasting performance among these models.

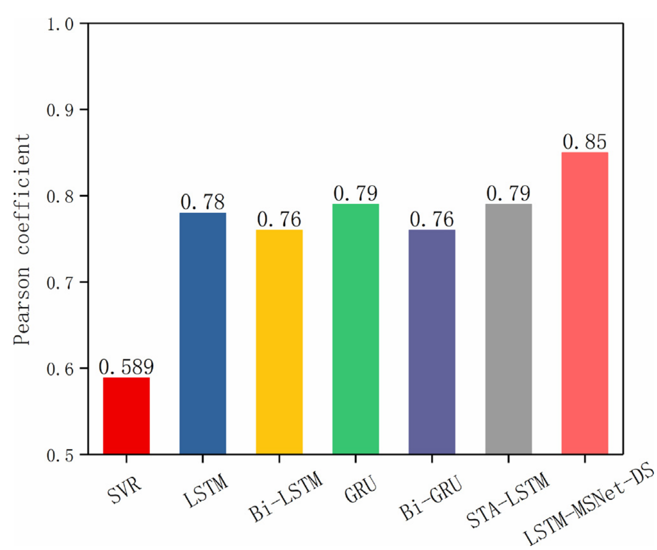

Fig. 6. The correlation coefficient graphs of all benchmark models.

5. Conclusion

The COVID-19 pandemic caused a sharp drop in power demand, and the demand composition and daily demand situation also changed. The change in the power balance situation and the increase of demand uncertainty have brought greater pressure to system operators, at the same time, they have also brought about challenges in system maintenance and management. Taking the example of Taixing, China, we focused on the effects of demand patterns during the lockdown period since the COVID-19 outbreak and non-lockdown period. We constructed a multiple linear regression model to depict the influence of the lockdown and the pandemic. For current demand, which is still influenced by COVID-19, we combined the multiple linear regression model and LSTM-MSNet-DS model to construct a novel forecasting framework. The experimental results demonstrate the proposed framework can generate accurate demand predictions. Furthermore, the electricity sectors can receive a deep insight into the influence the COVID-19 and the lockdown has. Meanwhile, accurate daily demand predictions generated by our framework are of great significance for reasonably arranging power grid

Table A.5
The VIF values for the 21 variables.

Variables	VIF	Variables	VIF	Variables	VIF
Weekend	5.37	Time	4.57	Lockdown	1.31
Weekend_COVID	6.51	WeeklySin	2.54	WeeklyCos	4.72
MonthlySin	1.89	MonthlyCos	1.88	SeasonlySin	1.88
SeasonlyCos	1.90	YearlySin	2.03	YearlyCos	1.99
WeeklySin_COVID	2.54	WeeklyCos_COVID	4.71	MonthlySin_COVID	1.89
MonthlyCos_COVID	1.88	SeasonlySin_COVID	1.88	SeasonlyCos_COVID	1.98
YearlySin_COVID	2.09	YearlyCos_COVID	1.94	COVID	5.83

operation modes and maintenance plans, saving coal and oil, reducing power generation costs, formulating reasonable power supply construction plans, and improving the economic and social benefits of the power system.

CRediT authorship contribution statement

Zhesen Cui: Software, Visualization, Formal analysis, Writing – original draft, Investigation, Project administration. **Jinran Wu:** Formal analysis, Writing – original draft, Writing – review & editing. **Wei Lian:** Writing – review & editing. **You-Gan Wang:** Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi, PR China (2022L517).

Appendix. The result of the multicollinearity test

In this section, we employed the variance inflation factor (VIF) to measure the inflation in the coefficient of the independent variable due to the collinearities among the other independent variables (Mansfield and Helms, 1982). According to the measurement, a VIF value larger than 10 indicates there is multicollinearity among the independent variables. The technical details of VIF can be found in Miles (2014). To conduct the statistical test, the function "ols_vif_tol()" in R package "olsrr" is used, and the corresponding VIF values for each variable are obtained. As shown in Table A.5, all VIF values are smaller than 10, thus we can conclude there is no multicollinearity among the 21 constructed variables.

