International Journal of Grey Systems

ISSN 2767-6412 / eISSN 2767-3308
2023 Volume 3 Issue 1: 17–28
https://doi.org/10.52812/ijgs.71

Posterior Variance Test: *Ex ante* Evaluation of Grey Forecasting model

Saad Ahmed Javed[1,*]

[1] *Operations Research Centre, GreySys Foundation, Lahore, Pakistan*

*Corresponding author: *saad.ahmed.javed@live.com*

Received 11 June 2023; *Revised* 27 July 2023; *Accepted* 29 July 2023

Abstract: Scholars frequently use *ex post* evaluation metrics such as the Mean Absolute Percentage Error (MAPE) to estimate the forecast accuracy. However, *ex ante* metrics are essential to know whether a given forecasting model is suitable for a given variable irrespective of the outcome of *ex post* evaluations. The *ex ante* measures help us ensure that the forecast is accurate, not by chance. The current study presents the Posterior Variance Test (PVT), which can serve as an *ex ante* measure of grey forecast accuracy. The study forecasted the methane emissions from Australia and India using a grey forecasting model and found that even though the MAPE generated "accurate forecasts" for both cases, the PVT invalidated the model's suitability for one of the two cases. The data visualization also corroborated the outcome of the PVT.

Keywords: Posterior error test; grey forecast; forecast accuracy; forecast error; methane emission.

1. Introduction

Ex ante and *ex post* are Latin words for *before* and *after* a particular event, respectively. According to the Collins English Dictionary, *ex ante* is something "based on what is expected to happen," while *ex post* is something "based on analysis of past performance" (Collins, 2023a; 2023b). Thus, *ex ante* means we look at future events based on possible predictions, while *ex post* means we look at results and events after they have occurred. The terminologies are particularly well-known in economics, planning, and forecasting literature (Steiger, 2018; Clements, 2014; Boardman *et al.*, 1994). In the forecasting context, *ex ante* measure of forecast accuracy implies testing the qualification of a forecasting model/technique before it is used to produce a simulation of a future event. *Ex post* measure of forecast accuracy implies testing the accuracy of that simulation (output of a forecasting model).

If a forecasting model fails to pass *ex ante* evaluation, its forecasting is deemed unscientific (and thus, unreliable) even if it did not fail *ex post* evaluation. Many "good" things may happen in real-life by "chance" (Rastrigin, 1984). A forecasting model can produce a "good" forecast by chance, and one would find *ex post* evaluation confirming the forecast accuracy. But there is no guarantee that the same model would produce a "good" forecast for another similar event. A forecast that happens to be accurate by chance is not scientific but an accurate guess. A forecasting model that is generally valid for all types of events is non-existent, as no single forecasting model can consistently outperform all other methods (Liu *et al.*, 2022). Nevertheless, a forecasting model that is generally valid for a particular type of event (e.g., forecasting project costs or duration) must be guided by theory, even though a theory (e.g., S-curve) is not exactly applicable to all types of problems (see, e.g., Al-Rasheed & Soliman, 2022). However, a theory can help the mathematical modellers develop more realistic models by allowing them to understand and incorporate general characteristics of a particular type of event (a set of similar events).

If one looks at the forecasting literature (see, e.g., Boamah, 2021; Ahmed *et al.*, 2020; Kırbaş *et al.*, 2020), including the grey forecasting literature (see, e.g., Cudjoe *et al.*, 2023; Singh *et al.*, 2022), one finds that most scholars have performed *ex post* evaluations. *Ex ante* evaluations are generally taken for granted. One reason can be that it is easier to compare multiple forecasting models (such as done by Khan and Osińska (2023), Wei and Xie (2022), and others) on *ex post* measures. *Ex ante* measures for different models can be different because of different parameters and guiding principles (and properties). Also, if necessary and sufficient conditions of a model are not known (Chen & Huang, 2013), constructing an *ex ante* measure is not easy. Deng Julong, the founder of the grey forecasting methodology, proposed the earliest known technique for *ex ante* evaluation of grey forecasting models (Javed & Cudjoe, 2022). Till today no better alternative has been proposed by the grey forecasters revealing the lack of interest in the *ex ante* evaluations. The current study aims to highlight the significance of this technique, the Posterior Variance Test (PVT), for grey forecasting. The PVT is a useful technique. By using it, data analysts can easily classify the forecasting models into qualified (valid) and unqualified (invalid) ones. Instead of applying a grey forecasting model first and then finding it unsuitable for a particular problem, the PVT helps them identify the unqualified models beforehand.

