Journal of Air Transport Management 112 (2023) 102439

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

Air passenger flow forecasting using nonadditive forecast combination with grey prediction

Yi-Chung Hu

Department of Business Administration, Chung Yuan Christian University, Taoyuan City, Taiwan

ARTICLE INFO

Keywords:
Passenger demand
Air transport
Forecast combination
Grey system
Fuzzy integral

ABSTRACT

Air passenger flow forecasting plays a critical role in managing air transportation. It is important to develop accurate methods for forecasting the flow of passengers to formulate relevant operational plans and the schedule of airports. Even though empirical evidence has shown that forecast combination can improve the prediction accuracy of single-model forecasts, no study in the literature on transportation to date has examined the performance of such combinations to predict the demand for air transport. In light of the usefulness of grey prediction, which does not require applying any statistical test to examine the data at hand, and relationships among forecasts to be reckoned with, this study proposes using the nonadditive Choquet fuzzy integral to nonlinearly synthesize forecasts from four commonly used univariate grey prediction models into composite forecasts. The results of experiments confirmed that the risk arising from the incorrect choice of forecasting models can be eliminated by using forecast combinations. The proposed combination method based on grey prediction significantly outperformed the other competitors considered.

1. Introduction

Due to the importance of accurate forecasts of passenger flow for both transportation policymakers and business practitioners, a large body of research has been dedicated to developing forecasting methods in the transportation literature. Forecasting the demand for air transport is also attracting increasing attention from the transportation industry because it can help air transport practitioners manage the system of transport through the implementation of appropriate operational plans and related strategies (Jin et al., 2020).

Passenger demand forecasting methods can be grouped into three categories: time series models, causal econometric approaches, and Artificial Intelligence (AI)-based methods (Dantas et al., 2017). Time series models relate past observations to the current forecasted variable. Commonly used models of this kind include the autoregressive integrated moving average (ARIMA) (Andreana et al., 2021; Anvari et al., 2016; Chen et al., 2009; Gong et al., 2014; Tsui et al., 2014) and smoothing techniques (Samagaio and Wolters, 2010). Econometric approaches relate independent variables with the demand for passengers. A number of such approaches have been used to forecast the demand for passengers, such as the logit model (Lu et al., 2020), autoregressive distributed lag models, time-varying parameter, and vector autoregressive model (Fildes et al., 2011; Sun et al., 2019). Both time series models and econometric approaches usually assume a linear relationship among the relevant variables. Furthermore, to deliver adequate performance, such statistics as the stationarity test are required by these methods to examine the available data (Liou and Tzeng, 2012; Tzeng and Shen, 2017). However, the requirement for agreement between the data and certain statistical properties may not be realistic for empirical data (Jiang and Hu, 2022; Liou, 2013; Peng and Tzeng, 2013; Tzeng and Shen, 2017). Moreover, the presumption of a linear relationship between independent variables and the dependent variable appears to deviate from reality (Sun et al., 2019).

Because time series models and econometric approaches have a limited ability to capture the nonlinear characteristics of the demand for passengers (Saayman and Botha, 2017; Jin et al., 2020), many studies have instead used AI-based methods, which are adaptable and nonlinear, and thus can approximate to arbitrary functions (Sun et al., 2019). Recent examples of forecasting the demand for transport include the use of neural networks (Li et al., 2019; Sun et al., 2019; Wang et al., 2021; Yang et al., 2019; Yao et al., 2018), support vector regression (Shi et al., 2020; Tang et al., 2018), deep learning (He et al., 2022; Liu et al., 2019; Yang et al., 2021; Zhang et al., 2021), **and** spatiotemporal modeling (Wei et al., 2022). Some studies have extracted meaningful patterns from the series of data and applied AI-based methods to integrate them and generate forecasts (Jin et al., 2020; Ma et al., 2014; Qin

E-mail address: ychu@cycu.edu.tw.

https://doi.org/10.1016/j.jairtraman.2023.102439
Received 19 January 2023; Received in revised form 30 April 2023; Accepted 30 May 2023
Available online 12 June 2023

Table 1
Descriptive statistics on air passenger flow of six major airports in Taiwan.

Airport	Maximum	Minimum	Mean	Median	Std. Dev.	Skewness	Kurtosis
Taoyuan	4312472	1883987	3149662	3220985.5	690796.10	−0.202	−1.243
Kaohsiung	688484	286215	480908.4	487381	103843.15	−0.056	−1.042
Songshan	629394	354286	494600	485037	61922.58	0.147	−0.457
Taichung	268484	86988	181610.3	184128	44201.78	−0.135	−0.787
Kinmen	249752	137456	193288.5	191536.5	22856.10	0.145	−0.220
Penghu	326637	99034	189283.3	177471.5	64152.32	0.379	−1.118

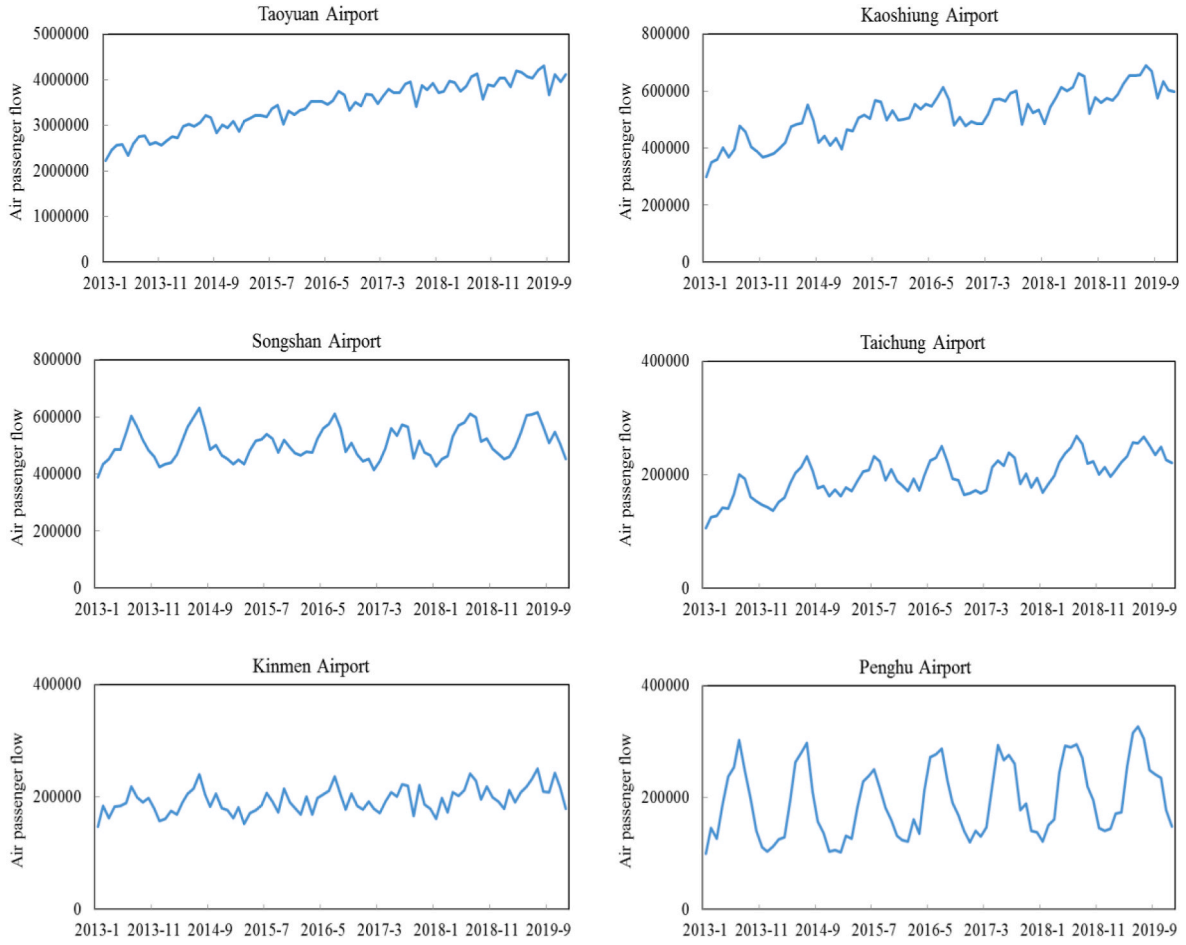

Fig. 1. Time series plot of passenger flows for Taiwan's primary airports.

et al., 2019; Wei and Chen, 2012).

Despite of the effectiveness of AI-based methods on passenger demand forecasting, a large number of observations are required by them to provide reasonably accurate forecasts (Shahrabi et al., 2013). In the context of machine learning, some empirical evidence has shown that well-behaved learning curves with respect to the accuracy and the size of the training set often have the shape of a power law (Viering and Loog, 2021). By contrast, grey prediction, an AI-based method, can capture the structure of nonlinear relationships based on a small sample size without requiring that the data satisfy any statistical property (Liu et al., 2017). In light of this advantage, grey prediction has been widely applied to many fields, such as for predicting the demand for energy (Hu and Wang, 2022; Wu et al., 2019), vehicle emissions (Xie et al., 2020), and traffic flow (Xiao et al., 2017). However, little attention has been paid to using grey prediction models to forecast the demand for passengers in transportation, with some exceptions such as the work by Chen et al. (2019), Hsu and Wen (1998), and Li and Han (2022).

