Forecasting China's wastewater discharge using dynamic factors and mixed-frequency data

Ding, L.; Lv, Z.; Han, Meng; Zhao, X.; Wang, W.

Published in:
Environmental Pollution

DOI:
10.1016/j.envpol.2019.113148

Environmental Pollution 255 (2019) 113148

Contents lists available at ScienceDirect

Environmental Pollution

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

Forecasting China's wastewater discharge using dynamic factors and mixed-frequency data☆

Lili Ding [a, b], Zhanlei Lv [a], Meng Han [c, *], Xin Zhao [a, b], Wei Wang [a, b]

[a] School of Economics, Ocean University of China, China
[b] Marine Development Studies Institute of OUC, Key Research Institute of Humanities and Social Sciences at Universities, Ministry of Education, Qingdao, China
[c] Department of Economics and Business, University of Groningen, Netherlands

A R T I C L E I N F O

Article history:
Received 10 January 2019
Received in revised form
16 July 2019
Accepted 30 August 2019
Available online 31 August 2019

Keywords:
Wastewater discharge
MIDAS regression
Forecast combination

A B S T R A C T

Forecasting wastewater discharge is the basis for wastewater treatment and policy formulation. This paper proposes a novel mixed-data sampling regression model, i.e., combination-MIDAS model to forecast quarterly wastewater emissions in China based on dynamic factors at different frequencies. The results show that a significant auto-correlation for wastewater emissions exists and that water consumption per ten thousand gross domestic product is the best predictor of wastewater emissions. The forecast performances of the combination-MIDAS models are robust and better than those of the benchmark models. Therefore, the combination-MIDAS models can better capture the characteristics of wastewater emissions, suggesting that the proposed method is a good method to deal with model misspecification and uncertainty for the control and management of wastewater discharge in China.

1. Introduction

Water is critical to economic development and human life (Horne, 2013). However, the per capita annual sustainable freshwater available in China is only about one-quarter of the world average (Qin et al., 2013). Water pollution, one of the environmental crises, has become a major concern in China. The rapid development of the economy, industrialization, and urbanization without adequate investment in wastewater treatment has resulted in degraded water quality. More than 30% of the river length, 66% of major lakes, and 70% of groundwater wells do not meet the quality criteria for drinking water sources (Lu et al., 2017a). Many people inevitably consume unsafe water, which is contaminated with heavy metals, organic compounds and pathogens from wastewater discharge. Over 70% of Chinese residents feel threatened by water pollution (Wang and Yang, 2016). Evidence suggests that wastewater has a large impact on public health and has a significantly negative impact on the overall mortality rate (Lu et al., 2017b), which can pose a risk to further development and social stability in China. Therefore, the control and management of water pollution have become an extremely important problem for the Chinese government.

Due to the severity of water pollution, reducing the impacts of wastewater discharge is one of the focuses in China. To address this problem and adhere to the ecological priorities and green development, the Water Pollution Prevention and Control Action Plan ("10-Point Water Plan") was released in April 2015 to tackle the nation's water pollution crisis, including the pollution of ground water and surface water, which are considered the most severely deteriorated natural sources in China (Han et al., 2016). Although substantial reductions in the pollutant emissions into the water have been achieved with the government's endeavour to cure water pollution, the current total amount of wastewater discharge remains massive (Zhang et al., 2017). To control and remedy wastewater emissions, and therefore gain better performance in the management of water pollution, precisely forecasting the amount of wastewater emission is required. Accurate forecasting of the amount of wastewater discharge is essential to making more effective policies relative to wastewater control and management to achieve more sustainable economic development.

The amount of wastewater emissions is affected by many factors which can be classified into three types, namely, economic growth, industrial structure and urbanization. Academic literature on the drivers of wastewater emissions primarily focuses on the influence

☆ This paper has been recommended for acceptance by Sarah Harmon.
* Corresponding author.
E-mail address: m.han@rug.nl (M. Han).

https://doi.org/10.1016/j.envpol.2019.113148
0269-7491/

of economic growth. The rapid development of the economy increases the demand for water use and therefore results in more wastewater discharge, especially in industrial sectors (Li et al., 2014). Therefore, economic development is considered the dominating factor of industrial wastewater discharge (Geng et al., 2014). Some evidence shows that there is a significant inverted U-shaped EKC (Environmental Kuznets Curve) between income growth and industrial wastewater emissions in rich provinces in China (Qian et al., 2012; Yang et al., 2010; Zhou and Sun, 2013). There exists in the EKC relationship a rising left leg of the unobserved curve between the economy and the total amount of industrial wastewater emissions in China (Zhang et al., 2011). Moreover, some studies show that water consumption per ten thousand gross domestic product exerts a larger influence on wastewater emissions in the near future than in the long term using Wuhan City in China as a case study (Zhang et al., 2016).

The industrial structure suggests the development level of industrial sectors and industrial growth. The industry accounts for approximately 20% of the decrease in total freshwater from a global perspective (Duarte et al., 2014). Although China enters the "New Normal" and needs to upgrade the level of industrialization (Yan and Liu, 2015), the industrial sewage still increases and always does not meet the discharge standards. The extension of the industrial sector will directly lead to an increase in wastewater emissions without a responding improvement of wastewater treatment. This rapid growth results in a dilemma between the development of industrialization and the preservation of water resources including the treatment of industrial wastewater emissions (Yao et al., 2018). Therefore, it is regarded as one of the main determinants of wastewater emissions, especially industrial wastewater emissions, which is closely related to industrial production and will increase with the growing industrial output (Li et al., 2016a).

Urbanization is also considered one of the main factors of wastewater discharge and has mixed effects on water conditions. On the one hand, rapid urbanization can directly influence the domestic water consumption, which cannot be avoided because of human activities (Elleuch et al., 2018). On the other hand, it will accelerate industrial production by increasing household consumption. However, urbanization also can relieve the increase in industrial wastewater through improved industrial labour productivity. Some evidence suggests that there exists a reverse U-shaped relationship between the urbanization rate and industrial wastewater discharge (Somorowska and Łaszewski, 2019), while some studies indicate that urbanization may have a negative impact on environmental quality in the long run by increasing wastewater emissions (Li et al., 2016b). Moreover, the population has a significant impact on the contamination level in surface water (Zhu et al., 2018) and therefore can affect wastewater emissions (Hu and Guan, 2018). It has been estimated that population growth is an important driver of water efficiency, especially in the long run (Distefano and Kelly, 2017; Gu et al., 2009).

However, there still exist some research gaps in the cognition on the forecast of wastewater emissions. Limited literature pays attention to wastewater emissions prediction, mainly focusing on the application of the Grey model and neural networks by the historical change trends of wastewater emissions. The Grey model has been applied to forecast wastewater discharge, because it can ensure reliability and stability with inadequate datasets (Wang et al., 2017). There are some extended grey models to improve the forecast accuracy (Wang et al., 2018). For example, the Nonlinear Grey Bernoulli Model NGBM (1,1) is used to simulate the annual qualified discharge rate of industrial water in China and yields accurate performance (Wang et al., 2011). Artificial neural networks are other methods frequently used to forecast wastewater emissions (Wang and Yu, 2012). This technique has been widely applied to research on global climate changes because it has larger advantages in portraying nonlinear characteristics. Similarly, it has been used to forecast the amount of wastewater discharge (Fernandeza et al., 2009) including methane emissions (Du et al., 2018) and performs well.