The rest of the study is organized as follows. The second section introduces grey forecasting, its one model, and an error metric. The third section presents the PVT. In the fourth section, the PVT is applied to two cases, and its performance is evaluated. In the fifth section, the study is concluded.

2. Grey forecasting

In 1982, Deng Julong proposed the Grey System Theory (GST), further refined by his pupil Liu Sifeng and others in succeeding decades. It is a practice-oriented research-intensive mathematical discipline (like operations research, artificial intelligence, etc.) that aims to solve complex problems and aid quantitative decision-making through a continuously evolving toolbox of intelligent mathematical models. However, unlike the Taguchi methods, it does not end at "Deng's methods" and can be extended with new information and developments, for instance, by updating older methods (e.g., the Dynamic Grey Relational Analysis is the generalized form of Deng's Grey Relational Analysis method), introducing new methods (e.g., the Grey Ordinal Priority Approach to multiple criteria decision analysis), and introducing new research directions (e.g., the concept of bidirectional grey relational analysis was introduced to the GST by Deng's successors). In his correspondence with Liu Sifeng, Lotfi A. Zadeh (the founder of Fuzzy Set Theory) noted the GST as "an important approach to uncertainty."[1]

Among the significant contributions of the GST is the methodology of grey forecasting, proposed by Deng (1984). Grey forecasting is a data-driven time series forecasting methodology that aims to discover the law governing the development of time series by transforming the original data's unobvious trend into a growth trend using the cumulative sum operator (Xie, 2022). In its simplest form, a grey forecasting model is a "grey differential equation," which is an organic (inseparable) integration of differential equation and regression model (Guo & Guo, 2009). It represents an essential contribution of the GST to differential equations (mathematics), regression (statistics), and forecasting. Since its introduction, grey forecasting has provided solutions to many real-world problems. For instance, a 1994 study reported its success in obtaining good representations of both temperature and precipitation for Cold Lake, Alberta, Canada (Bass *et al.*, 1996). In 2014, a grey forecasting model was used to forecast the traffic volume at the Qinhuangdao Port, China (Bin, 2014). In 2017, it was shown that with a small sample size and multiple failure

[1] The copy of the email by Zadeh to Liu can be obtained from the author at a reasonable request.

modes, a grey forecasting model outperformed the Crow/AMSAA reliability growth model in predicting system reliability growth parameters (Talafuse & Pohl, 2017). In 2023, Italian scholars performed the first diffusion analysis of ISO/IEC 27001 using the grey forecasting models (Podrecca & Sartor, 2023). Several grey forecasting models have been proposed to date. Most of them share some common characteristics. Based on these common characteristics, one can record the key steps of a grey forecasting model as:

1. Data collection (and preparation, if data is not in the appropriate form),
   a. The data sets must be positive and equidistant in classical models and their modern variants.
2. Data pre-processing through a cumulative sum operator (or its variants),
3. The grey differential equation (and its discrete approximation as a regression-like equation)
   a. Usually, these equations contain a set of parameters whose values are yet to be estimated.
4. A procedure to estimate the values of the aforementioned unknown parameters (such as through the Ordinary Least Square method),
5. Reverse data accumulation (unprocessing),
   a. It is usually done through a difference equation.
6. The solution to the differential equation.
7. Simulation (forecasting)

These are the procedures generally present in every grey forecasting model though some models have extra, and sometimes very complicated, procedures. Probably because of computational complexity, many grey forecasting models have not seen any application after being proposed, and only a few models have seen several applications. The next section presents a representative grey forecasting model to which the PVT will be applied later.