Point forecasting dominates the literature on forecasting the demand for passengers. No single model has been found to be superior to all others in all situations (Li et al., 2019; Song et al., 2019; Sun et al., 2021). Relying on a single forecasting model may incur the risk of the failure of forecasting because it may generate the best forecasts in one time period but fail in other periods (Sun et al., 2021). Combination forecasting is a feasible approach for reducing such risk by synthesizing forecasts generated by different models based on weighting schemes (Qiu et al., 2021; Sun et al., 2021; Wu et al., 2017). Empirical evidence has shown that combining forecasting can improve accuracy of point forecasts of the demand for tourism (Cang, 2014; Hu and Wu, 2022; Shen et al., 2011; Song et al., 2019; Sun et al., 2021; Wu and Blake, 2022). The combination methods can be classified into linear and nonlinear methods (Cang, 2014). Linear combination methods, such as weighted average, assume that the sources of information are not preferentially independent of one another. However, because such an assumption ignores interactions among the inputs (Liou and Tzeng, 2012; Tzeng and Huang, 2011), it is reasonable to use nonlinear combination methods that consider preferential dependence among the

inputs while combining individual forecasts. The commonly used nonlinear combination methods such as support vector regression (SVR) (Cang, 2014), the radial-basis function network (RBFN) (Zheng et al., 2006), and multi-layer perceptron (MLP) (Cang, 2014; Wei and Chen, 2012), have shown to be more accurate than linear combinations. The nonlinear Choquet fuzzy integral, which considers interactions among the inputs as well, has been widely applied to many fields as well, such as alternative assessment (Liou et al., 2014; Liao et al., 2020) and classifier ensembles (Batista et al., 2022). Despite the usefulness of the Choquet integral for performance evaluation, little attention has been paid to its use in forecasting the passenger demand, with some exceptions such as tourism demand forecasting by Hu et al. (2021), and Hu and Wu (2022), among others. In summary, despite the advantages of combining forecasting models, little research has been devoted to this in the context of passenger demand forecasting, with some exceptions such as Guo et al. (2019), Ni et al. (2017), Tan et al. (2009) and Zhai et al. (2020).

To provide more accurate forecasts of air passenger demand for policymakers on transportation and businesses, this research aims to fill the aforementioned research gap by developing a forecast combination method for passenger demand forecasting by generating single-model forecasts from commonly used grey prediction models. The Choquet integral is then applied to nonlinearly synthesize individual forecasts into composite forecasts. To the best of our knowledge, no study in the literature on transportation to date has examined the performance of combination forecasting on air passenger demand.

The remainder of this paper is structured as follows. Section 2 introduces the proposed nonlinear combined model with grey prediction (NCMG), including candidate models in a combination set and GM(1, *N*). Section 3 presents commonly used combination methods. Section 4 verifies forecasting accuracy of the proposed combination forecasting methods, and Section 5 summarizes the conclusions of this study.

2. Methodology

2.1. Data

The total number of passengers can be treated as a proxy for passenger demand (Sun et al., 2019), and the data used for the analysis are air passenger flow in Taiwan. In the case of Taiwan's airports, The Taoyuan, Kaohsiung, and Songshan airports are the main hubs for international passenger arrivals and departures, while the Kinmen and Penghu airports are important for transport to outlying islands of Taiwan. Taoyuan airport in particular is among the top 10 international airports worldwide in terms of international passenger flow and freight traffic. To investigate the forecasting performance of the proposed combination forecasting method, we used data on the total number of passengers going through each of six major civil aviation facilities in Taiwan: the Taoyuan, Kaohsiung, Songshan, Taichung, Kinmen, and Penghu airports. Monthly data were collected from the civil air transportation statistics reported by the Taiwan Civil Aeronautics Administration Tourism Bureau and spanned from January 2013 to December 2019.

The time series of air passenger flow usually exhibit characteristics of irregularity and volatility. Table 1 shows the descriptive statistics on the air passenger flow data. The skewness statistic indicates that all data series analyzed in this study are skewed. The distribution of passenger flow data from each airport is **platykurtic because** each series of data has negative kurtosis such that it has a flatter peak and thinner tails compared with a normal distribution. Therefore, the fluctuation of each data series is unstable. Fig. 1 shows the time series plot of air passenger flow at the six airports in Taiwan considered in this study. It is evident that there is a strong growth in flows of Taoyuan and Kaohsiung airports. The passenger flows at the Songshan and Kinmen airports were relatively stagnant, while that at the Penghu airport was highly volatile.

A time series of the demand is characterized by distinct seasonal patterns and volatility (Claveria and Torra, 2014). Therefore, seasonal adjustment is required to eliminate seasonality from the data (Claveria et al., 2017). Given an original sequence $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$, let s_i (i = 1, 2, …, 12) denote the seasonal index for the i-th month, where

$$s_i = \frac{\bar{s}_i}{\bar{s}} \tag{1}$$

$\bar{s}_i$ and $\bar{s}$ can be expressed as follows (Zhou et al., 2021):

$$\bar{s}_i = \frac{1}{\lfloor (n-i+12)/12 \rfloor}\left(x_i^{(0)} + x_{i+12}^{(0)} + x_{i+24}^{(0)} + \ldots + x_{n'}^{(0)}\right), n' \le n \tag{2}$$

$$\bar{s} = \frac{1}{n}\sum_{k=1}^{n} x_k^{(0)} \tag{3}$$

Seasonality is then removed from $x^{(0)}$ as follows:

$$x_k^{(0)} = \frac{x_k^{(0)}}{s_u} \tag{4}$$

where

$$u = \begin{cases} k \bmod 12, k \bmod 12 \neq 0 \\ 12, k \bmod 12 = 0 \end{cases} \tag{5}$$

2.2. Candidate models

Although GDP, exchange rate, and population are commonly used independent variables to estimate passenger demand, other variables can also influence air passenger flow (Sun et al., 2019). Because it is difficult to statistically determine the independent variables that influence demand (Cang, 2014), we do not use causal grey models that consider the relationship between the relevant factors and air passenger flow as constituents. We instead use univariate grey prediction models, which are frequently applied to tourism demand forecasting, including the first-order grey model with one variable (GM(1,1)), nonlinear grey Bernoulli model (NGBM(1,1)), fractional GM(1,1) (FGM(1,1)), and fractional NGBM(1,1) (FNGBM(1,1)) (He et al., 2022), to forecast air passenger flow. There are thus 11 ($C_2^4 + C_3^4 + C_4^4$) different model combinations.

The above-mentioned grey models have their own differential equations to characterize the dynamic behavior of an unknown system, and their forecasts can be expressed by the corresponding time response equations (Liu et al., 2017). The relevant parameters in a time response equation are usually estimated by using a set of discrete equations. Xie and Liu (2009) have noted that this is likely to lead to forecasting errors because the time response and discrete equations could not equal accurately. This motivated us to use the discrete grey models as individual constituents. Individual single-model forecasts obtained from the candidate models are synthesized by the Choquet integral into a combined forecast.

2.2.1. GM(1,1)

Given a series ($x^{(0)}(1)$, $x^{(0)}(2)$, …, $x^{(0)}(n)$), the first-order accumulated generating operation (1-AGO) can be used to derive ($x^{(1)}(1)$,$x^{(1)}(2)$, …,$x^{(1)}(n)$) as (Wang, 2002)

$$x^{(1)}(k) = \sum_{j=1}^{k} x^{(0)}(j), k = 1, 2, \ldots, n \tag{6}$$

Xie and Liu (2009) expressed the relationship between $x^{(1)}(k-1)$ and $x^{(1)}(k)$ by the following discrete formulation:

$$x^{(1)}(k) = \beta_1 x^{(1)}(k-1) + \beta_2 \tag{7}$$

where β_1 and β_2 can be estimated by ordinary least squares (OLS). The non-recursive function with respect to $\hat{x}^{(1)}(k)$ is written as follows:

$$\hat{x}^{(1)}(k)=\beta_1^{k-1}\left(x^{(0)}(1)-\frac{\beta_2}{1-\beta_1}\right)+\frac{\beta_2}{1-\beta_1} \tag{8}$$

The forecast of $\hat{x}^{(0)}(k)$ is then obtained by using the first-order inverse AGO (1-IAGO) as

$$\hat{x}^{(0)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1) \tag{9}$$

2.2.2. NGBM(1,1)

$x^{(1)}(k)$ in NGBM(1,1) is described by the following discrete formulation (Kong and Ma, 2018):

$$\left(x^{(1)}(k)\right)^{1-\eta}=\beta_1\left(x^{(1)}(k-1)\right)^{1-\eta}+\beta_2 \tag{10}$$

where η denotes a power exponent, and β_1 and β_2 can be derived by OLS. As a result, the non-recursive function can be written as

$$\hat{x}^{(1)}(k)=\left[\beta_1^{k-1}\left(x^{(0)}(1)\right)^{1-\eta}+\frac{1-\beta_1^{k-1}}{1-\beta_1}\beta_2\right]^{\frac{1}{1-\eta}} \tag{11}$$

and $\hat{x}_k^{(0)}$ can be obtained using 1-IAGO.