Indeed, the above Grey model and neural networks utilizing the past change trends of wastewater discharge do not take relative factors into consideration. Wastewater generation is a complex process influenced by many drivers. The fluctuation of latent factors intensifies the fluctuation and uncertainty of wastewater emissions, which is a considerable obstacle for accurate predictions. Thus, it is important to take into account the considerable influence of exogenous variables when performing trend analysis of wastewater emissions. Therefore, multivariate models that can integrate the effects of several factors have gained growing attention when forecasting wastewater emissions. Specifically, the multivariate nonlinear regression model is applied to forecast industrial wastewater discharge in China considering the industrial gross product and industrial wastewater consumption and yields a higher precision (Lei and Pan, 2011). Moreover, an artificial neural network that uses generic indicators performs well in forecasting municipal waste generation for countries with highly diversified levels of economic development, industrial structure, productivity and output (Antanasijević et al., 2013). However, this method has the disadvantage of overfitting, which limits its application. Furthermore, the hidden Markov model considers latent factors and performs well when forecasting the fluctuation of wastewater generation (Jiang and Liu, 2016). Lastly, the Grey multivariable model based on the nonlinear least squares method can identify the nonlinear relationship between wastewater emissions and related drivers and present greater precision than the traditional Grey multivariable model (Zeng et al., 2019).

Nevertheless, the multivariate regression models mentioned above pay more attention to forecast annual wastewater using the same frequency predictors. Due to unanticipated changes in economic growth and the factors that affect wastewater discharge, the predictor determinations for wastewater emission should have dynamic characters. Since there are time lags between environmental policies and economic activities changes, the high-frequency predictors can help the forward-looking decision makers react before the annual wastewater emissions actually occur. It is meaningful to use monthly or quarterly factors such as economic growth, industries and consumption to forecast low-frequency wastewater emissions. On the other hand, the influence of some factors on wastewater discharge may last for even shorter periods than one quarter, and the effects will be ignored if the quarterly data are used directly or is added to annual data. Therefore, it is necessary to explore new methods that can make full use of the high-frequency data to forecast wastewater emissions with higher accuracy.

The mixed-data sampling (MIDAS) regression model is introduced to forecast wastewater emissions to solve the problem of different frequencies mentioned above. The MIDAS regression model provides an attractive way to deal with datasets with different frequencies because it involves parsimonious specification based on distributed lag polynomials (Ghysels and Qian, 2019). It is designed to keep a balance between retaining the effective information of high-frequency data and reducing the number of parameters which need to be estimated (Tsui et al., 2018). Therefore, it can enhance the prediction accuracy for a low-frequency variable through effectively utilizing high-frequency information, thus alleviating information loss resulting from data accumulation. However, the MIDAS regression model has not yet been applied to forecast wastewater emissions. In China, the data on wastewater emissions and some drivers are published annually, whereas the data on other drivers are published monthly or quarterly. The quarterly wastewater emissions measurements are more meaningful for policy making because they can provide better guidance to manufacturers and residents to change their plan of wastewater disposal according to relevant policies, even though most of the existing research focuses on annual forecasts because of data availability. Therefore, this paper focuses on the forecasting of quarterly wastewater emissions considering monthly multiple drivers based on mixed datasets., In order to improve the forecast accuracy, this paper constructs a combination-MIDAS regression model based on the single MIDAS forecast results to better capture the dynamic factors of wastewater emissions.

The intention of this paper is to forecast quarterly wastewater emissions in China considering monthly datasets of relevant factors. The main contributions can be classified into two categories. First, this paper analyses the relationship between quarterly wastewater emissions and monthly factors. The more comprehensive factor system is presented, which is critical to fully capturing the dynamic factors of wastewater emissions. Namely, the gross domestic product (GDP), water consumption per ten thousand gross domestic product (WC/TTGDP), value added of secondary industry (VASI), water consumption per value added of secondary industry (WC/VASI), value added of tertiary industry (VATI), urban population (UP), and household consumption (HC) are chosen to forecast the wastewater emissions. The results show that there exists a significant auto-correlation for wastewater emissions and that WC/TTGDP has the best predictive ability.

The second contribution of this paper is methodology. This paper constructs a combination-MIDAS regression model to deal with model uncertainty because the dynamic drivers may have different information sets and modelling structures. We combine the forecasts from the best single MIDAS models for each factor with five weight-type schemes to alleviate the problem of model uncertainty. The results show that the proposed combination-MIDAS models perform better than benchmark models, e.g. autoregressive (AR), moving average (MA), and autoregressive moving average (ARMA) models. Furthermore, the robustness analysis suggests that the forecast performances of the combination-MIDAS models are robust, which indicates that the combination-MIDAS model is a promising method for wastewater emissions forecasting.

The remainder of the paper is organized as follows. Section 2 introduces the data resources and methods employed in this paper. Section 3 presents an analysis of the driving factors of wastewater discharge. Section 4 demonstrates the analysis of the forecast of wastewater discharge, which suggests the efficacy of the models developed in this study. The research conclusions and future work are discussed in section 5.

2. Data resources and research methodology

2.1. Data resources

This paper aims to forecast the quarterly wastewater emissions in China, including industrial and domestic wastewater emissions and considering monthly factors. The existing literature discussed above suggests that wastewater discharge is a complex process influenced by many factors, such as economic growth, industrial structure and urbanization. To construct a comprehensive factor system for wastewater emissions, this paper chooses GDP and WC/TTGDP from the economic aspect, chooses VASI, WC/VASI and VATI from the industrial aspect, and chooses UP and HC from the urbanization aspect. The data are available in the China Statistical Yearbook, the Report on the State of the Environment in China, and the China Water Resources Bulletin. However, most data are published annually or quarterly in China. To detect the relationship among wastewater emissions and its factors in detail, this paper applies the “quadratic match sum” principle to obtain the quarterly wastewater emissions and monthly factors. The sample covers the period from January 1997 to December 2016. This paper chooses the period from January 1997 to March 2014 as the estimation sample to train single MIDAS regression models and chooses the period from April 2014 to December 2016 as the out-of-sample period for forecasting to compare the forecast performance of single MIDAS models.

The growth rates of variables are considered in the empirical analysis to eliminate heteroscedasticity. Namely, $growth_{it} = \ln(value_{it}/value_{it-1}) \times 100$, where $growth_{it}$ refers to the growth rate of the $i-th$ variable at time t. $value_{it}$ is the observation of the $i-th$ variable at time t.

2.2. MIDAS regression model

There exists some econometric analysis providing theoretical evidence for the advantage of the MIDAS models (Andreou et al., 2010; Ghysels et al., 2007; Kvedaras and Račkauskas, 2010). Moreover, the MIDAS model has been widely adopted to forecast economic variables and has yielded good performance. Many studies have revealed overwhelming support indicating that MIDAS has the ability to improve the forecast accuracy of low-frequency variables and outperforms linear time series models in both the in-sample and the out-of-sample forecast periods. Specifically, much evidence suggests that MIDAS performs better in forecasting economic variables than do traditional models, such as traditional linear regressions (Bangwayo-Skeete and Skeete, 2015; Jiang et al., 2017) and generalized autoregressive conditional heteroskedasticity (GARCH)-class models (Alper et al., 2012). Recently, some evidence has demonstrated that the MIDAS model performs well in energy and environmental research due to its advantages in addressing the problem of different sampling frequencies. It has been used to forecast energy demands (He and Lin, 2018), oil prices (Baumeister et al., 2015; Pan et al., 2017), carbon prices (Zhao et al., 2018b) and carbon dioxide emissions (Zhao et al., 2018a).