2.1 EGM (1, 1, α, θ)

EGM $(1, 1, \alpha, \theta)$ represents the even form of a grey model with a first-order differential equation containing one variable and weighted background value involving conformable fractional accumulation. It was proposed by Javed *et al.* (2020), and has seen several interesting applications (see, e.g., Tulkinov, 2023; Podrecca & Sartor, 2023; Singh *et al.*, 2022). Let the set of actual data is

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \dots\ x^{(0)}(n)\right), x^{(0)}(k) \geq 0, n \geq 4 \tag{1}$$

And the set of conformable fractional accumulated data of $X^{(0)}$ is

$$X^{(\alpha)} = \left(x^{(\alpha)}(1), x^{(\alpha)}(2), \dots\ x^{(\alpha)}(n)\right) \tag{2}$$

where, $x^{(\alpha)}(k) = \sum_{i=1}^{k}\left(\frac{x^{(0)}(i)}{i^{1-\alpha}}\right), k = 1,2, \dots, n$ and $\alpha \in (0,1]$. The adjacent neighbor average set of $X^{(1)}$ will be

$$Z^{(1)} = \left(z^{(1)}(1), z^{(1)}(2), \dots\ z^{(1)}(n)\right) \tag{3}$$

where the background value $z^{(1)}(k) = \theta . x^{(\alpha)}(k) + (1 - \theta). x^{(\alpha)}(k - 1)$ and $\theta \in [0,1]$.

The even form of GM (1, 1), a first-order, single-variable grey forecasting model with parameters a and b, is a continuous-time grey differential equation defined as

$$\frac{dx^{(1)}(k)}{dk} + ax^{(1)}(k) = b, k \geq 1 \tag{4}$$

Its discrete approximation is a linear regression-like equation given by

$$x^{(0)}(k) + az^{(1)}(k) = b \tag{5}$$

Or,

$$x^{(0)}(k) = -az^{(1)}(k) + b \tag{6}$$

The parameters a and b can be estimated through the Ordinary Least Square method, i.e.,

$$[a, b]^T = [B^T B]^{-1} B^T Y \tag{7}$$

where,

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \text{ and } Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$$

The time response function of the grey forecasting model is expressed as

$$\hat{x}^{(\alpha)}(k) = \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-a(k-1)} + \frac{b}{a} \ , k = 1,2, \dots, n \tag{8}$$

The inverse conformable fractional accumulation, which is needed to extract the simulation of the actual data set $\hat{x}^{(0)}(k)$ from the simulation of the accumulated data set $\hat{x}^{(\alpha)}(k)$, is executed through the following approximate regressive reduction formula.

$$\hat{x}^{(0)}(k) = k^{1-\alpha}\left(\hat{x}^{(\alpha)}(k) - \hat{x}^{(\alpha)}(k-1)\right), \qquad k = 1,2, \dots, n; \ \hat{x}^{(0)}(0) = 0 \tag{9}$$

This procedure of data unprocessing is unique to grey forecasting models. The time-response function of $X^{(0)}$, which is an exponential function of time, is given by

$$\hat{x}^{(0)}(k) = k^{1-\alpha}(1 - e^{a})\left(x^{(0)}(1) - \frac{b}{a}\right) e^{-a(k-1)} \, , k = 1,2, \dots, n \tag{10}$$

whereas, $\hat{x}^{(0)}(1) = x^{(\alpha)}(1) = x^{(0)}(1)$. For the complete detail on the even grey model, its parameters, and properties, Liu *et al.* (2022) can be referred to, and for complete information on conformable fractional accumulation, Ma *et al.* (2020) should be consulted.