References

Abdeen, Ahmed, Kharvari, Farzam, O'Brien, William, Gunay, Burak, 2021. The impact of the covid-19 on households' hourly electricity consumption in canada. Energy Build. 250, 111280.
Bandara, Kasun, Bergmeir, Christoph, Hewamalage, Hansika, 2020. Lstm-msnet: Leveraging forecasts on sets of related time series with multiple seasonal patterns. IEEE Trans. Neural Netw. Learn. Syst. 32 (4), 1586–1599.
Baniasad, Maryam, Mofrad, Morvarid Golrokh, Bahmanabadi, Bahare, Jamshidi, Sajad, 2021. Covid-19 in asia: Transmission factors, re-opening policies, and vaccination simulation. Environ. Res. 202, 111657.
Chung, Junyoung, Gulcehre, Caglar, Cho, Kyunghyun, Bengio, Yoshua, 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. In: NIPS 2014 Workshop on Deep Learning, December 2014.
Cihan, Pınar, 2022. Impact of the covid-19 lockdowns on electricity and natural gas consumption in the different industrial zones and forecasting consumption amounts: Turkey case study. Int. J. Electr. Power Energy Syst. 134, 107369.
Elavarasan, Rajvikram Madurai, Shafiullah, G.M., Raju, Kannadasan, Mudgal, Vijay, Arif, Mohammad Taufiqul, Jamal, Taskin, Subramanian, Senthilkumar, Sriraja Balaguru, V.S., Reddy, K.S., Subramaniam, Umashankar, 2020. Covid-19: Impact analysis and recommendations for power sector operation. Appl. Energy 279, 115739.
Fu, Rui, Zhang, Zuo, Li, Li, 2016. Using lstm and gru neural network methods for traffic flow prediction. In: 2016 31st Youth Academic Annual Conference of Chinese Association of Automation. YAC, IEEE, pp. 324–328.
Guan, Dabo, Wang, Daoping, Hallegatte, Stephane, Davis, Steven J., Huo, Jingwen, Li, Shuping, Bai, Yangchun, Lei, Tianyang, Xue, Qianyu, Coffman, D'Maris, et al., 2020. Global supply-chain effects of covid-19 control measures. Nat. Hum. Behav. 4 (6), 577–587.
Haxhimusa, Adhurim, Liebensteiner, Mario, 2021. Effects of electricity demand reductions under a carbon pricing regime on emissions: lessons from covid-19. Energy Policy 156, 112392.
Huang, Shoujun, Liu, Hezhe, 2021. Impact of covid-19 on stock price crash risk: Evidence from chinese energy firms. Energy Econ. 101, 105431.
Isaifan, R.J., 2020. The dramatic impact of coronavirus outbreak on air quality: has it saved as much as it has killed so far? Glob. J. Environ. Sci. Manag. 6 (3), 275–288.
Li, Cong, Dong, Di, Li, Liang, Gong, Wei, Li, Xiaohu, Bai, Yan, Wang, Meiyun, Hu, Zhenhua, Zha, Yunfei, Tian, Jie, 2020. Classification of severe and critical covid-19 using deep learning and radiomics. IEEE J. Biomed. Health Inf. 24 (12), 3585–3594.
Li, Lechen, Meinrenken, Christoph J., Modi, Vijay, Culligan, Patricia J., 2021. Impacts of covid-19 related stay-at-home restrictions on residential electricity use and implications for future grid stability. Energy Build. 251, 111330.
Liu, Dong, Yang, Hong, Thompson, Julian R., Li, Junli, Loiselle, Steven, Duan, Hongtao, 2022. Covid-19 lockdown improved river water quality in china. Sci. Total Environ. 802, 149585.
Mahmud, Tanvir, Alam, Md Jahin, Chowdhury, Sakib, Ali, Sham Nafisa, Rahman, Md Maisoon, Fattah, Shaik Anowarul, Saquib, Mohammad, 2020. Covtanet: A hybrid tri-level attention-based network for lesion segmentation, diagnosis, and severity prediction of covid-19 chest ct scans. IEEE Trans. Ind. Inform. 17 (9), 6489–6498.
Mansfield, Edward R., Helms, Billy P., 1982. Detecting multicollinearity. Amer. Statist. 36 (3a), 158–160.
Meng, Lingwei, Dong, Di, Li, Liang, Niu, Meng, Bai, Yan, Wang, Meiyun, Qiu, Xiaoming, Zha, Yunfei, Tian, Jie, 2020. A deep learning prognosis model help alert for covid-19 patients at high-risk of death: a multi-center study. IEEE J. Biomed. Health Inf. 24 (12), 3576–3584.
Miles, Jeremy, 2014. Tolerance and variance inflation factor. In: Wiley StatsRef: Statistics Reference Online.
Obst, David, Vilmarest, Joseph De, Goude, Yannig, 2021. Adaptive methods for short-term electricity load forecasting during covid-19 lockdown in france. IEEE Trans. Power Syst. 36 (5), 4754–4763.
Peng, Yifan, Tang, Yuxing, Lee, Sungwon, Zhu, Yingying, Summers, Ronald M., Lu, Zhiyong, 2020. Covid-19-ct-cxr: a freely accessible and weakly labeled chest x-ray and ct image collection on covid-19 from biomedical literature. IEEE Trans. Big Data 7 (1), 3–12.
Qin, Lan, Li, Weide, 2020. A combination approach based on seasonal adjustment method and echo state network for energy consumption forecasting in usa. Energy Effic. 13 (7), 1505–1524.
Şahin, Utkucan, Ballı, Serkan, Chen, Yan, 2021. Forecasting seasonal electricity generation in european countries under covid-19-induced lockdown using fractional grey prediction models and machine learning methods. Appl. Energy 302, 117540.
Selvin, Sreelekshmy, Vinayakumar, R., Gopalakrishnan, E.A., Menon, Vijay Krishna, Soman, K.P., 2017. Stock price prediction using lstm, rnn and cnn-sliding window model. In: 2017 International Conference on Advances in Computing, Communications and Informatics. IEEE, pp. 1643–1647.
Shahid, Farah, Zameer, Aneela, Muneeb, Muhammad, 2020. Predictions for covid-19 with deep learning models of lstm, gru and bi-lstm. Chaos Solitons Fractals 140, 110212.