To improve the forecast accuracy of quarterly wastewater emissions using monthly drivers, MIDAS regression models are applied in this paper. MIDAS depends on parsimonious polynomials to reflect the dynamic relationship among data with different frequencies. The basic MIDAS model ($MIDAS(m,k)$) can be defined simply as:

$$Y_t = \alpha + \beta W(L^{1/m}, \theta) x_t^{(m)} + \varepsilon_t \quad (1)$$

where Y_t refers to quarterly wastewater emissions and $t = 1, 2, \cdots, T$. $x_t^{(m)}$ refers to the $i-th$ monthly indicators, which can be observed m times between quarters $t-1$ and t. Thus, in this paper, $m = 3$. α and β are unknown parameters. $L^{1/m}$ is the lag order, and $L^{k/m} x_t^{(m)} = x_{t-k/m}^{(m)}$. When $k = 0$, $x_{t-k/m}^{(m)}$ refers to the observation of the third month in a quarter. When $k = 1$, $x_{t-k/m}^{(m)}$ refers to the observation of the second month in a quarter, and so on. $W(L^{1/m}, \theta)$ is decided by the lag operator $L^{1/m}$ and a parameter vector of limited dimension, θ, which can be defined as $W(L^{1/m}, \theta) = \sum_{k=0}^{K} w(k;\theta) L^{k/m}$, where $w(k;\theta)$ is the polynomial weight and K is the maximum lag order of the relevant drivers.

Furthermore, to take advantage of the readily available higher frequency data, the MIDAS regression model with leads $MIDAS(m,k,h)$ is introduced. Suppose that we can obtain data on drivers in the second month of a quarter and during which we aim to achieve a forecast for wastewater emissions. Then, we can employ $MIDAS(m,k,h)$ with one-month leads or one-step ahead to incorporate real-time information. Therefore, $MIDAS(m,k,h)$ can provide forecast updates when new data are available. The $MIDAS(m,k,h)$ model can be represented by:

$$Y_t = \alpha + \beta W(L^{1/m}, \theta) x_{t-h/m}^{(m)} + \varepsilon_t \tag{2}$$

where h refers to the number of steps ahead. When $h = 1$, $MIDAS(m,k,h)$ can achieve the forecast of quarterly data in the quarter t with monthly data until the second month of the quarter t. Similarly, when $h > 3$, $MIDAS(m,k,h)$ can obtain the forecast for a quarter $t+1$ or later by using the data in the quarter t.

To fully utilize the information concerning high-frequency and low-frequency variables in the distributed lags, we employ $MIDAS(m,k,h)$ with an autoregressive distributed lag, namely, $AR-MIDAS(m, k, h)$. It takes the lag effects of Y_t into consideration. Therefore, it can minimize information loss and enhance forecast accuracy. The $AR-MIDAS(m,k,h)$ model can be represented by the following equation:

$$Y_t = \alpha + \sum_{j=1}^{p} \gamma_j Y_{t-j} + \beta W(L^{1/m}, \theta) x_{t-h/m}^{(m)} + \varepsilon_t \tag{3}$$

The MIDAS regression model depends on polynomial weights to capture the dynamic relationship between high-frequency data and low-frequency data. Thus, a suitable function form for the MIDAS model is important to achieve an accurate forecast. This paper considers various parsimonious polynomial specifications, including the beta density function with zero lag (*Beta*), beta density function with non-zero lag (*BetaNN*), exponential Almon lag polynomial (*ExpAlmon*), Almon lag polynomial (*Almon*), step function (*Step*) and unrestricted weight function (*UMIDAS*).

The beta density function is expressed as follows:

$$w(k;\theta) = w(k;\theta_1,\theta_2,\theta_3) = f(k/K,\theta_1,\theta_2) \Big/ \sum_{k=1}^{K} f(k/K,\theta_1,\theta_2) + \theta_3 \tag{4}$$

where $f(x_i,\theta_1,\theta_2) = x_i^{\theta_1 - 1}(1-x_i)^{\theta_2 - 1}\Gamma(\theta_1+\theta_2)/\Gamma(\theta_1)\Gamma(\theta_2)$ and $\Gamma(\theta) = \int_0^\infty e^{-x}x^{\theta-1}dx$. When $\theta_3 = 0$, $w(k;\theta) = w(k;\theta_1,\theta_2) = f(k/K,\theta_1,\theta_2)/\sum_{k=1}^{K} f(k/K,\theta_1,\theta_2)$, which is the *Beta* weight function. When $\theta_1 = 1$, $w(k;\theta) = w(k;1,\theta_2,\theta_3) = f(k/K,1,\theta_2)/\sum_{k=1}^{K} f(k/K,1,\theta_2) + \theta_3$, which is the *BetaNN* weight function.

The normalized exponential Almon lag polynomial (*ExpAlmon*) can be written as follows

$$w(k;\theta) = e^{(\theta_1 k + \theta_2 k^2 + \cdots + \theta_p k^p)} \Big/ \sum_{k=1}^{K} e^{(\theta_1 k + \theta_2 k^2 + \cdots + \theta_p k^p)} \tag{5}$$

The Almon lag polynomial specification (*Almon*) can be defined as

$$\beta w(k;\theta_0,\theta_1,\theta_2,\theta_3) = \sum_{p=0}^{3} \theta_p k^p \tag{6}$$

The polynomial specification with a step function (*Step*) can be expressed as

$$\beta w(k;\theta) = \theta_1 I_{i\in[a_0,a_1]} + \sum_{p=2}^{P} \theta_p I_{i\in[a_{p-1},a_p]}, I_{i\in[a_{p-1},a_p]} = \begin{cases} 1, & a_{p-1} \le i \le a_p \\ 0, & otherwise \end{cases} \tag{7}$$

where $a_0 = 1 < a_1 < \cdots < a_P = K$.

The unrestricted weight function (*UMIDAS*) generalizes to

$$Y_t = \alpha + B(\beta, L^{1/m}) x_t^{(m)} + \varepsilon_t \tag{8}$$

where $B(\beta, L^{1/m}) = \sum_{k=0}^{K} \beta_k L^{k/m}$.

2.3. Forecast combination method

In section 2.2, the forecast models use single predictor named by single MIDAS regression models. However, some evidences show that multiple predictors may provide more accurate results and more stable performance over time (Kuzin et al., 2013). There are two ways to include multiple predictors into to single MIDAS regression models to form the forecast. One is the multivariate MIDAS regression method, which extends the single MIDAS regression by adding more predictors as explanatory variables in the regression model (Andreou et al., 2013). But this method has to suffer the parameter proliferation problem, when there is a great deal of predictors. The other is the forecast combination method, which constructs a weighted average of forecasts using single MIDAS regression with different predictors (please see the detail in Fig. 1). This approach can solve the large number of predictors and achieve more stable forecasts. Specifically, it provides a way to deal with model uncertainty and is more robust to misspecification biases, measurement errors and structural breaks (Ghysels and Ozkan, 2015) since it does not depend on identifying a single best model (Pettenuzzo et al., 2016). In this paper, we use forecast combination method to introduce multiple predictors into single MIDAS regression models. Given N forecast results from the single MIDAS models, the forecast combination can be represented as:

$$\hat{f}_{N,T+s|T} = \sum_{j=1}^{N} \hat{w}_{j,T}\, \hat{y}_{j,T+s|T} \tag{9}$$

where $\hat{f}_{N,T+s|T}$ refers to the $s-th$ forecast results from the forecast combination method, denoted by combination-MIDAS model. T represents the number of estimation samples to train single MIDAS models. $\hat{y}_{j,T+s|T}$ is the $s-th$ forecast result from the best single MIDAS model considering the $j-th$ factor. $\hat{w}_{j,T}$ suggests that the weight given to the forecasts results depends on the best single MIDAS model considering the $j-th$ factor. This paper considers five weight types for $\hat{w}_{j,T}$.