2.2.3. Fractional GM(1,1)

By using the p-order AGO ($0<p<1$), $(x^{(p)}(1), x^{(p)}(2), \ldots, x^{(p)}(n))$ can be derived for FGM(1,1) as (Hu, 2021; Liu et al., 2022)

$$x^{(p)}(k)=\sum_{i=1}^{k}\binom{k-i+p-1}{k-i}x^{(0)}(i), k=1,2,\ldots,n \tag{12}$$

in which

$$\binom{k-i+p-1}{k-i}=\frac{(k-i+p-1)(k-i+p-2)\ldots(p+1)p}{(k-i)!} \tag{13}$$

$x^{(p)}(k)$ is written as

$$x^{(p)}(k)=\beta_1 x^{(p)}(k-1)+\beta_2 \tag{14}$$

where β_1 and β_2 can be obtained by using OLS. The non-recursive function with respect to $\hat{x}^{(p)}(k)$ is then written as follows:

$$\hat{x}^{(p)}(k)=\beta_1^{k-1}\left(x^{(0)}(1)-\frac{\beta_2}{1-\beta_1}\right)+\frac{\beta_2}{1-\beta_1} \tag{15}$$

$\hat{x}^{(0)}(k)$ can be obtained by using the p-order IAGO:

$$\hat{x}^{(0)}(k)=\sum_{i=1}^{k}\binom{k-i-p-1}{k-i}\hat{x}^{(p)}(i) \tag{16}$$

2.2.4. Fractional NGBM(1,1)

The discrete form of FNGBM(1,1) is as follows:

$$(x^{(p)}(k))^{1-\eta}=\beta_1(x^{(p)}(k-1))^{1-\eta}+\beta_2 \tag{17}$$

where β_1 and β_2 can be derived by OLS. The non-recursive function with respect to $\hat{x}^{(p)}(k)$ is written as

$$\hat{x}^{(p)}(k)=\left[\beta_1^{k-1}\left(x^{(0)}(1)\right)^{1-\eta}+\frac{1-\beta_1^{k-1}}{1-\beta_1}\beta_2\right]^{\frac{1}{1-\eta}} \tag{18}$$

As in FGM(1,1), $\hat{x}^{(0)}(k)$ can be obtained using the p-order IAGO. Note that $\hat{x}^{(0)}(k)$ in each of the above discrete grey models is nonlinearly expressed in exponential form.

2.3. Forecasting performance evaluation

The periods of estimation and forecasting spanned from January 2013 to April 2016, and from May 2016 to December 2019, respectively. We ended up with the generation of 52 one- and three-months-ahead forecasts from September 2014 to December 2019. The latest 32 forecasts derived by the individual grey models were used for a comparison of their accuracy while the other forecasts were used to measure the relative weights of the individual models. The forecasting accuracy was evaluated by commonly used root mean square error (RMSE) and mean absolute percentage error (MAPE), which are defined as follows:

$$\text{RMSE}=\sqrt{\sum_{k=1}^{n}\frac{\left(\hat{x}_k^{(0)}-x_k^{(0)}\right)^2}{n}} \tag{19}$$

$$\text{MAPE}=\sum_{k=1}^{n}\frac{\left|\hat{x}_k^{(0)}-x_k^{(0)}\right|}{x_k^{(0)}} \tag{20}$$

where $\hat{x}_k^{(0)}$ is a forecast at time point k. Each single model was constructed by minimizing the MAPE.

To construct a single model, we used 20 observations from January 2013 to August 2014 and applied a rolling mechanism to determine the optimal length of the series to construct a single model. We use four-point rolling for one-month-ahead (1-MA) forecasting as an example. A forecast for September 2014 was produced by constructing a single model based on data from May 2014 to August 2014. Data from June 2014 to September 2014 were then used to construct a model to produce a forecast for October 2014. This process continued until a forecast for April 2016 had been generated. The average MAPE was computed over 20 forecasts from September 2014 to April 2016 for four-point rolling. In case v-point ($4\le v\le 20$) rolling has the minimum average MAPE among all considered rolls, v-point rolling is taken to yield all forecasts in the forecasting periods. Note that the construction of a grey prediction model requires at least four samples (Sun et al., 2016).

3. Combination forecasting

3.1. Linear combination methods

For each combination method, forecasts produced by candidate models were used as its inputs. Simple average (SA), variance-covariance (VACO), discounted mean-squared forecasting error (DMSFE), and quadratic programming (QP) which are commonly used in tourism demand forecasting are adopted to combine individual forecasts generated by the candidate models. Let $\hat{x}_{jk}$ denote the forecast generated by model j at time point k ($1=k\le n$). In the case of a z-model combination ($z=2, 3, 4$), a combined forecast, $\hat{e}_k$, generated at time point k by the SA can be expressed as follows:

$$\hat{e}_k=\sum_{j=1}^{z}\frac{\hat{x}_{jk}}{z} \tag{21}$$

The VACO generates $\hat{e}_k$ as follows:

$$\hat{e}_k=\sum_{j=1}^{z}\frac{\left[\sum_{t=1}^{n}\left(\hat{x}_{it}-x_t\right)^2\right]^{-1}}{\sum_{i=1}^{z}\left[\sum_{t=1}^{n}\left(\hat{x}_{it}-x_t\right)^2\right]^{-1}}\hat{x}_{jk} \tag{22}$$

$\hat{e}_k$ generated by the DMSFE takes the following form:

$$\hat{e}_k=\sum_{j=1}^{z}\frac{\left[\sum_{t=1}^{n}\beta^{n-t+1}\left(\hat{x}_{it}-x_t\right)^2\right]^{-1}}{\sum_{i=1}^{z}\left[\sum_{t=1}^{n}\beta^{n-t+1}\left(\hat{x}_{it}-x_t\right)^2\right]^{-1}}\hat{x}_{jk} \tag{23}$$

where β is a discounting factor ($0<\beta\le 1$) that assigns a greater weight to more recent data. For QP, the optimal weights are determined as (Chan et al., 2010)

Table 2
Forecasting accuracy of single-model forecasts.

Single model	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu	Average
1-MA (MAPE)							
GM(1,1)	0.1896	0.5195	0.8379	0.7574	0.6797	2.2405	0.8708
NGBM(1,1)	0.1840	0.4831	0.8427	0.8201	0.6803	2.6287	0.9398
FGM(1,1)	0.2247	0.5544	0.6981	0.8749	0.7622	2.6074	0.9536
FNGBM(1,1)	0.2035	0.5390	0.9523	0.9166	0.7890	2.6319	1.0054
Average	0.2005	0.5240	0.8328	0.8422	0.7278	2.5271	
3-MA (MAPE)							
GM(1,1)	0.2087	0.5349	0.8696	0.9709	0.8159	2.6738	1.0123
NGBM(1,1)	0.2016	0.5854	0.7937	0.9055	0.7671	2.4503	0.9506
FGM(1,1)	0.2447	0.6360	0.8471	0.9829	0.8850	2.4520	1.0080
FNGBM(1,1)	0.3112	0.8613	1.4157	1.7446	1.0017	4.8312	1.6943
Average	0.2416	0.6544	0.9815	1.1510	0.8674	3.1018	
1-MA (RMSE)							
GM(1,1)	0.0380	0.0854	0.1401	0.1228	0.1134	0.3705	0.1450
NGBM(1,1)	0.0375	0.0790	0.1417	0.1247	0.1102	0.4121	0.1509
FGM(1,1)	0.0400	0.0908	0.1223	0.1432	0.1233	0.4128	0.1554
FNGBM(1,1)	0.0428	0.0877	0.1573	0.1433	0.1290	0.4213	0.1636
Average	0.0396	0.0857	0.1404	0.1335	0.1190	0.4042	
3-MA (RMSE)							
GM(1,1)	0.0398	0.0908	0.1411	0.1457	0.1268	0.4176	0.1603
NGBM(1,1)	0.0391	0.0970	0.1322	0.1422	0.1189	0.3963	0.1543
FGM(1,1)	0.0406	0.1062	0.1423	0.1592	0.1384	0.4075	0.1657
FNGBM(1,1)	0.0617	0.1366	0.2230	0.2595	0.1530	0.7432	0.2628
Average	0.0453	0.1077	0.1597	0.1767	0.1343	0.4912	

Table 3
MAPE of 1-MA combined forecasts.

Combination method	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu	Average
Linear combination							
SA							
Two-model	0.1972	0.4723	0.8091	0.7567	0.7217	2.2172	0.8624
Three-model	0.1949	0.4597	0.8028	0.7356	0.7184	2.0963	0.8346
Four-model	0.1930	0.4553	0.7992	0.7269	0.7157	2.0587	0.8248
VACO							
Two-model	0.1953	0.4726	0.7992	0.7587	0.7233	2.2168	0.8610
Three-model	0.1927	0.4605	0.7918	0.7425	0.7204	2.0898	0.8330
Four-model	0.1918	0.4565	0.7872	0.7369	0.7177	2.0683	0.8264
DMSFE							
Two-model	0.1905	0.4837	0.7994	0.8035	0.7347	2.3190	0.8885
Three-model	0.1889	0.4642	0.7748	0.8122	0.7404	2.3331	0.8856
Four-model	0.1887	0.4517	0.7600	0.8091	0.7450	2.3523	0.8845
QP							
Two-model	0.1907	0.4852	0.7839	0.7672	0.7386	2.2068	0.8621
Three-model	0.1896	0.4704	0.7566	0.7393	0.7506	2.0336	0.8234
Four-model	0.1892	0.4750	0.7265	0.7145	0.7596	1.9516	0.8027
Nonlinear combination							
MLP							
Two-model	0.8349	0.7098	0.6318	1.1069	0.8554	2.2989	1.0730
Three-model	0.8369	0.9832	0.5818	1.2953	0.8854	2.1963	1.1298
Four-model	0.9201	1.4943	0.5339	1.2451	0.7991	3.1458	1.3564
RBFN							
Two-model	1.4734	1.9012	1.7504	2.9329	1.5734	6.6230	2.7091
Three-model	1.4432	1.8770	1.7277	2.8817	1.5322	6.3995	2.6436
Four-model	1.4074	1.8320	1.7041	2.8573	1.4597	6.5828	2.6406
SVR							
Two-model	0.9625	0.9070	0.6160	1.1044	0.8303	2.0673	1.0813
Three-model	0.8783	0.9209	0.5980	1.2961	0.8545	1.8463	1.0657
Four-model	1.0200	1.0200	1.0200	1.0200	1.0200	1.0200	1.0200
NCMG							
Two-model	0.1893	0.4689	0.7815	0.7531	0.6935	2.2121	0.8497
Three-model	0.1854	0.4579	0.7516	0.7219	0.7001	2.0351	0.8087
Four-model	0.1872	0.4511	0.7262	0.7026	0.7000	1.9526	0.7866

$$\text{Minimize} \sum_{t=1}^{n}\left(\sum_{i=1}^{z} w_i \widehat{x}_{it} - x_t\right)^2 \qquad \text{subject to} \sum_{i=1}^{z} w_i = 1 \text{ and } w_i \geq 0 \tag{24}$$

Table 4
MAPE of 3-MA combined forecasts.