EGM (1, 1, α, θ) is a generalized form of the classical model EGM (1, 1). When $\alpha = 1$ and $\theta = 0.5$, the EGM (1, 1, α, θ) reduces to the classical even grey forecasting model EGM (1, 1), which is available in Liu *et al.* (2022). The advantage of EGM (1, 1, α, θ) lies in its ability to adjust the parameters of θ and α with the variation in data. The noise in data can have different forms. Unlike the classical model EGM (1, 1), in the model EGM (1, 1, α, θ), the two parameters are not static but rather dynamic, and their values changes as the noise in data varies, thus helping us produce more accurate forecasts.

2.1.1 Finding the optimal values of θ and α. The relationship between θ and α is nonlinear. For the model EGM (1, 1, α, θ), the values of θ and α need to be estimated through an optimization problem. By following a few steps, as stipulated by Javed *et al.* (2020), their values can be predicted. The first step is to convert the MAPE formula (see the sub-section 2.1.2) into an optimization problem with the objective function containing the minimization of the MAPE (%) value, i.e.

$$\min_{\alpha,\theta} MAPE = \frac{1}{n-1} \times \sum_{k=2}^{n} \left| \frac{x^{(0)}(k) \ - \ \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \times 100$$

Subject to

$x^{(0)}(k)$ as defined in Section 2.1.

$\hat{x}^{(0)}(k)$ as defined in Section 2.1.

$\alpha \in (0,1]$

$\theta \in [0,1]$

The second step is to save the result in a temporary memory, such as

$$MEMORY = MAPE$$

The third step involves developing a constraint using the information stored in the memory. Then, the model in the first step is revised with the new constraint, as shown below, and the revised first step is executed.

$$MAPE \leq MEMORY$$

The fourth step involves repeating the second and third steps until no feasible solution is possible. Record the final value of MAPE (%) and corresponding optimum values of θ and α.

2.1.2 Forecast error measurement. The Mean Absolute Percentage Error (MAPE) is among the most popular goodness-of-fit measures widely used for *ex post* evaluation of grey forecasting models. It can easily be converted into accuracy (in percentage), and is defined as (Wu *et al.*, 2022; Guo & Guo, 2009),

$$MAPE(\%) = \frac{1}{n-1} \times \sum_{k=2}^{n} \left| \frac{x(k) - \hat{x}(k)}{x(k)} \right| \times 100 \tag{11}$$

where, $x(k)$ and $\hat{x}(k)$ represents actual observation and its simulation generated through the grey forecasting model, respectively.

C. D. Lewis developed a scale in 1982 that is being widely used to interpret the MAPE (%) values (Boamah, 2021; Klimberg *et al.*, 2015). According to the Lewis scale, a forecast error as high as 50% can be considered a "reasonable forecast." Javed and Cudjoe (2022) contested this view and updated the Lewis scale, which is given below.

$$MAPE(\%) = \begin{cases} < 10 & Highly\ accurate\ forecast \\ [10, 20) & Good\ forecast \\ [20, 30) & Reasonable\ forecast \\ \geq 30 & Inaccurate\ forecast \end{cases}$$

Literature (Gowrisankar *et al.*, 2022; Tulkinov, 2023) has confirmed the rationality and validity of the revised scale. The scale does not represent absolute truth, and a user may adjust it depending on the sensitivity of the forecasting problem.

3. The Posterior Variance Test

It has been observed that the fields (such as information systems) that emphasize the development of new methods address the evaluation of methods rather in a limited fashion (Moody, 2003). The methodology of the grey system, a system of "grey information" (Savić, 2019; Lin *et al.*, 2004), is no exception. Deng Julong, the founder of the Grey System Theory, knew the significance of this issue, which has not received much attention to date. The Posterior Variance Test (PVT), or posterior-error test, is a novel system of statistical hypothesis testing involving four hypotheses in two categories (two hypotheses in each category) and represents one of the significant scientific contributions of Deng. He also used the statistical theory of standard deviation for uncertainty quantification in his Grey Relational Analysis method (see, Javed *et al.*, 2022; Tan *et al.*, 2007). These facts show that statistical reasoning is essential in studying and evaluating novel grey models. This practice is consistent with scientific methods, where a new method must pass the evaluation by established method(s) to be considered valid (Loh *et al.*, 2022). Thus, in general, the grey system theory is not an alternative to statistics for data analysis; their role is complementary.