(i) MSFE-weighted type

MSFE refers to the mean squared forecast error, which is employed to provide a reference to define the weight to each best single MIDAS model. The weight is given as

$$w_{j,T} = m_{j,T}^{-1} \Big/ \sum_{j=1}^{n} m_{j,T}^{-1} \tag{10}$$

where $m_{j,T} = \sum_{i=T_0}^{t} (\delta^{i-T_0}(y_{j,T+s} - \hat{y}_{j,T+s|T}))^2/(t - T_0 + 1)$. When $\delta = 1$, $m_{j,T}$ is the MSFE of the best single MIDAS model that considers the $j-th$ factor and that is trained by estimation samples including T observations. $y_{j,T+s}$ is the $s-th$ real observation in the out-of-sample data. $t-T_0+1$ refers to the number of out-of-sample data.

(ii) DMSFE-weighted type

DMSFE is the discounted mean squared forecast error based on MSFE. For example, set $\delta = 0.9$ shown above.

Namely, when $m_{j,T} = \sum_{i=T_0}^{t} (.0.9^{i-T_0} (y_{j,T+s} - \hat{y}_{j,T+s|T}))^2 / (t - T_0 + 1)$, $w_{j,T} = m_{j,T}^{-1} / \sum_{j=1}^{n} m_{j,T}^{-1}$ refers to the discounted mean squared forecast error (DMSFE)-weighted type.

(iii) AIC-weighted type

AIC suggests Akaike information criteria (AICs). The AIC-weighted type is defined as follows:

$$w_{j,T} = \exp(-AIC_j) \Bigg/ \sum_{j=1}^{N} \exp(-AIC_j) \tag{11}$$

(iv) BIC-weighted type

BIC refers to Bayesian information criteria. This kind of weight is shown as follows:

$$w_{j,T} = \exp(-BIC_j) \Bigg/ \sum_{j=1}^{N} \exp(-BIC_j) \tag{12}$$

(v) Equal-weighted type

The equal-weighted type is simply defined as

$$w_{j,T} = 1/N \tag{13}$$

Therefore, the process of combination-MIDAS includes 3 steps to forecast quarterly wastewater emission, as shown in Fig. 1.

Step 1 refers to the establishment of large numbers of single MIDAS models with single predictor, e.g., GDP, WC/TTGDP, VASI, WC/VASI, VATI, UP or HC. Step 2 chooses the best polynomial weight and lag orders for wastewater emissions and the relevant factors by comparing root mean squared errors (RMSE) with single MIDAS regression models, which is consistent with Andreou et al. (2010), Bai et al. (2013), Han et al. (2019). RMSE is a good index that can test the performance of out-of-sample forecasts. The single MIDAS regression model with the smallest RMSE is the one with the highest forecast accuracy. We call it the best single MIDAS regression model. Therefore, the indicator with the best predictive ability can also be chosen by comparing RMSE values. RMSE is defined as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{14}$$

where n refers to the number of out-of-sample data, y_i suggests the $i-th$ real observation, and $\hat{y}_i$ is the $i-th$ out-of-sample forecast results.

Step 3 combines the best single MIDAS models for each factor as chosen by step 2 with the weight schemes shown above to construct the combination-MIDAS model and applies it to forecast wastewater emissions.

3. Analysis of driving factors of wastewater discharge

3.1. Selection of the single MIDAS models

This section analyses the forecast performance of the single MIDAS regression models discussed in section 2.2. The single MIDAS models with the highest accuracy and the indicators with the best predictive ability are chosen by comparing RMSE values, which is in line with Han et al. (2019). The selection of the best single MIDAS model is a critical step to obtain final results with higher accuracy because it is the basis of the combination-MIDAS model and thus influences the final forecast results.

In line with Han et al. (2019), Jiang et al. (2017), and Zhao et al. (2018b), the best lag orders and polynomial weights for wastewater discharge and its factors are selected by comparing the RMSE values of relevant single MIDAS models. To reflect the change trends of RMSE values, the maximum lag orders for monthly factors are changed from 1 to 30 months, the maximum lag order for wastewater emissions ranges from 0 to 4 quarters, and h ranges from 0 to 3 months. This paper takes the single MIDAS models considering the gross domestic product (GDP) as an example to demonstrate the mechanism of choosing the best polynomial weight and the best lag orders for wastewater emissions and GDP, which is shown in Table 1. Table 1 shows the RMSEs of single MIDAS models with 2-month leads of wastewater emissions, namely, models in which the h step is 2 months, which suggests that the best polynomial weight and the best lag orders for GDP may be different under different lag orders for wastewater emissions. Generally, the polynomial weight of UMIDAS performs better in reflecting the relationship between wastewater emissions and GDP with the condition of 2-month leads of wastewater emissions. Moreover, there exists an auto-correlation for wastewater emissions since the RMSE is smaller when we consider the lag order for wastewater emissions. The smaller RMSE means that the forecast accuracy is improved. Furthermore, in the case considering MIDAS models with 2-month leads, the RMSE is the smallest when the lag orders for wastewater emissions and GDP are 4 quarters and 8 months, respectively.

3.2. Comparisons among the dynamic factors

Considering the complex processes among wastewater emissions and factors, this paper chooses the best polynomial weight, the best h steps, and the best lag orders for wastewater emissions and factors, respectively, to decide the best single MIDAS models with different leads. Similar to the process shown above, the best single MIDAS models are selected, which are shown in Table 2. The results in Table 2 suggest that there exists an auto-correlation

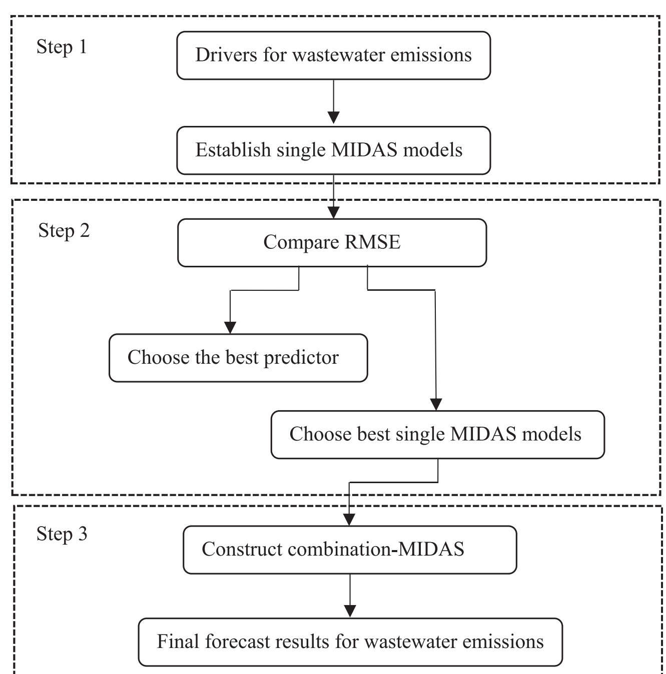

Fig. 1. Flow chart of the combination-MIDAS model.

among wastewater emissions. This auto-correlation provides evidence for the research of Wang et al. (2011), which only consider wastewater discharge time series. Moreover, the mechanisms by which factors influence wastewater emissions are different, which provide evidence that we should determine the best polynomial weight to reflect the influence of each factor on wastewater emissions rather than predetermining one before analysis for all factors.