Combination method	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu	Average
Linear combination							
SA							
Two-model	0.2154	0.6215	0.9589	1.0892	0.8527	3.0263	1.1273
Three-model	0.2052	0.6085	0.9508	1.0647	0.8495	2.9966	1.1126
Four-model	0.2008	0.6042	0.9468	1.0493	0.8446	2.9827	1.1047
VACO							
Two-model	0.2060	0.6010	0.9146	0.9885	0.8509	2.7245	1.0476
Three-model	0.2019	0.5860	0.8952	0.9696	0.8463	2.6439	1.0238
Four-model	0.1994	0.5806	0.8836	0.9569	0.8407	2.6106	1.0120
DMSFE							
Two-model	0.2101	0.5954	0.9212	1.2776	0.8693	3.4634	1.2228
Three-model	0.2036	0.5940	0.9035	1.3830	0.8705	3.7327	1.2812
Four-model	0.2039	0.5806	0.8920	1.4585	0.8704	3.9375	1.3238
QP							
Two-model	0.2084	0.5761	0.8311	0.9437	0.8334	2.5200	0.9855
Three-model	0.2042	0.5670	0.8284	0.9288	0.8292	2.4942	0.9753
Four-model	0.2043	0.5793	0.8323	0.9131	0.8199	2.4332	0.9637
Nonlinear combination							
MLP							
Two-model	0.7836	1.3413	0.6761	1.4701	0.8517	2.5710	1.2823
Three-model	0.7447	1.5200	0.7127	1.6008	0.8241	3.3748	1.4629
Four-model	0.7389	1.6719	0.8949	1.2715	0.7742	2.6212	1.3288
RBFN							
Two-model	1.4776	1.8855	1.7764	2.9594	1.5605	6.7346	2.7323
Three-model	1.4544	1.8548	1.7518	2.9487	1.4963	6.5454	2.6752
Four-model	1.4502	1.8475	1.6684	2.7879	1.4931	6.4388	2.6143
SVR							
Two-model	0.7576	1.1989	0.8329	1.3893	0.9903	2.7414	1.3184
Three-model	0.9665	1.2812	0.7831	1.2868	0.9907	2.7527	1.3435
Four-model	0.7481	2.0340	0.8421	1.2609	1.0298	2.5841	1.4165
NCMG							
Two-model	0.1983	0.5589	0.8324	0.9166	0.8006	2.5288	0.9726
Three-model	0.1893	0.5553	0.8094	0.9140	0.7821	2.4591	0.9515
Four-model	0.1802	0.5468	0.7931	0.8862	0.7696	2.4291	0.9342

3.2. Nonlinear combination methods

Linear combination methods presuppose preferential independence among the individual variables (Liou and Tzeng, 2012; Liou et al., 2014; Tzeng and Shen, 2017). This assumption is unrealistic because it deviates from reality. This calls for the use of nonlinear methods in combining the forecasts. Such methods, including the MLP, SVR, and RBFN, have attracted considerable interest because they can capture the intrinsically complex and nonlinear characteristics of passenger demand (Jin et al., 2020).

Forecasts generated by individual grey prediction models are used as the inputs to the MLP, SVR, and RBFN. A neural network with one hidden layer is used because it can sufficiently approximate any continuous function (Goodfellow et al., 2016). The number of nodes in the hidden layer should be smaller than twice the number of inputs, as suggested by Heaton (2015). Python with the Scikit-Learn package was used to construct the MLP and the SVR for a given period of weight estimation, and the PyTorch package was used to set-up the RBFN. A grid search was performed to determine the appropriate hyper-parameters, including the learning rate of the MLP and the kernel function for SVR. For the sake of simplicity, introductions to the MLP, SVR, and RBFN are omitted here, but the interested reader can refer to work by Cang (2014) for them. The details of PyTorch and Scikit-Learn can refer to Raschka et al. (2022).

We applied the Choquet fuzzy integral, which is effective for dealing with preferential dependence among variables in multi-attribute decision making (Liao et al., 2020; Tzeng and Shen, 2017), to develop the NCMG. Let $\widehat{x}_k = (\widehat{x}_{k1}, \widehat{x}_{k2}, \widehat{x}_{kn})$ denote the sequence generated by a single model k $(1 = k \leq z)$ in a z-model combination. Let f be a measurable function defined on a feature space $X = \{x_1, x_2, \ldots, x_z\}$ such that $f(x_k) = \widehat{x}_{kt}$ at time point t. The elements of X are renumbered such that $f(x_z) \leq f(x_{z\text{-}1}) \leq \ldots \leq f(x_1)$. Along with the λ-fuzzy measure μ, where μ: $P(X) \rightarrow [0, 1]$ and $P(X)$ denotes the power set of X, a combined forecast $(\widehat{e}_k)$ generated by the Choquet integral is expressed as follows (Hu, 2010a; Hu and Chen, 2010; Kuncheva, 2000):

$$\widehat{e}_t = \sum_{i=1}^{z} f(x_i)[\mu(E_i) - \mu(E_{i-1})] \tag{25}$$

where $\mu(E_0) = 0$ and $\mu(E_s)$ is

$$\mu(E_s) = \frac{1}{\lambda}\left[\prod_{i=1..s}(1 + \lambda\mu_i) - 1\right] \tag{26}$$

where $E_s = \{x_1, x_2, \ldots, x_s\}$ $(1 = s \leq z)$, and $\mu_i = \mu(\{x_i\})$ is called the fuzzy density $(0 \leq \mu_i \leq 1)$. λ can be uniquely determined from the condition $\mu(X) = 1$ in case all fuzzy densities are known. Given R, $S \in P(X)$, where $R \cap S = \varnothing$, λ can indicate the interaction between R and S. That is, there is a multiplicative effect between R and S in case of $\lambda > 0$ and a substitutive effect between them in case of $\lambda < 0$. $\lambda = 0$ means that R and S are noninteractive. The nonadditivity of μ means that $\mu(E_z)$ may be smaller or greater than $\mu_1 + \mu_2 + \ldots + \mu_z$, and expresses interactions among the elements in E_z.

With the aim of minimizing the RMSE, the genetic algorithm toolbox in **MATLAB** was used to determine optimized values of the fuzzy densities for combining z individual forecasts. Values of 100, 500, 0.8, and 0.01 were assigned to the population, total number of generations, and the rates of crossover and mutation, respectively. The average accuracy of each nonlinear combination method was obtained over 10 independent trials. Genetic algorithm implementation and its use of **MATLAB** can refer to works by Hu (2010b) and Sivanandam and Deepa (2008), respectively.

Table 5
RMSE of 1-MA combined forecasts.

Combination method	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu	Average
Linear combination							
SA							
Two-model	0.0386	0.0788	0.1343	0.1228	0.1180	0.3601	0.1421
Three-model	0.0383	0.0764	0.1323	0.1190	0.1176	0.3454	0.1382
Four-model	0.0381	0.0752	0.1314	0.1172	0.1175	0.3373	0.1361
VACO							
Two-model	0.0386	0.0788	0.1331	0.1224	0.1181	0.3592	0.1417
Three-model	0.0383	0.0767	0.1303	0.1196	0.1178	0.3443	0.1378
Four-model	0.0382	0.0757	0.1288	0.1185	0.1177	0.3365	0.1359
DMSFE							
Two-model	0.0385	0.0798	0.1344	0.1293	0.1202	0.3768	0.1465
Three-model	0.0380	0.0762	0.1294	0.1319	0.1212	0.3743	0.1452
Four-model	0.0380	0.0741	0.1257	0.1336	0.1220	0.3746	0.1447
QP							
Two-model	0.0381	0.0803	0.1312	0.1208	0.1199	0.3575	0.1413
Three-model	0.0376	0.0788	0.1260	0.1162	0.1214	0.3374	0.1362
Four-model	0.0375	0.0804	0.1209	0.1125	0.1228	0.3335	0.1346
Nonlinear combination							
MLP							
Two-model	0.1388	0.1155	0.1083	0.1786	0.1321	0.3743	0.1746
Three-model	0.1394	0.1579	0.1030	0.2230	0.1376	0.3647	0.1876
Four-model	0.1532	0.2446	0.0963	0.2039	0.1268	0.4673	0.2154
RBFN							
Two-model	0.2279	0.2681	0.2612	0.3878	0.2234	0.9172	0.3809
Three-model	0.2234	0.2650	0.2584	0.3817	0.2180	0.8926	0.3732
Four-model	0.2181	0.2591	0.2567	0.3790	0.2094	0.9130	0.3726
SVR							
Two-model	0.1548	0.1443	0.1078	0.1730	0.1299	0.3405	0.1751
Three-model	0.1423	0.1462	0.1052	0.1975	0.1321	0.3085	0.1720
Four-model	0.1638	0.1444	0.1034	0.1984	0.1262	0.3126	0.1748
NCMG							
Two-model	0.0380	0.0787	0.1309	0.1202	0.1141	0.3532	0.1392
Three-model	0.0376	0.0762	0.1256	0.1155	0.1150	0.3291	0.1332
Four-model	0.0374	0.0752	0.1203	0.1128	0.1148	0.3249	0.1309