Sharma, Ketaki, Koirala, Archana, Nicolopoulos, Katrina, Chiu, Clayton, Wood, Nicholas, Britton, Philip N., 2021. Vaccines for covid-19: Where do we stand in 2021? Paediatr. Respir. Rev. 39, 22–31.
Sun, Shufang, Lin, Danhua, Operario, Don, 2021. Interest in covid-19 vaccine trials participation among young adults in china: Willingness, reasons for hesitancy, and demographic and psychosocial determinants. Prev. Med. Rep. 22, 101350.
Wang, Qiang, Li, Shuyu, Jiang, Feng, 2021. Uncovering the impact of the covid-19 pandemic on energy consumption: New insight from difference between pandemic-free scenario and actual electricity consumption in china. J. Clean. Prod. 313, 127897.
Wei, Runchu, Chen, Xia, Chang, Chun-Ping, 2021. Does covid-19 pandemic hurt stock prices of solar enterprises? Econ. Anal. Policy 72, 41–57.
Wu, Yu-Huan, Gao, Shang-Hua, Mei, Jie, Xu, Jun, Fan, Deng-Ping, Zhang, Rong-Guo, Cheng, Ming-Ming, 2021. Jcs: An explainable covid-19 diagnosis system by joint classification and segmentation. IEEE Trans. Image Process. 30, 3113–3126.
Wu, Jinran, Levi, Noa, Araujo, Robyn, Wang, You-Gan, 2023. An evaluation of the impact of covid-19 lockdowns on electricity demand. Electr. Power Syst. Res. 216, 109015.
Wu, Jinran, Wang, You-Gan, 2022. A working likelihood approach to support vector regression with a data-driven insensitivity parameter. Int. J. Mach. Learn. Cybern. 1–17.
Yang, Ailing, Li, Weide, Yang, Xuan, 2019. Short-term electricity load forecasting based on feature selection and least squares support vector machines. Knowl.-Based Syst. 163, 159–173.
Yuan, Xiaofeng, Li, Lin, Shardt, Yuri A.W., Wang, Yalin, Yang, Chunhua, 2020. Deep learning with spatiotemporal attention-based lstm for industrial soft sensor model development. IEEE Trans. Ind. Electron. 68 (5), 4404–4414.
Zhao, Zixi, Wu, Jinran, Cai, Fengjing, Zhang, Shaotong, Wang, You-Gan, 2022. A statistical learning framework for spatial–temporal feature selection and application to air quality index forecasting. Ecol. Indic. 144, 109416.
Zhong, Haiwang, Tan, Zhenfei, He, Yiliu, Xie, Le, Kang, Chongqing, 2020. Implications of covid-19 for the electricity industry: A comprehensive review. CSEE J. Power Energy Syst. 6 (3), 489–495.
Zoran, Maria A., Savastru, Roxana S., Savastru, Dan M., Tautan, Marina N., Baschir, Laurentiu A., Tenciu, Daniel V., 2022. Assessing the impact of air pollution and climate seasonality on covid-19 multiwaves in madrid, spain. Environ. Res. 203, 111849.