Specifically, the best MIDAS models with each factor under different conditions are the ones whose RMSEs are bolded in Table 2. According to the best MIDAS models, we suggest that UMIDAS can better capture the impacts of GDP and urban population (UP) on wastewater emissions. The influence of GDP and UP lasts for 8 and 16 months respectively, which is different from the findings of Zhang et al. (2016). Zhang et al. (2016) hold that GDP growth rate impacts wastewater emissions more significantly in the long term, which may because that they only consider the municipal wastewater emissions in Wuhan, China using annual data directly. For water consumption per ten thousand gross domestic product (WC/TTGDP), water consumption per value added of secondary industry (WC/VASI) and the household consumption (HC), their impacts on wastewater emissions can be represented by a step function. The impacts of WC/TTGDP, WC/VASI, and HC last for 30, 28 and 24 months. Regarding the value added of secondary industry (VASI), the BetaNN polynomial weight can better capture the relationship between this factor and wastewater emissions, which lasts for 26 months. For the value added of tertiary industry (VATI), the Almon polynomial weight performs better, and this factor's influence on wastewater emissions lasts for 3 months. Furthermore, the best single MIDAS models relevant to WC/TTGDP have the smallest RMSE value among the best single MIDAS models with different factors, which means that WC/TTGDP, as an economic proxy, is the best predictor for wastewater emissions among the factors discussed above. This is to some degree in line with the findings of Geng et al. (2014), which suggests that economic factors are the main determinants for industrial wastewater emissions.

Table 1
RMSEs of single MIDAS models considering GDP with 2-month leads.

Weights	Lag order for GDP						
	3	7	8	16	20	25	30
Lag order for wastewater emissions: 0							
Beta	0.2549	0.2540	0.2542	0.2677	0.2694	0.2682	0.2669
BetaNN	0.2553	0.2534	0.2570	0.2479	0.2705	0.2679	0.2636
ExpAlmon	0.2542	0.2545	0.2642	0.2787	0.2732	0.2745	0.2745
Almon	0.2494	0.3366	0.2591	0.2742	0.2516	0.2391	0.2631
Step	0.2592	0.2543	0.2537	0.3169	0.3095	0.2960	0.2840
UMIDAS	0.2408	**0.2194**	0.2373	0.4905	0.4359	0.5723	1.5260
Lag order for wastewater emissions: 1							
Beta	0.1753	0.1753	0.1753	0.1837	0.1850	0.1835	0.1826
BetaNN	0.1733	0.1712	0.1740	0.1703	0.1871	0.1790	0.1803
ExpAlmon	0.1753	0.1753	0.1753	0.1753	0.1753	0.1753	0.1753
Almon	0.2411	0.2334	0.1881	0.1850	0.1759	0.1784	0.1875
Step	0.1741	0.1738	0.1739	0.2264	0.2217	0.2656	0.2821
UMIDAS	**0.1681**	0.1682	0.1687	0.5117	0.5648	0.5960	1.1605
Lag order for wastewater emissions: 2							
Beta	0.1710	0.1710	0.1710	0.1813	0.1816	0.1797	0.1788
BetaNN	0.1689	0.1669	0.1698	0.1670	0.1686	0.1781	0.1856
ExpAlmon	0.1710	0.1710	0.1710	0.1710	0.1710	0.1710	0.1710
Almon	0.6229	0.2252	0.1871	0.1822	0.1723	0.1750	0.1848
Step	0.1698	0.1699	0.1702	0.2219	0.2171	0.2582	0.2800
UMIDAS	0.1644	**0.1637**	0.1648	0.5093	0.5594	0.6193	1.2090
Lag order for wastewater emissions: 3							
Beta	0.1760	0.1760	0.1760	0.1828	0.1848	0.1836	0.1828
BetaNN	0.1740	0.1719	0.1746	0.1709	0.1723	0.1792	0.1888
ExpAlmon	0.1760	0.1760	0.1760	0.1760	0.1760	0.1760	0.1760
Almon	0.1964	0.2363	0.1855	0.1841	0.1744	0.1773	0.1865
Step	0.1748	0.1744	0.1745	0.2224	0.2194	0.2699	0.2896
UMIDAS	**0.1692**	0.1724	0.1731	0.5301	0.5769	0.6235	1.1863
Lag order for wastewater emissions: 4							
Beta	0.2094	0.2094	0.2094	0.2094	0.1956	0.1989	0.2006
BetaNN	0.2062	0.2032	0.2046	0.2007	0.1987	0.1987	0.2090
ExpAlmon	0.2094	0.2094	0.2093	0.2094	0.2094	0.2094	0.2094
Almon	0.7042	0.2401	0.1803	0.1878	0.1864	0.1909	0.2031
Step	0.2091	0.2058	0.2053	0.2219	0.2242	0.2551	0.2668
UMIDAS	0.2021	0.1544	**0.1540**	0.3638	0.4883	0.6196	1.1386

Notes: The bold values represent the smallest RMSEs under the condition of different lag orders for wastewater emissions.

4. Analysis of forecast of wastewater discharge

4.1. Forecast results of the combination-MIDAS model

One model cannot maintain a dominant performance all the time because of model uncertainty. Therefore, this paper constructs combination-MIDAS models based on the best single MIDAS models under each condition shown in Table 2, which is discussed in section 2.2. The best polynomial weight and best lag orders for each factor are different, thus, the combination-MIDAS model can make full use of the advantages of the best single MIDAS models. Consistent with Han et al. (2019) and Zhao et al. (2018b), this paper applies the best single MIDAS models considering each factor under the condition of different leads shown in Table 2 to achieve the forecast for wastewater emissions. Then, the forecast results obtained from the best single MIDAS models are given corresponding weights according to the five weight schemes to achieve the final forecasts of the combination-MIDAS models for wastewater emissions.

Table 3 demonstrates the out-of-sample forecast performance of the combination-MIDAS models. Panel A shows the RMSEs of the combination-MIDAS models, which suggests that the DMSFE weight can better capture the advantages of the best single MIDAS models since the combination-MIDAS model with this type of weight has the smallest RMSE value. This finding is consistent with that of Ghysels and Ozkan (2015), which only applies DMSFE weight to construct forecast combination method since the discount factor attaches greater weight to the recent predictive power of individual factors. Moreover, comparing with forecast performance of the best single MIDAS models shown in Table 2, we find that the combination-MIDAS models outperform most of the best single MIDAS models, which is in line with the findings of Zhao et al. (2018b). Especially, the forecast performances of the combination-MIDAS models given the best combination weight perform better than do the best single MIDAS models considering all other factors except WC/TTGDP, which is slightly better than the combination-MIDAS models.

Furthermore, we compare the forecast performances of the combination-MIDAS models with those of the benchmark models using quarterly data, e.g. AR(1), MA(1) and ARMA(1, 1) models in order to illustrate the effectiveness and feasibility of the combination-MIDAS models. Table 4 shows the RMSE ratios of the combination-MIDAS models to that of the benchmark models. It is demonstrated that the forecast performances of the combination-MIDAS models are obviously better than those of the benchmark models. This finding suggests that related factors can help forecast wastewater emissions and that the combination-MIDAS models can take advantage of the useful information in monthly factors and thus can significantly improve the forecast accuracy of wastewater emissions.

4.2. Robustness checks

In this section, we investigate the robustness of the forecast performance of the combination-MIDAS models discussed in section 4.1. We consider not only the randomness of sample selection

Table 2
The best single MIDAS models with different leads.