4. Empirical analysis

The average results of each single grey model in Table 2 show that no single model outperformed all others in all situations involving all six airports and on both forecasting horizons. For example, in terms of the MAPE for 1-MA forecasting, GM(1,1) yielded the best forecast (0.6797) for Kingmen airport but NGBM(1,1) delivered the best performance (0.1840) for Taoyuan airport. Table 2 confirms that any single model was unlikely to yield the best forecasts in different time periods, where this has been noted in some studies, such as Hu and Wu (2022), Li et al. (2019), Shen et al. (2011), and Sun et al. (2021).

To show the forecasting accuracy of different combination methods, the average MAPEs are summarized in Tables 3 and 4, and the average RMSEs are summarized in Tables 5 and 6. The results from Tables 3–6 show that, the average accuracy of three- and four-model combinations of a majority of combination methods including, the SA, VACO, QP, RBFN, and NCMG, were better than those of two- and three-model combinations, respectively, for all accuracy measures and over all forecasting horizons. The forecasting accuracy of the other methods of combining single-model forecasts did not exhibit this characteristic. For example, in the case of the MLP, Table 4 shows that the three-model combination (1.4629) was inferior to both the four-model (1.3288) and the two-model (1.2823) combinations. These findings appear to be inconsistent with the results reported by Song et al. (2009), whereby the forecasting accuracy of combined models increases with the number of single models used.

For each accuracy measure, the results in Tables 3–6 show that the average accuracy of the proposed NCMG with either model combination was superior to those of single-model forecasts for all airports and over all forecasting horizons. Likewise, Table 7 shows that the worst forecasts of the NCMG with either model combination were superior to the worst single-model forecasts for all airports, accuracy measures, and forecasting horizons as well. One may sometimes encounter the problem of model selection when the final single model selected is not necessarily the best for future use, whereas the results here indicate that the NCMG is a suitable choice for decision-makers, especially those involved in air transport.

The Friedman test was used to statistically examine whether there was any difference in forecasting accuracy between the NCMG and the other nonlinear methods considered. The Friedman test checks whether the measured average ranks are significantly different from the mean rank (2.5). The methods of combining forecasts were ranked separately for each airport, with the best-performing method ranked as “1,” the second-best ranked as “2,” and so on. For both accuracy measures, the “Friedman statistic” column in Table 8 shows that the null hypothesis was rejected for all model combinations. The post-hoc Nemenyi test was then conducted to detect differences between the matched pairs by comparing differences in average rank between them with the critical difference. The “Matched comparison” columns in Table 8 show intriguing matched pairs considered between the NCMG and nonlinear combination methods. The results in the “Nemenyi test” column show that the NCMG significantly outperformed the RBFN across all model combinations and accuracy measures. Even though the NCMG was not significantly superior to the MLP and SVR, it had the lowest average rank among the nonlinear methods of combining forecasts considered.

On both accuracy measures, the “Friedman statistic” column in Table 9 shows that the average ranks of the NCMG and the other linear competitive methods were significantly different from the mean rank (3). The “Nemenyi test” column shows that the NCMG significantly outperformed the VACO, DMSFE, and SA across all model combinations. Even though the NCMG was not significantly superior to the QP in terms of the RMSE across all model combinations, its average rank was lower than that of the QP for each model combination. That is, the NCMG was superior to the QP in terms of rankings of forecasting performance.

Table 6
RMSE of 3-MA combined forecasts.

Combination method	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu	Average
Linear combination							
SA							
Two-model	0.0417	0.1022	0.1567	0.1704	0.1318	0.4801	0.1805
Three-model	0.0403	0.1001	0.1557	0.1679	0.1310	0.4755	0.1784
Four-model	0.0396	0.0990	0.1552	0.1666	0.1306	0.4730	0.1773
VACO							
Two-model	0.0393	0.0996	0.1505	0.1565	0.1315	0.4374	0.1691
Three-model	0.0384	0.0976	0.1482	0.1538	0.1306	0.4286	0.1662
Four-model	0.0383	0.0966	0.1471	0.1520	0.1301	0.4249	0.1648
DMSFE							
Two-model	0.0394	0.0996	0.1514	0.1950	0.1344	0.5445	0.1941
Three-model	0.0387	0.0998	0.1489	0.2099	0.1345	0.5831	0.2025
Four-model	0.0386	0.0983	0.1477	0.2216	0.1346	0.6116	0.2087
QP							
Two-model	0.0393	0.0966	0.1384	0.1497	0.1291	0.4082	0.1602
Three-model	0.0382	0.0957	0.1389	0.1501	0.1285	0.4078	0.1599
Four-model	0.0379	0.0974	0.1400	0.1493	0.1279	0.4039	0.1594
Nonlinear combination							
MLP							
Two-model	0.1314	0.2160	0.1075	0.2163	0.1320	0.4220	0.2042
Three-model	0.1250	0.2384	0.1127	0.2463	0.1291	0.5262	0.2296
Four-model	0.1248	0.2497	0.1414	0.2279	0.1278	0.4183	0.2150
RBFN							
Two-model	0.2286	0.2661	0.2646	0.3911	0.2216	0.9298	0.3836
Three-model	0.2251	0.2624	0.2616	0.3896	0.2138	0.9083	0.3768
Four-model	0.2245	0.2614	0.2519	0.3719	0.2133	0.8979	0.3702
SVR							
Two-model	0.1309	0.1913	0.1367	0.2171	0.1490	0.4389	0.2107
Three-model	0.1557	0.1989	0.1266	0.1941	0.1491	0.4447	0.2115
Four-model	0.1219	0.2899	0.1419	0.1923	0.1542	0.4133	0.2189
NCMG							
Two-model	0.0384	0.0934	0.1382	0.1444	0.1242	0.3993	0.1563
Three-model	0.0370	0.0934	0.1347	0.1413	0.1222	0.3953	0.1540
Four-model	0.0364	0.0940	0.1343	0.1370	0.1200	0.3945	0.1527

Table 7
Worst single-model forecasts and worst combined forecasts of NCMG.

1-MA (MAPE)	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu
Single-model forecasts	0.2247	0.5544	0.9523	0.9166	0.7890	2.6319
NCMG						
Two-model	0.1981	0.4906	0.8546	0.8073	0.7274	2.6009
Three-model	0.1937	0.4698	0.8491	0.7679	0.7062	2.1541
Four-model	0.1872	0.4511	0.7262	0.7026	0.7000	1.9526
3-MA (MAPE)	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu
Single-model forecasts	0.3112	0.8613	1.4157	1.7446	1.0017	4.8312
NCMG						
Two-model	0.2134	0.5956	0.8781	0.9743	0.8702	2.7370
Three-model	0.2095	0.5675	0.8286	0.9327	0.8066	2.5618
Four-model	0.1802	0.5468	0.7931	0.8862	0.7696	2.4291
1-MA (RMSE)	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu
Single-model forecasts	0.0428	0.0908	0.1573	0.1433	0.1290	0.4213
NCMG						
Two-model	0.0392	0.0825	0.1427	0.1262	0.1202	0.4167
Three-model	0.0383	0.0798	0.1413	0.1227	0.1170	0.3397
Four-model	0.0374	0.0752	0.1203	0.1128	0.1148	0.3249
3-MA (RMSE)	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu
Single-model forecasts	0.0617	0.1366	0.2230	0.2595	0.1530	0.7432
NCMG						
Two-model	0.0406	0.0980	0.1451	0.1522	0.1333	0.4202
Three-model	0.0391	0.0952	0.1364	0.1438	0.1265	0.4081
Four-model	0.0364	0.0940	0.1343	0.1370	0.1200	0.3945

Overall, Tables 8 and 9 verify the superior forecasting accuracy of the proposed combination method compared with the linear and nonlinear methods considered here.

The Diebold–Mariano test, which uses the square prediction error as the loss function, was used to investigate whether the forecasting accuracy of the proposed NCMG was significantly superior to the other competitive combination methods for each airport as well. The results in Tables 10 and 11 are positive for most cases, suggesting that the NCMG was more accurate than the other combination methods. The ratios of the statistically positive results were 57.14% and 78.57% in cases of one-

Table 8
Results of the statistical test for NCMG and nonlinear combination methods.