The PVT is a statistical technique used by Deng Julong to estimate the accuracy of grey forecasting models (Deng, 1996). Despite its known advantages (Liu *et al.*, 2014), the test has seen relatively fewer applications (Yu *et al.*, 2022; Ayvaz & Kusakci, 2017) and is generally underused as most of the grey forecasting literature only involves *ex post* evaluation of forecast accuracy, which is usually done through Mean Absolute Percentage Error (Tian *et al.*, 2021; Guo & Guo, 2009). Deng's other works, such as grey relational analysis, grey forecasting, and grey numbers, have received much greater attention. While refining the PVT, Javed and Cudjoe (2020) have recently emphasized the significance of the PVT as an *ex ante* measure for grey forecasting models.

Ex ante evaluations are indispensable for forecasting models (Steece, 1986). If the result of *ex ante* evaluation is unsatisfactory, it provides an opportunity to revisit data or modify the model. The PVT provides an *ex ante* performance measure of a time-series forecasting model to see whether the selected model is suitable for forecast or not (Yu *et al.*, 2022). The PVT is guided by two metrics: the "ratio of root-mean-square deviations" (C) and the "small-error probability" (P). For highly accurate forecasts, C should be the minimum, and P should be the maximum. If needed, the root-mean-square deviation can be replaced by the standard deviation. The PVT's goal is to confirm whether the simulations (forecasts) are C-satisfactory or P-satisfactory, or both. A highly reliable forecast is both C-satisfactory and P-satisfactory. The procedure to execute the PVT is explained below.

If $X = \{x(1), x(2), \dots, x(n)\}$ represents the set of actual observations, $\hat{X} = \{\hat{x}(1), \hat{x}(2), \dots, \hat{x}(n)\}$ represents the model-generated simulation of X, and $\epsilon = \{\epsilon(1), \epsilon(2), \dots, \epsilon(n)\}$, $\epsilon(k) = x(k) - \hat{x}(k),\ k = 2,3, \dots, n$ represents the set of absolute errors (residuals) then the root-mean-square deviations S_1 and S_2 are calculated as

$$S_1 = \sqrt{\frac{1}{n-1}\sum_{k=2}^{n}(x(k) - \bar{x})^2}\,, \bar{x} = \frac{1}{n-1}\sum_{k=2}^{n} x(k) \quad (12)$$

And

$$S_2 = \sqrt{\frac{1}{n-1}\sum_{k=2}^{n}(\epsilon(k) - \bar{\epsilon})^2}\,, \bar{\epsilon} = \frac{1}{n-1}\sum_{k=2}^{n} \epsilon(k) \quad (13)$$

The square of S_1 and S_2 gives variance in the data sets. If the ratio of root-mean-square deviations $C = \frac{S_2}{S_1} < 0.35$, then the model (or simulation) is said to be C-satisfactory, or the model is significant with respect to C. In the aforementioned equations, k is starting from 2 because, in most grey forecasting models, the first simulated value is usually the same as the first actual observation and thus can be ignored. If this is not the case, k should start from 1. C is a cost-type metric ("lower, the better").