Factors	Model	RMSE	Model	RMSE
	H = 0		H = 1	
GDP	AR(2)-BetaNN-MIDAS(3,21,0)	0.1556	AR(2)-BetaNN-MIDAS(3,15,1)	0.1563
WC/TTGDP	AR(1)-Step-MIDAS(3,28,0)	0.0762	AR(1)-Step-MIDAS(3,28,1)	0.0622
VASI	AR(2)-Almon-MIDAS(3,3,0)	0.1156	AR(2)-BetaNN-MIDAS(3,26,1)	0.1189
WC/VASI	AR(2)-U-MIDAS(3,9,0)	0.0982	AR(2)-Step-MIDAS(3,28,1)	**0.0976**
VATI	AR(2)-Beta-MIDAS(3,4,0)	0.1682	AR(2)-Almon-MIDAS(3,3,1)	**0.1648**
UP	AR(3)-U-MIDAS(3,16,0)	**0.1184**	AR(3)-U-MIDAS(3,15,1)	0.1206
HC	AR(2)-Step-MIDAS(3,24,0)	**0.1440**	AR(2)-Almon-MIDAS(3,16,1)	0.1439
	H = 2		H = 3	
GDP	AR(4)-U-MIDAS(3,8,2)	**0.1540**	AR(2)-U-MIDAS(3,5,3)	0.1541
WC/TTGDP	AR(2)-Step-MIDAS(3,30,2)	**0.0573**	AR(4)-Almon-MIDAS(3,29,3)	0.0873
VASI	AR(1)-BetaNN-MIDAS(3,26,2)	**0.1078**	AR(3)-BetaNN-MIDAS(3,29,3)	0.1366
WC/VASI	AR(2)-U-MIDAS(3,8,2)	0.1130	AR(1)-U-MIDAS(3,7,3)	0.1255
VATI	AR(2)-Almon-MIDAS(3,3,2)	0.1666	AR(2)-Step-MIDAS(3,3,3)	0.1696
UP	AR(2)-Step-MIDAS(3,3,2)	0.1669	AR(2)-Almon-MIDAS(3,5,3)	0.1659
HC	AR(2)-Almon-MIDAS(3,7,2)	0.1448	AR(2)-Almon-MIDAS(3,10,3)	0.1446

Notes: GDP refers to gross domestic product; WC/TTGDP refers to water consumption per ten thousand gross domestic product; VASI refers to value added of secondary industry; WC/VASI refers to water consumption per value added of secondary industry; VATI refers to value added of tertiary industry; UP refers to urban population; HC refers to household consumption. The bold value is the smallest RMSE obtained through comparison among the best single MIDAS models. The models are displayed as AR-weight-MIDAS(m,k,h). The bolded values are the smallest RMSEs of the best MIDAS models with each factor among different conditions of h.

Table 3
Forecast performance for combination-MIDAS models.

Weight	h = 0	h = 1	h = 2	h = 3
Panel A: RMSE				
MSFE	0.0888	0.0805	0.0905	0.1144
DMSFE	**0.0848**	**0.0795**	**0.0832**	0.1051
AIC	0.0881	0.1400	0.1426	**0.0874**
BIC	0.1612	0.1439	0.1507	0.1270
Equal Weights	0.1062	0.1036	0.1242	0.1319

Notes: The bolded values refer to the smallest RMSEs under the condition of different leads.

Table 4
Comparison of the combination-MIDAS models with benchmark models.

Weight	h = 0	h = 1	h = 2	h = 3
Panel A: AR(1)				
MSFE	0.1740	0.1578	0.1773	0.2242
DMSFE	0.1662	0.1558	0.1630	0.2059
AIC	0.1728	0.2744	0.2796	0.1713
BIC	0.3160	0.2821	0.2954	0.2490
Equal Weights	0.2081	0.2030	0.2435	0.2585
Panel B: MA(1)				
MSFE	0.1029	0.0933	0.1049	0.1326
DMSFE	0.0983	0.0921	0.0964	0.1218
AIC	0.1022	0.1623	0.1653	0.1013
BIC	0.1868	0.1668	0.1747	0.1472
Equal Weights	0.1230	0.1200	0.1440	0.1529
Panel C: ARMA(1,1)				
MSFE	0.2181	0.1977	0.2223	0.2810
DMSFE	0.2083	0.1952	0.2043	0.2581
AIC	0.2165	0.3439	0.3504	0.2146
BIC	0.3960	0.3535	0.3702	0.3121
Equal Weights	0.2608	0.2544	0.3052	0.3240

Note: If the RMSE ratios are smaller than 1, the forecast accuracy of the combination-MIDAS models is higher than benchmark models.

but also the frequency of wastewater emissions.

Because the forecast performance of a model may be sensitive to sample selections (Zhang et al., 2015), we test the robustness of the combination-MIDAS models by selecting a different estimation sample. In particular, we choose the period from January 1997 to December 2010 as the estimation sample and choose the period from January 2011 to December 2016 as the out-of-sample period.

Table 5
Forecast results analysis of the combination-MIDAS models based on the new sample selection.

Weight	h = 0	h = 1	h = 2	h = 3
Panel A: AR(1)				
MSFE	0.3468	0.3513	0.3386	0.3627
DMSFE	**0.2891**	**0.2888**	**0.2711**	**0.3091**
AIC	0.5856	0.5769	0.5794	0.5293
BIC	0.5693	0.5913	0.5935	0.5994
Equal Weights	0.4695	0.4747	0.4585	0.4379
Panel B: MA(1)				
MSFE	0.2164	0.2192	0.2113	0.2264
DMSFE	**0.1804**	**0.1802**	**0.1692**	**0.1929**
AIC	0.3655	0.3601	0.3616	0.3304
BIC	0.3553	0.3690	0.3704	0.3741
Equal Weights	0.2930	0.2963	0.2861	0.2733
Panel C: ARMA(1,1)				
MSFE	0.4319	0.4376	0.4218	0.4518
DMSFE	**0.3601**	**0.3597**	**0.3376**	**0.3850**
AIC	0.7295	0.7186	0.7217	0.6594
BIC	0.7092	0.7365	0.7393	0.7466
Equal Weights	0.5848	0.5913	0.5711	0.5454

Note: If the RMSE ratios are smaller than 1, the forecast accuracy of the combination-MIDAS models is higher than benchmark models. The bolded values are the smallest RMSE rations under each condition of h.

Then we construct combination-MIDAS models with similar mechanism discussed above and compare the forecast performance with benchmark models estimated with the new estimation sample. The results of the robustness analysis are shown in Table 5, which suggests that the DMSFE outperforms among the five combination weights because the RMSE ratios of the combination-MIDAS models with DMSFE weight are smaller than those with other weights under each condition of *h*. Moreover, the forecast performances of the combination-MIDAS models based on the new sample selection are better than the benchmark models.

Furthermore, we investigate whether the combination-MIDAS models discussed in section 4.1 are robust to annual wastewater emissions. Similarly, we reconstruct the combination-MIDAS models to forecast annual wastewater emissions with monthly factors. Because the combination-MIDAS models are robust to sample selection as discussed above and in order to have enough observations to evaluate the forecast accuracy of the combination-MIDAS models, we determine the out-samples according to Turhan

Table 6
Forecast results analysis of the combination-MIDAS models based on annual wastewater emissions.

Weight	h = 0	h = 1	h = 2	h = 3
Panel A: AR(1)				
MSFE	0.3587	0.3579	0.3078	0.3945
DMSFE	**0.3518**	**0.3510**	**0.3038**	**0.3876**
AIC	0.5636	0.6445	0.4392	0.9076
BIC	0.5636	0.6445	0.4392	0.9083
Equal Weights	0.4722	0.4665	0.4243	0.5267
Panel B: MA(1)				
MSFE	0.2262	0.2256	0.1940	0.2487
DMSFE	**0.2218**	**0.2213**	**0.1916**	**0.2444**
AIC	0.3554	0.4063	0.2769	0.5723
BIC	0.3554	0.4063	0.2769	0.5727
Equal Weights	0.2977	0.2941	0.2675	0.3321
Panel C: ARMA(1,1)				
MSFE	0.2134	0.2129	0.1831	0.2347
DMSFE	**0.2093**	**0.2088**	**0.1808**	**0.2306**
AIC	0.3353	0.3834	0.2613	0.5400
BIC	0.3353	0.3834	0.2613	0.5404
Equal Weights	0.2809	0.2775	0.2524	0.3133

Note: If the RMSE ratios are smaller than 1, the forecast accuracy of the combination-MIDAS models is higher than benchmark models. The bolded values are the smallest RMSE rations under each condition of h.