Combination	Average rank				Friedman statistic	Matched-pair comparison	Nemenyi test
	MLP	RBFN	SVR	NCMG			
Two-model(MAPE)	2.417	4.000	2.333	1.250	36.711***	MLP vs. NCMG	1.167
						RBFN vs. NCMG	2.750***
						SVR vs. NCMG	1.083
Three-model(MAPE)	2.333	4.000	2.250	1.417	26.009***	MLP vs. NCMG	0.917
						RBFN vs. NCMG	2.583***
						SVR vs. NCMG	0.833
Four-model(MAPE)	2.417	3.917	2.417	1.250	4.936***	MLP vs. NCMG	1.167
						RBFN vs. NCMG	2.667***
						SVR vs. NCMG	1.167
Two-model(RMSE)	2.333	4.000	2.250	1.417	26.009***	MLP vs. NCMG	0.917
						RBFN vs. NCMG	2.583***
						SVR vs. NCMG	0.833
Three-model(RMSE)	2.417	4.000	2.167	1.417	26.714***	MLP vs. NCMG	1.000
						RBFN vs. NCMG	2.583***
						SVR vs. NCMG	0.750
Four-model(RMSE)	2.500	3.833	2.417	1.250	22.277***	MLP vs. NCMG	1.250*
						RBFN vs. NCMG	2.583***
						SVR vs. NCMG	1.167

Note: *** in 'Friedman statistic' column denotes the statistical significance of Friedman statistic at a 1% significance level. ***, **, and * in 'Nemenyi test' column denote significant differences between the average ranks of two methods in 'Matched comparison' column at 1%, 5%, and 10% significance levels, respectively.

Table 9
Results of the statistical test for NCMG and linear combination methods.

Combination	Average rank					Friedman statistic	Matched-pair comparison	Nemenyi test
	VACO	DMSFE	QP	SA	NCMG			
Two-model(MAPE	3.083	4.167	2.583	3.917	1.250	13.146***	VACO vs. NCMG	1.833**
							DMSFE vs. NCMG	2.917**
							QP vs. NCMG	1.333
							SA vs. NCMG	2.667***
Three-model(MAPE)	3.167	4.083	2.750	3.917	1.083	15.053***	VACO vs. NCMG	2.083**
							DMSFE vs. NCMG	3.000***
							QP vs. NCMG	1.667*
							SA vs. NCMG	2.833***
Four-model(MAPE)	3.375	3.958	2.750	3.833	1.083	13.388***	VACO vs. NCMG	2.292**
							DMSFE vs. NCMG	2.875***
							QP vs. NCMG	1.667*
							SA vs. NCMG	2.750***
Two-model(RMSE)	3.083	4.375	2.458	4.083	1.000	30.739***	VACO vs. NCMG	2.033**
							DMSFE vs. NCMG	3.375***
							QP vs. NCMG	1.458
							SA vs. NCMG	3.083***
Three-model(RMSE)	3.375	4.083	2.458	4.042	1.042	20.680***	VACO vs. NCMG	2.333**
							DMSFE vs. NCMG	3.042***
							QP vs. NCMG	1.417
							SA vs. NCMG	3.000***
Four-model(RMSE)	3.333	4.000	2.583	3.958	1.125	14.693***	VACO vs. NCMG	2.208**
							DMSFE vs. NCMG	2.875***
							QP vs. NCMG	1.458
							SA vs. NCMG	2.833***

Note: *** in 'Friedman statistic' column denotes the statistical significance of Friedman statistic at a 1% significance level. ***, **, and * in "Nemenyi test" column denote significant differences between the average ranks of two methods in 'Matched comparison' column at 1%, 5%, and 10% significance levels, respectively.

step-ahead and three-step-ahead forecasts, respectively.

5. Concluding remarks

In light of the importance of accurate predictions of air passenger flow for the aviation industry and the government, this paper proposed a nonadditive combination forecasting method by using the Choquet fuzzy integral along with the λ-fuzzy measure to synthesize point forecasts from individual grey models into combined forecasts. Empirical results based on two accuracy measures showed that the proposed NCMG is superior to the other commonly used combination methods considered here for forecasting air passenger demand in Taiwan. The main reason for this is the ability of the fuzzy integral to capture interactions among individual forecasts in forecast combination.

In contrast to univariate time series models that perform well in modeling linear time series (Wei and Chen, 2012) including such as the ARIMA and exponential smoothing, the univariate grey models considered here do not require that the data follow any statistical property, and can nonlinearly approximate them. Jin et al. (2020) have also noted that traditional time series models cannot accurately capture the trend of air passenger demand. It is thus reasonable to use univariate grey models instead of traditional time series models as candidate models. The results here indicate that univariate grey models have a positive influence on the forecasting accuracy of the NCMG.

The results of the NCMG were superior to the worst single-model forecasts and its average accuracy was higher than those of the

Table 10
Results of Diebold-Mariano test statistic in 1-MA.

Combination	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu
Two-model						
SA	0.832	−0.250	2.390***	0.435	1.832**	0.784
VACO	0.948*	0.086	2.179**	0.976	1.850**	0.741
DMSFE	0.946	−2.173	0.864	2.308**	2.086**	1.352*
QP	−0.158	1.146	1.229	0.922	1.712**	0.726
MLP	6.374***	3.082***	−2.111	2.382***	1.169	−0.587
RBFN	14.698***	8.604***	5.253***	8.515***	4.787***	6.794***
SVR	8.127***	3.971***	−1.762	2.933***	1.296**	−0.178
Three-model						
SA	0.887	0.354	2.389***	1.328*	1.504*	1.371*
VACO	1.026	1.122	2.325***	1.767**	1.544*	1.548*
DMSFE	1.224	−1.614	1.672**	2.358***	2.012**	1.520*
QP	0.060	1.477*	0.704	0.386	1.732**	1.100
MLP	6.263***	4.269***	−2.468	2.096**	1.180	−0.538
RBFN	14.362***	8.554***	5.317***	8.325***	4.643***	6.726***
SVR	7.586***	4.084***	−2.020	3.486***	1.286**	−0.805
Four-model						
SA	1.304*	−0.043	1.950**	1.695**	1.717**	0.938
VACO	1.288	0.546	1.825**	1.632*	1.785**	0.972
DMSFE	0.878	−0.449	1.648**	2.358***	2.244**	1.732**
QP	0.229	2.146**	0.682	−0.261	1.993**	0.659
MLP	6.510***	4.714***	−2.602	2.363***	0.795	2.635**
RBFN	14.013***	8.379***	5.244***	8.281***	4.442***	6.644***
SVR	8.353***	4.084***	−1.953	3.725***	1.169	−0.233

Note: A negative sign of the statistic implies that NCMG has greater forecasting errors. ***, **, and * denote significance at 1%, 5%, and 10% significance levels, respectively.

Table 11
Results of Diebold-Mariano test statistic in 3-MA.

Combination	Taoyuan	Kaohsiung	Songshan	Taichung	Kingmen	Penghu
Two-model						
SA	1.498*	3.878***	4.849***	4.674***	2.442***	4.718***
VACO	0.771	3.481***	4.521***	4.246***	2.350***	3.706***
DMSFE	0.570	1.242	4.587***	5.043***	2.213**	5.040***
QP	1.353*	1.926**	2.972***	1.065	1.451*	0.906
MLP	6.034***	4.380***	−2.287	3.020***	0.533	−0.662
RBFN	14.699***	8.169***	6.240***	8.554***	4.983***	6.758***
SVR	4.748***	3.795***	0.740	3.131***	2.012**	1.231
Three-model						
SA	2.403***	3.901***	5.224***	4.246***	2.264**	4.804***
VACO	2.180**	4.250***	4.802***	2.943***	2.162**	4.100***
DMSFE	2.312***	1.613*	5.027***	5.174***	1.986**	5.155***
QP	1.548*	1.693**	0.864	1.006	1.329*	1.180
MLP	5.915***	4.812***	−2.425	0.867	0.417	0.565
RBFN	14.482***	8.114***	5.895***	8.355***	4.693***	6.708***
SVR	8.092***	4.551***	−1.492	2.828***	2.010**	1.761**
Four-model						
SA	2.371***	1.868**	5.556***	3.937***	2.497***	4.618***
VACO	2.089**	1.293*	5.438***	2.748***	2.404***	3.857***
DMSFE	2.330***	1.125	5.149***	5.243***	2.115**	5.196***
QP	1.093	1.191	1.422*	1.592*	1.506*	1.001
MLP	5.654***	5.429***	0.361	1.966**	0.466	0.479
RBFN	14.399***	7.938***	5.645***	7.649***	4.800***	6.615***
SVR	6.921***	7.265***	1.066	2.962***	2.476***	0.716

Note: A negative sign of the statistic implies that NCMG has greater forecasting errors. ***, **, and * denote significance at 1%, 5%, and 10% significance levels, respectively.

individual models considered. This finding is consistent with the claim that combining forecasts can reduce the risk of forecasting failure arising from the inappropriate choice of a single model (Li et al., 2019). Interestingly, some studies also indicate that it is less risky to apply forecast combination to tourism demand forecasting (Cang, 2014; Li et al., 2019; Wu et al., 2017; Sun et al., 2021). The NCMG is a useful combination forecasting method for aviation practitioners, and this study provides empirical support for the advantage of applying the NCMG in air passenger demand forecasting practice.

Undoubtedly, passenger flow forecasting is a crucial component of transportation systems. Transportation system management of the private and public sectors can benefit from a more accurate combination method because its forecasts reveal reliable needs for transport infrastructure, including the service quality and efficiency of airports, aviation equipment, operation planning, revenue planning, facility improvement, new route opening, and the need for talent cultivation. After all, accurately forecasting of passenger demand benefits the revenue management for the aviation industry (Pereira and Cerqueira, 2022; Yu, 2022).