The small-error probability $P = \mathrm{P}\{|\epsilon(k) - \bar{\epsilon}| < 0.6745S_1\}$. Here, 0.6745 represents the root-mean-square value for the 75th percentile in the z-scores statistical table. P is a profit-type metric ("higher, the better"). A larger P value implies that at more points (k) the residuals are below the threshold of $0.6745S_1$. In short, if $P > 0.95$, this means the probability of smaller errors (P) is higher than that of larger errors, and therefore the model is P-satisfactory, or the model is significant with respect to P. $P = 1$, when $|\epsilon(k) - \bar{\epsilon}| < 0.6745S_1$ for every value of k, and $P = 0$ when the inequality does not hold for any value of k. If a grey forecasting model is both C-satisfactory and P-satisfactory, then its predictions are considered reliable. Otherwise, it shows that the model is unsuitable for forecasting the given variable because of considerable variations in data

(Javed & Cudjoe, 2022). The Deng scales (Deng, 1996), shown in Table 1, can guide us in judging whether a grey forecasting model is suitable for use or not.

Unlike C, the precise formula to estimate P was unavailable in the literature. By building upon the frequentist approach to probability, Javed and Cudjoe (2022) defined the small-error probability P as a function of the ratio of the number of favourable outcomes and the total number of outcomes such as,

$$P = \frac{F}{U + F} \tag{14}$$

where, F and U and represent the number of favourable and unfavourable outcomes, whereas favourability imply $|\epsilon(k) - \bar{\epsilon}| < 0.6745S_1$. Thus, $F = f_F\{|\epsilon(k) - \bar{\epsilon}| < 0.6745S_1\}$, and $U = f_U\{|\epsilon(k) - \bar{\epsilon}| \geq 0.6745S_1\}$, where f represents the number of times (frequency) the inequality holds (f_F) or not (f_U). The sum of f_F and f_U equals the total number of simulations. In most grey forecasting models, the first observation is not simulated and is used as a seed in the forecasting process. Thus, $f_F + f_U = n - 1$, where n represents the sample size. For further details about the PVT, Javed and Cudjoe (2020) can be consulted.

4. Data and results

Methane is the second most abundant anthropogenic greenhouse gas after carbon dioxide (CO_2), accounting for about 20% of global emissions, and its warming capacity is 80 times greater than CO_2 over a 20-year period (UNEP, 2021; EPA, 2023). Forecasting methane emissions is of great significance for energy policy makers and environmentalists. To test the PVT on the grey forecasting model, data on "Methane emissions (kt of CO_2 equivalent)" was collected from the World Bank Indicators for Australia and India. Data from 2009 to 2017 was used for generating simulations from 2018 to 2028. Data from 2018 to 2020 played no role in the forecast and was kept for out-of-sample testing. The actual data and simulations are shown in Table 2.

The in-sample MAPE (%) was 9.11% for Australia and 0.16% for India. One may think the model is suitable for both cases; however, out-of-sample testing is still needed to monitor the problem of overfitting. The out-of-sample MAPE (%) was 6.26% for Australia and 0.36% for India. According to the MAPE scale (Javed & Cudjoe, 2022), in both cases, the forecast is "highly accurate," and *ex post* evaluation is passed. It is very tempting to discuss the results without touching the difficult question; was the chosen model suitable for the given problem? One can hardly answer the question without performing *ex ante* evaluation.

The *ex ante* assessments done through the PVT for Australia and India are shown in Tables 3 and 4, respectively. The assessments are summarized in Table 5. One can see that C is close to 1, and P is much lower than 0.7 for Australia. Thus, the model is neither C-satisfactory nor P-satisfactory. The model is unqualified to forecast methane emissions (kt of CO_2 equivalent) from Australia. For India, $C < 0.35$ and $P > 0.95$. Thus, the model is both C-satisfactory and P-satisfactory. The model is good for forecasting India's methane emissions (kt of CO_2 equivalent).