Table 7
Forecast results for wastewater emissions (%).

Date	March-2017	June-2017	September-2017	December-2017
H	3	6	9	12
Forecast results	1.1349	1.1480	1.1354	1.0926

et al. (2015). We select the year 1997–2010 as the estimation sample and select year 2011–2016 as the out-of-sample period. Then, we compare the forecast performance of the reconstructed combination-MIDAS models with the benchmark models using annual data. The results of the robustness analysis are shown in Table 6, which suggests that the DMSFE weight remains better performance when constructing the combination-MIDAS models and that the combination-MIDAS models are still robust to annual wastewater emissions given the better performance than the benchmark models.

4.3. Application of the combination-MIDAS models

The robustness analysis discussed above suggests that the combination-MIDAS models constructed in section 4.1 to forecast quarterly wastewater emissions with monthly factors are robust. Therefore, we apply the combination-MIDAS models to achieve short-term forecast for the wastewater emissions in the future one year, which is of interest to policy-makers to manage the wastewater emissions. Consistent with the way introduced by Zhao et al. (2018b), we apply the combination-MIDAS models based on the best single MIDAS models with three-, six-, nine-, and twelve-month leads of wastewater emissions to conduct the short-term forecasts with the monthly factors until December 2016. We construct the combination-MIDAS models with DMSFE weight because this type of weight performs better and are robust suggested by the previous analysis. The forecast results are shown in Table 7, which shows that the total amount of wastewater emissions in China keeps increasing with a stable growth rate in 2017.

5. Conclusions

This paper proposes combination-MIDAS regression models to forecast quarterly wastewater emissions in China with monthly factors based on datasets with different frequencies. To better capture the complex process of wastewater emissions including industrial and domestic emissions, this paper selects GDP and WC/TTGDP from the economic aspect, VASI, WC/VASI, and VATI from the industrial aspect, and UP and HC from the urbanization aspect. The combination-MIDAS model is based on the best single MIDAS models, therefore, it can deal with model misspecification and solve the weakness of wastewater emissions forecasting to some degree.

First, large numbers of single MIDAS models are analysed. The forecast performances of single MIDAS models suggest that there exists a significant auto-correlation for wastewater emissions and that WC/TTGDP is the best predictor for wastewater emissions. Moreover, the impacts of WC/TTGDP, VASI, WC/VASI, UP, and HC on wastewater emissions last longer than 1 year. An alternative explanation is that the industrial production pattern, urban population and the household consumption habit are hard to change in the short term, which should draw more attention when the government makes policies relevant to wastewater emissions. Furthermore, the influence of GDP and VATI on wastewater emissions lasts only for several months, which is less than 1 year. This result suggests the reasonability of employing MIDAS models with mixed datasets since the short influence may be ignored if annual data are directly used to forecast wastewater emissions. Furthermore, the polynomial weight and h steps vary for the best single MIDAS models considering different factors, which shows the rationality of constructing combination-MIDAS models to make full use of the advantages of the best single MIDAS models.

Second, this paper constructs the combination-MIDAS model with five different weight schemes on the basis of the forecast results from the best single MIDAS models. The DMSFE weight performs better than do the other weight types. The forecast accuracy of the combination-MIDAS model is higher than those of the benchmark models. Then we test the robustness analysis from two aspects. The results suggest that the forecast performances of the combination-MIDAS model are robust not only to sample selection but also to the frequency of wastewater emissions. Therefore, the combination-MIDAS models can capture the characteristics of wastewater emissions and reflect the complex processes among wastewater discharge and its factors. This method provides a reference for the forecasting and the forecast update of wastewater emissions. The short-term forecast results indicate that the total amount of wastewater emissions in China will increase steadily.

In future research, we will expand our model and apply it to other countries, especially developing countries, where the problem of water pollution is more serious. Furthermore, we will also aim to establish a more comprehensive factor system for wastewater emissions forecasts, because wastewater emission is a complex process including many other factors in addition to those considered in this paper.

Acknowledgements

This work is supported by National Science Foundation of China (71973132, 71471105, 71701189), Shandong Society Science funds (No. 19CGLJ30), Fundamental Research Funds for the Central Universities (201762024), Natural Resource ministry Fund (No. CAMA201815), Taishan Scholar Program (tsqn20161014, ts201712014), and Research Fund (No. 15ZDB171), China Scholarship Council (CSC).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2019.113148.