Interestingly, the average accuracy listed in Tables 3–6 show that

linear combination methods outperformed nonlinear methods for each model combination. These results contradict the findings by Cang (2014), possibly because the forecasting performance of AI-based methods might have been sensitive to such hyper-parameters as the number of samples and the network architecture (Claveria and Torra, 2014). Besides, the findings of this study also highlight directions that ought to be pursued in future research. Increasing the diversity of candidate models is expected to benefit the performance of forecast combination (Song et al., 2009). Forecasting models in addition to grey prediction models—for example, ARIMA and neural networks—should thus be used in forecast combinations. Moreover, because multivariate forecasting models may yield better prediction than univariate models (Hu, 2020; Liu et al., 2023; Ma et al., 2019; Wu et al., 2019; Xie et al., 2020), it is possible to examine the effect of multivariate grey models on the improvement in accuracy of combination forecasting. In particular, based on the weakened accumulation grey optimization model and the multiobjective grey wolf optimizer, Liu et al. (2023) developed a new grey intelligent prediction algorithm with multiobjective correction strategy. Their methods addressed the uniformity, ill-condition and overfitting of the grey intelligent prediction algorithm. Although these issues are not the focus of this study, incorporation of multiobjective correction strategy into the NCMG remains to be future work.

Compliance with ethical standards

The author declares that he has no conflict of interest.

CRediT authorship contribution statement

Yi-Chung Hu: Conceptualization, Methodology, Validation, Data curation, Writing – original draft, Writing – review & editing.

Data availability

Data will be made available on request.

Acknowledgements

This research is supported by the Ministry of Science and Technology, Taiwan under grant MOST 110-2410-H-033-013-MY2. The author would like to thank the anonymous referees for their valuable comments.

References

Andreana, G., Gualini, A., Martini, G., Porta, F., Scotti, D., 2021. The disruptive impact of COVID-19 on air transportation: an ITS econometric analysis. Res. Transport. Econ. 90 https://doi.org/10.1016/j.retrec.2021.101042.
Anvari, S., Tuna, S., Canci, M., Turkay, M., 2016. Automated Box-Jenkins forecasting tool with an application for passenger demand in urban rail systems. J. Adv. Transport. 50 (1), 25–49.
Batista, T., Bedregal, B., Moraes, R., 2022. Constructing multi-layer classifier ensembles using the Choquet integral based on overlap and quasi-overlap functions. Neurocomputing 500, 413–421.
Cang, S., 2014. A comparative analysis of three types of tourism demand forecasting models: individual, linear combination and non-linear combination. Int. J. Tourism Res. 16, 596–607.
Chan, C.K., Witt, S.F., Lee, Y.C.E., Song, H., 2010. Tourism forecast combination using the CUSUM technique. Tourism Manag. 31, 891–897.
Chen, C.F., Chang, Y.H., Chang, Y.W., 2009. Seasonal ARIMA forecasting of inbound air travel arrivals to Taiwan. Transportmetrica 5 (2), 125–140.
Chen, Y.Y., Liu, H.T., Hsieh, H.L., 2019. Time series interval forecast using GM(1,1) and NGBM(1, 1) models. Soft Comput. 23, 1541–1555.
Claveria, O., Torra, S., 2014. Forecasting tourism demand to Catalonia: neural networks vs. time series models. Econ. Modell. 36, 220–228.
Claveria, O., Monte, E., Torra, S., 2017. Data pre-processing for neural network-based forecasting: does it really matter? Technol. Econ. Dev. Econ. 23 (5), 709–725.
Dantas, T.M., Cyrino Oliveira, F.L., Varela Repolho, H.M., 2017. Air transportation demand forecast through Bagging Holt Winters methods. J. Air Transport. Manag. 59, 116–123.
Fildes, R., Wei, Y., Ismail, S., 2011. Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures. Int. J. Forecast. 27 (3), 902–922.
Gong, M., Fei, X., Wang, Z.H., Qiu, Y.J., 2014. Sequential framework for short-term passenger flow prediction at bus stop. Transport. Res. Rec. 2417, 58–66.
Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press, Cambridge, MA.
Guo, J., Xie, Z., Qin, Y., Jia, L., Wang, Y., 2019. Short-Term abnormal passenger flow prediction based on the fusion of SVR and LSTM. IEEE Access 7, 42946–42955.
Guo, Z., Zhao, X., Chen, Y., Wu, W., Yang, J., 2019. Short-term passenger flow forecast of urban rail transit based on GPR and KRR. IET Intell. Transp. Syst. 13 (9), 1374–1382.
He, Y., Zhao, Y., Luo, Q., Tsui, K.-L., 2022. Forecasting nationwide passenger flows at city-level via a spatiotemporal deep learning approach. Phys. Stat. Mech. Appl. https://doi.org/10.1016/j.physa.2021.126603.
Heaton, J., 2015. Artificial Intelligence for Humans: Deep Learning and Neural Networks. Heaton Research, St. Louis, MO.
Hsu, C.I., Wen, Y.H., 1998. Improved grey prediction models for the trans-pacific air passenger market. Transport. Plann. Technol. 22 (2), 87–107.
Hu, Y.C., 2010a. Pattern classification by multi-layer perceptron using fuzzy integral-based activation function. Appl. Soft Comput. 10 (3), 813–819.
Hu, Y.C., 2010b. Analytic network process for pattern classification problems using genetic algorithms. Inf. Sci. 180 (13), 2528–2539.
Hu, Y.C., 2020. A multivariate grey prediction model with grey relational analysis for bankruptcy prediction problems. Soft Comput. 24, 4259–4268.
Hu, Y.C., 2021. Forecasting tourism demand using fractional grey prediction models with Fourier series. Ann. Oper. Res. 300, 467–491.
Hu, Y.C., Chen, H.C., 2010. Choquet integral-based hierarchical networks for evaluating customer service perceptions on fast food stores. Expert Syst. Appl. 37 (12), 7880–7887.
Hu, Y.C., Wang, W.B., 2022. Nonlinear interval regression analysis with neural networks and grey prediction for energy demand forecasting. Soft Comput. 26 (14), 6529–6545.
Hu, Y.C., Wu, G., 2022. The impact of Google Trends index and encompassing tests on forecast combinations in tourism. Tourism Rev. 77 (5), 1276–1298.
Hu, Y.C., Wu, G., Jiang, P., 2021. Tourism demand forecasting using nonadditive forecast combinations. J. Hospit. Tourism Res. https://doi.org/10.1177/10963480211047857.
Jiang, P., Hu, Y.C., 2022. Constructing interval models using neural networks with non-additive combinations of grey prediction models in tourism demand. Grey Syst. Theor. Appl. https://doi.org/10.1108/GS-11-2021-0180.
Jin, F., Li, Y., Sun, S., Li, H., 2020. Forecasting air passenger demand with a new hybrid ensemble approach. J. Air Transport. Manag. 83 https://doi.org/10.1016/j.jairtraman.2019.101744.
Kong, L., Ma, X., 2018. Comparison study on the nonlinear parameter optimization of nonlinear grey Bernoulli model (NGBM(1,1)) between intelligent optimizers. Grey Syst. Theor. Appl. 8 (2), 210–226.
Kuncheva, L.I., 2000. Fuzzy Classifier Design. Physica-Verlag, Heidelberg.
Li, Y., Han, H., 2022. Forecast and analysis of passenger flow at Sanya Airport based on gray system theory. Commun. Comput. Inform. Sci. 1629, 427–434.
Li, H., Wang, Y., Xu, X., Qin, L., Zhang, H., 2019. Short-term passenger flow prediction under passenger flow control using a dynamic radial basis function network. Appl. Soft Comput. 83 https://doi.org/10.1016/j.asoc.2019.105620.
Li, G., Wu, C., Zhou, M., Liu, A., 2019. The combination of interval forecasts in tourism. Ann. Tourism Res. 75, 363–378.
Liao, Z., Liao, H., Tang, M., Al-Barakati, A., Llopis-Albert, C., 2020. A Choquet integral-based hesitant fuzzy gained and lost dominance score method for multi-criteria group decision making considering the risk preferences of experts: case study of higher business education evaluation. Inf. Fusion 62, 121–133.
Liou, J.J.H., 2013. New concepts and trends of MCDM for tomorrow – in honor of Professor Gwo-Hshiung Tzeng on the occasion of his 70th birthday. Technol. Econ. Dev. Econ. 19, 367–375.
Liou, J.J.H., Tzeng, G.H., 2012. Comments on "Multiple criteria decision making (MCDM) methods in economics: an overview". Technol. Econ. Dev. Econ. 18, 672–695.
Liou, J.J.H., Chuang, Y.C., Tzeng, G.H., 2014. A fuzzy integral-based model for supplier evaluation and improvement. Inf. Sci. 266, 199–217.
Liu, S., Yao, E., 2017. Holiday passenger flow forecasting based on the modified least-square support vector machine for the metro system. J. Transport. Eng. Part A: Systems 143 (2). https://doi.org/10.1061/JTEPBS.0000010.
Liu, S., Yang, Y., Forrest, J., 2017. Grey Data Analysis: Methods, Models and Applications. Springer, Berlin.
Liu, Y., Liu, Z., Jia, R., 2019. DeepPF: a deep learning based architecture for metro passenger flow prediction. Transport. Res. Part C 101, 18–34.
Liu, L., Liu, S., Wu, L., Zhu, J., Shang, G., 2022. Forecasting the development trend of new energy vehicles in China by an optimized fractional discrete grey power model. J. Clean. Prod. https://doi.org/10.1016/j.jclepro.2022.133708.
Liu, C., Wu, W.Z., Xie, W., 2023. A new grey intelligent prediction algorithm with multiobjective correction strategy. Appl. Math. Model. 118, 692–708.
Lu, J., Lv, C., Feng, T., Wang, Z., 2020. Evaluating the effects of Hong Kong-zhuhai-Macao bridge on international air travel: demand analysis of the air-bridge-air path. Eur. J. Int. Manag. 14 (4), 595–616.
Ma, Z., Xing, J., Mesbah, M., Ferreira, L., 2014. Predicting short-term bus passenger demand using a pattern hybrid approach. Transport. Res. C Emerg. Technol. 39, 148–163.