References

Alper, C.E., Fendoglu, S., Saltoglu, B., 2012. MIDAS volatility forecast performance under market stress: evidence from emerging stock markets. Econ. Lett. 117, 528–532.
Andreou, E., Ghysels, E., Kourtellos, A., 2010. Regression models with mixed sampling frequencies. J. Econom. 158, 246–261.
Andreou, E., Ghysels, E., Kourtellos, A., 2013. Should macroeconomic forecasters use daily financial data and how? J. Bus. Econ. Stat. 31, 240–251.
Antanasijević, D., Pocajt, V., Popović, I., Redžić, N., Ristić, M., 2013. The forecasting of municipal waste generation using artificial neural networks and sustainability indicators. Sustain. Sci. 8, 37–46.
Bai, J., Ghysels, E., Wright, J.H., 2013. State space models and MIDAS regressions. Econom. Rev. 32, 779–813.
Bangwayo-Skeete, P.F., Skeete, R.W., 2015. Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. Tour. Manag. 46, 454–464.
Baumeister, C., Guérin, P., Kilian, L., 2015. Do high-frequency financial data help forecast oil prices? The MIDAS touch at work. Int. J. Forecast. 31, 238–252.
Distefano, T., Kelly, S., 2017. Are we in deep water? Water scarcity and its limits to economic growth. Ecol. Econ. 142, 130–147.
Du, M., Zhu, Q., Wang, X., Li, P., Yang, B., Chen, H., Wang, M., Zhou, X., Peng, C., 2018. Estimates and predictions of methane emissions from wastewater in China from 2000 to 2020. Earth's Future 6, 252–263.
Duarte, R., Pinilla, V., Serrano, A., 2014. Looking backward to look forward: water use and economic growth from a long-term perspective. Appl. Econ. 46, 212–224.
Elleuch, B., Bouhamed, F., Elloussaief, M., Jaghbir, M., 2018. Environmental sustainability and pollution prevention. Environ. Sci. Pollut. Control Ser. 25, 18223–18225.
Fernandeza, F.J., Secob, A., Ferrerc, J., Rodrigoa, M.A., 2009. Use of neurofuzzy networks to improve wastewater flow-rate forecasting. Environ. Model. Softw 24, 686–693.
Geng, Y., Wang, M., Sarkis, J., Xue, B., Zhang, L., Fujita, T., Yu, X., Ren, W., Zhang, L., Dong, H.J., 2014. Spatial-temporal patterns and driving factors for industrial wastewater emission in China. J. Clean. Prod. 76, 116–124.
Ghysels, E., Ozkan, N., 2015. Real-time forecasting of the US federal government budget: a simple mixed frequency data regression approach. Int. J. Forecast. 31, 1009–1020.
Ghysels, E., Qian, H., 2019. Estimating MIDAS regressions via OLS with polynomial parameter profiling. Econometrics and Statistics 9, 1–16.
Ghysels, E., Sinko, A., Valkanov, R., 2007. MIDAS Regressions: further results and new directions. Econom. Rev. 26, 53–90.
Gu, K., Liu, J., Wang, Y., 2009. Relationship between economic growth and water environmental quality of Anshan city in Northeast China. Chin. Geogr. Sci. 19, 17–24.
Han, D., Currell, M.J., Cao, G., 2016. Deep challenges for China's war on water pollution. Environ. Pollut. 218, 1222–1233.
Han, M., Ding, L., Zhao, X., Kang, W., 2019. Forecasting carbon prices in the Shenzhen market, China: the role of mixed-frequency factors. Energy 171, 69–76.
He, Y., Lin, B., 2018. Forecasting China's total energy demand and its structure using ADL-MIDAS model. Energy 151, 420–429.
Horne, J., 2013. Economic approaches to water management in Australia. Int. J. Water Resour. Dev. 29, 526–543.
Hu, J., Guan, H., 2018. Research on driving factors of wastewater discharge based on the STIRPAT model and Tapio decoupling model: a case in Chongqing. Environ. Pollut. Control 3, 355–359.
Jiang, P., Liu, X., 2016. Hidden Markov model for municipal waste generation forecasting under uncertainties. Eur. J. Oper. Res. 250, 639–651.
Jiang, Y., Guo, Y., Zhang, Y., 2017. Forecasting China's GDP growth using dynamic factors and mixed-frequency data. Econ. Modell. 66, 132–138.
Kuzin, V., Marcellino, M., Schumacher, C., 2013. Pooling versus model selection for nowcasting GDP with many predictors: empirical evidence for six industrialized countries. J. Appl. Econom. 28, 392–411.
Kvedaras, V., Račkauskas, A., 2010. Regression models with variables of different Frequencies: the case of a fixed frequency ratio. Oxf. Bull. Econ. Stat. 72, 600–620.
Lei, L., Pan, H.L., 2011. Prediction of industrial wastewater discharge amount based on multivariate nonlinear regression. Journal of Jiangnan University (Natural Science Edition) 10, 309–313.
Li, Z., Liu, H., Luo, C., Li, P., Li, H., Xiong, Z., 2014. Industrial wastewater discharge retrieval based on stable nighttime light imagery in China from 1992 to 2010. Remote Sens. 6, 7566–7579.
Li, N., Wei, Y., Wang, L., Zeng, C., Ma, X., Wu, H., 2016a. Impact of industrialization on water protection in the Huai river basin within Shandong province, China. Nat. Hazards 81, 1193–1207.
Li, T., Wang, Y., Zhao, D., 2016b. Environmental Kuznets Curve in China: new evidence from dynamic panel analysis. Energy Policy 91, 138–147.
Lu, B., Du, X., Huang, S., 2017a. The economic and environmental implications of wastewater management policy in China: from the LCA perspective. J. Clean. Prod. 142, 3544–3557.
Lu, Z.N., Chen, H., Hao, Y., Wang, J., Song, X., Mok, T.M., 2017b. The dynamic relationship between environmental pollution, economic development and public health: evidence from China. J. Clean. Prod. 166, 134–147.
Pan, Z., Wang, Y., Wu, C., Yin, L., 2017. Oil price volatility and macroeconomic fundamentals: a regime switching GARCH-MIDAS model. J. Empir. Financ. 43, 130–142.
Pettenuzzo, D., Timmermann, A., Valkanov, R., 2016. A MIDAS approach to modelling first and second moment dynamics. J. Econom. 193, 315–334.
Qian, X., Zhang, S.H., Zheng, L., Hua, H., 2012. Study on the econometric model of industrial wastewater pollutants and economic growth in Hubei Province. Water & Wastewater Engineering 38, 51–55.
Qin, C.B., Su, B., Bressers, H., Jia, Y.W., Wang, H., 2013. Assessing the economic impact of North China's water scarcity mitigation strategy: a multi-region, water-extended computable general equilibrium analysis. Water Int. 38, 701–723.
Somorowska, U., Łaszewski, M., 2019. Quantifying streamflow response to climate variability, wastewater inflow, and sprawling urbanization in a heavily modified river basin. Sci. Total Environ. 656, 458–467.
Tsui, A.K., Xu, C.Y., Zhang, Z., 2018. Macroeconomic forecasting with mixed data sampling frequencies: evidence from a small open economy. J. Forecast. 37, 666–675.
Turhan, I.M., Sensoy, A., Hacihasanoglu, E., 2015. Shaping the manufacturing industry performance: MIDAS approach. Chaos, Solit. Fractals 77, 286–290.
Wang, Q., Yang, Z., 2016. Industrial water pollution, water environment treatment, and health risks in China. Environ. Pollut. 218, 358–365.
Wang, Z.C., Yu, H.F., 2012. Research on predictive model of water pollution emissions based on BP neural network. Journal of Qinghai Normal University (Natural Science Edition) 4, 13–16.
Wang, Z.X., Hipel, K.W., Wang, Q., He, S.W., 2011. An optimized NGBM (1, 1) model for forecasting the qualified discharge rate of industrial wastewater in China. Appl. Math. Model. 35, 5524–5532.
Wang, Z.X., Li, Q., Pei, L.L., 2017. Grey forecasting method of quarterly hydropower production in China based on a data grouping approach. Appl. Math. Model. 51, 302–316.
Wang, Z.X., Li, Q., Pei, L.L., 2018. A seasonal GM (1,1) model for forecasting the electricity consumption of the primary economic sectors. Energy 154, 522–534.
Yan, K., Liu, C., 2015. Estimation on GDP growth in China's 'New Normal' period. Finance & Trade Economics 137, 701–710.
Yang, H., Zhou, Y., Abbaspour, K.C., 2010. An analysis of economic growth and industrial wastewater pollution relations in China. Consilience-The Journal of Sustainable Development 4, 60–79.
Yao, X., Feng, W., Zhang, X., Wang, W., Zhang, C., You, S., 2018. Measurement and decomposition of industrial green total factor water efficiency in China. J. Clean. Prod. 198, 1144–1156.
Zeng, B., Duan, H.M., Zhou, Y.F., 2019. A new multivariable grey prediction model with structure compatibility. Appl. Math. Model. 75, 385–397.
Zhang, J., Deng, S.H., Zhang, Y.Z., Tang, Q., 2011. A new model concerning the relationship between industrial wastewater generation, abatement rate, discharge and economy in China. Procedia Environmental Sciences 11, 803–809.
Zhang, J.L., Zhang, Y.J., Zhang, L., 2015. A novel hybrid method for crude oil price forecasting. Energy Econ. 49, 649–659.
Zhang, Y., Zhao, T., Zhou, A., Zhang, Z., Liu, W., 2016. Scenario based municipal wastewater estimation: development and application of a dynamic simulation model. Model. Simul. Eng. 1–13, 2016.
Zhang, C., Wang, Y., Song, X., Kubota, J., He, Y., Tojo, J., Zhu, X., 2017. An integrated specification for the nexus of water pollution and economic growth in China: panel cointegration, long-run causality and environmental Kuznets curve. Sci. Total Environ. 609, 319–328.
Zhao, X., Han, M., Ding, L., Calin, A.C., 2018a. Forecasting carbon dioxide emissions based on a hybrid of mixed data sampling regression model and back propagation neural network in the USA. Environ. Sci. Pollut. Control Ser. 25, 2899–2910.
Zhao, X., Han, M., Ding, L., Kang, W., 2018b. Usefulness of economic and energy data at different frequencies for carbon price forecasting in the EU ETS. Appl. Energy 216, 132–141.
Zhou, X., Sun, H., 2013. The industrial wastewater emissions and economic growth in regional differentiation in China. J. Arid Land Resour. Environ. 27, 15–19.
Zhu, Y., Price, O.R., Kilgallon, J., Qi, Y., Tao, S., Jones, K.C., Sweetman, A.J., 2018. Drivers of contaminant levels in surface water of China during 2000–2030: relative importance for illustrative home and personal care product chemicals. Environ. Int. 115, 161–169.