Ma, X., Liu, Z., Wang, Y., 2019. Application of a novel nonlinear multivariate grey Bernoulli model to predict the tourist income of China. J. Comput. Appl. Math. 347, 84–94.
Ni, M., He, Q., Gao, J., 2017. Forecasting the subway passenger flow under event occurrences with social media. IEEE Trans. Intell. Transport. Syst. 18 (6), 1623–1632.
Peng, K.H., Tzeng, G.H., 2013. A hybrid dynamic MADM model for problems-improvement in economics and business. Technol. Econ. Dev. Econ. 19, 638–660.
Pereira, L.N., Cerqueira, V., 2022. Forecasting hotel demand for revenue management using machine learning regression methods. Curr. Issues Tourism 25 (17), 2733–2750.
Qin, L., Li, W., Li, S., 2019. Effective passenger flow forecasting using STL and ESN based on two improvement strategies. Neurocomputing 356, 244–256.
Qiu, R.T.R., Wu, D.C., Dropsy, V., Petit, S., Pratt, S., Ohe, Y., 2021. Visitor arrivals forecasts amid COVID-19: a perspective from the Asia and Pacific team. Ann. Tourism Res. 88 https://doi.org/10.1016/j.annals.2021.103155.
Raschka, S., Liu, Y., Mirjalili, V., Dzhulgakov, D., 2022. Machine Learning with PyTorch and Scikit-Learn: Develop Machine Learning and Deep Learning Models with Python. Packt Publishing, Birmingham, UK.
Saayman, A., Botha, I., 2017. Non-linear models for tourism demand forecasting. Tourism Econ. 23 (3), 594–613.
Samagaio, A., Wolters, M., 2010. Comparative analysis of government forecasts for the Lisbon Airport. J. Air Transport. Manag. 16 (4), 213–217.
Shahrabi, J., Hadavandi, E., Asadi, S., 2013. Developing a hybrid intelligent model for forecasting problems: case study of tourism demand time series. Knowl. Base Syst. 43, 112–122.
Shen, S.J., Li, G., Song, H., 2011. Combination forecasts of international tourism demand. Ann. Tourism Res. 38 (1), 72–89.
Shi, Z., Zhang, N., P, M., Zhang, J., 2020. Short-term metro passenger flow forecasting using ensemble-chaos support vector regression. Transportmetrica: Transport. Sci. 16 (2), 194–212.
Sivanandam, S., Deepa, S., 2008. Genetic algorithm implementation using matlab. In: Introduction to Genetic Algorithms. Springer, Heidelberg. https://doi.org/10.1007/978-3-540-73190-0_8.
Song, H., Witt, S.F., Wong, K.F., Wu, D.C., 2009. An empirical study of forecast combination in tourism. J. Hospit. Tourism Res. 33, 3–29.
Song, H., Qiu, R.T.R., Park, J., 2019. A review of research on tourism demand forecasting: launching the Annals of Tourism Research Curated Collection on tourism demand forecasting. Ann. Tourism Res. 75, 338–362.
Sun, X., Sun, W., Wang, J., Gao, Y., 2016. Using a Grey-Markov model optimized by Cuckoo search algorithm to forecast the annual foreign tourist arrivals to China. Tourism Manag. 52, 369–379.
Sun, S., Lu, H., Tsui, K.L., Wang, S., 2019. Nonlinear vector auto-regression neural network for forecasting air passenger flow. J. Air Transport. Manag. 78, 54–62.
Sun, Y., Zhang, J., Li, X., Wang, S., 2021. Forecasting tourism demand with a new time-varying forecast averaging approach. J. Trav. Res. https://doi.org/10.1177/00472875211061206.
Tan, M.C., Wong, S.C., Xu, J.M., Guan, Z.R., Zhang, P., 2009. An aggregation approach to short-term traffic flow prediction. IEEE Trans. Intell. Transport. Syst. 10 (1), 3613–3622.
Tang, L., Zhao, Y., Cabrera, J., Ma, J., Tsui, K.L., 2018. Forecasting short-term passenger flow: an empirical study on Shenzhen Metro. IEEE Trans. Intell. Transport. Syst. 20 (10), 3613–3622.
Tsui, W.H.K., Ozer Balli, H., Gilbey, A., Gow, H., 2014. Forecasting of Hong Kong airport's passenger throughput. Tourism Manag. 42, 62–76.
Tzeng, G.H., Huang, J.J., 2011. Multiple Attribute Decision Making: Methods and Applications. CRC Press, Boca Raton, FL.
Tzeng, G.H., Shen, K.Y., 2017. New Concepts and Trends of Hybrid Multiple Criteria Decision Making. CRC Press, NY.
Viering, T., Loog, M., 2021. The shape of learning curves: a review. Comput. Sci. https://doi.org/10.48550/arXiv.2103.10948.
Wang, Y.F., 2002. Predicting stock price using fuzzy grey prediction system. Expert Syst. Appl. 22 (1), 33–38.
Wang, J., Zhang, Y., Wei, Y., Hu, Y., Piao, X., Yin, B., 2021. Metro passenger flow prediction via dynamic hypergraph convolution networks. IEEE Trans. Intell. Transport. Syst. 22 (12), 7891–7903.
Wei, Y., Chen, M.C., 2012. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. Transport. Res. Part C 21 (1), 148–162.
Wei, Q., Qiu, Y., Wen, Y., 2022. Cluster-based spatiotemporal dual self-adaptive network for short-term subway passenger flow forecasting. Appl. Intell. 52 (12), 14137–14152.
Wu, X., Blake, A., 2022. Does the combination of models with different explanatory variables improve tourism demand forecasting performance? Tourism Econ. https://doi.org/10.1177/13548166221132.
Wu, D.C., Song, H., Shen, S., 2017. New developments in tourism and hotel demand modeling and forecasting. Int. J. Contemp. Hospit. Manag. 29 (1), 507–529.
Wu, W., Ma, X., Zeng, B., Wang, Y., Cai, W., 2019. Forecasting short-term renewable energy consumption of China using a novel fractional nonlinear grey Bernoulli model. Renew. Energy 140, 70–87.
Xiao, X., Yang, J., Mao, S., Wen, J., 2017. An improved seasonal rolling grey forecasting model using a cycle truncation accumulated generating operation for traffic flow. Appl. Math. Model. 51, 386–404.
Xie, N.M., Liu, S.F., 2009. Discrete grey forecasting model and its optimization. Appl. Math. Model. 33 (2), 1173–1186.
Xie, M., Wu, L., Li, B., Li, Z., 2020. A novel hybrid multivariate nonlinear grey model for forecasting the traffic-related emissions. Appl. Math. Model. 77, 1242–1254.
Yang, D., Chen, K., Yang, M., Zhao, X., 2019. Urban rail transit passenger flow forecast based on LSTM with enhanced long-term features. IET Intell. Transp. Syst. 13 (10), 1475–1482.
Yang, X., Xue, Q., Ding, M., Wu, J., Gao, Z., 2021. Short-term prediction of passenger volume for urban rail systems: a deep learning approach based on smart-card data. Int. J. Prod. Econ. https://doi.org/10.1016/j.ijpe.2020.107920.
Yao, Y., Cao, Y., Ding, X., Zhai, J., Liu, J., Luo, Y., Ma, S., Zou, K., 2018. A paired neural network model for tourist arrival forecasting. Expert Syst. Appl. 114, 588–614.
Yu, J., 2022. Short-term airline passenger flow prediction based on the attention mechanism and gated recurrent unit model. Cognitive Comput. 14 (2), 693–701.
Zhai, H., Tian, R., Cui, L., Xu, X., Zhang, W., 2020. A novel hierarchical hybrid model for short-term bus passenger flow forecasting. J. Adv. Transport. 2020 https://doi.org/10.1155/2020/7917353.
Zhang, J., Chen, F., Cui, Z., Guo, Y., Zhu, Y., 2021. Deep learning architecture for short-term passenger flow forecasting in urban rail transit. IEEE Trans. Intell. Transport. Syst. 22 (11), 7004–7014.
Zheng, W.Z., Lee, D.H., Shi, Q.X., 2006. Short-term freeway traffic flow prediction: bayesian combined neural network approach. J. Transport. Eng. 132 (2), 114–121.
Zhou, W., Wu, X., Ding, S., Ji, X., Pan, W., 2021. Predictions and mitigation strategies of $PM_{2.5}$ concentration in the Yangtze River Delta of China based on a novel nonlinear seasonal grey Model. Environ. Pollut. 276 https://doi.org/10.1016/j.envpol.2021.116614.