Data visualization also sheds important insights into this matter. For instance, if one looks at Figure 1 (Australia) and Figure 2 (India), it can be seen that in the case of Australia, the trend of methane emissions is less linear and more unpredictable, as compared to the case of India. In the case of Australian methane emissions, the simulation of actual data generated through the model

Table 1. The evaluation scales of the Posterior Variance Test

Forecast Accuracy	*P*	*C*
Good	> 0.95	< 0.35
Qualified	0.80 ~ 0.95	0.35 ~ 0.50
Barely Qualified	0.70 ~ 0.80	0.50 ~ 0.65
Unqualified	≤ 0.7	≥ 0.65

Table 2. Forecasting methane emissions (kt of CO_2 equivalent) from Australia and India

	Australia		**India**	
	Actual	**EGM (1, 1, α, θ)**	**Actual**	**EGM (1, 1, α, θ)**
2009	137198	137198	651519	651519
2010	129412	156129	658933	659079
2011	177597	155136	666188	662984
2012	176798	154149	665448	666902
2013	140884	153168	669582	670837
2014	154126	152194	674058	674792
2015	151225	151225	678829	678768
2016	134687	150263	684446	682766
2017	157518	149307	686785	686785
2018	*156428*	148357	*695494*	690828
2019	*144212*	147413	*696432*	694893
2020	*131485*	146475	*697655*	698981
2021		145544		703093
2022		144618		707229
2023		143697		711388
2024		142783		715571
2025		141875		719779
2026		140972		724011
2027		140075		728268
2028		139184		732549
MAPE % (in-sample)		9.11%		0.16%
MAPE % (out of sample)		6.26%		0.36%
a		0.006382625		-0.005855098
b		157503.6963		653278.457
α		1		0.999871968
θ		0.411832364		0.576532684

NOTE: Italic numbers did not participate in modelling and forecasting processes, and were only used for out-of-sample testing.

Table 3. *Ex ante* forecast error analysis by the PVT for Australia

	$\epsilon(k)$	$(\epsilon(k)-\bar{\epsilon})^2$	$(x(k)-\bar{x})^2$	$\lvert\epsilon(k)-\bar{\epsilon}\rvert$	$\lvert\epsilon(k)-\bar{\epsilon}\rvert< 0.6745S_1$	Favourability
2010	-26717	718319361	546095062	26801	No	U
2011	22461	500688568	615810730	22376	No	U
2012	22649	509160436	576817511	22565	No	U
2013	-12284	152989808	141543471	12369	No	U
2014	1933	3415047	1809519	1848	Yes	F
2015	0	7152	2420112	85	Yes	F
2016	-15576	245247576	327379282	15660	No	U
2017	8211	66043203	22443286	8127	Yes	F

Table 4. *Ex ante* forecast error analysis by the PVT for India

	$\epsilon(k)$	$(\epsilon(k)-\bar{\epsilon})^2$	$(x(k)-\bar{x})^2$	$\lvert\epsilon(k)-\bar{\epsilon}\rvert$	$\lvert\epsilon(k)-\bar{\epsilon}\rvert< 0.6745S_1$	Favourability
2010	-146	99383	198816266	315	Yes	F
2011	3204	9207525	46865978	3034	Yes	F
2012	-1454	2635095	57544003	1623	Yes	F
2013	-1255	2029105	11913213	1424	Yes	F
2014	-734	817065	1048754	904	Yes	F
2015	61	11765	33587745	108	Yes	F
2016	1680	2281929	130237202	1511	Yes	F
2017	0	28763	189109033	170	Yes	F

Table 5. Key parameters and final assessment of the suitability of the model through the PVT

Parameters	Australia	India
S_1	16711.97	9145.51
S_2	16567.56	1462.47
C	0.991 > 0.65 (Unqualified)	0.160 < 0.35 (Good)
$0.6745\ S_1$	11272.2	6168.6
f_F	3	8
f_U	5	0
P	0.375 < 0.7 (Unqualified)	1.00 > 0.95 (Good)

is not well-fitted with the actual data. Thus, the conclusions drawn from the PVT are reliable, while those from the MAPE are debatable. It is found that by using the PVT, poor forecasts and interpretations can be avoided.

5. Conclusion and recommendations

Forecasting is an important problem in several disciplines and industries. Grey forecasting models represent an emerging methodology of forecasting. They have seen applications in various areas of practical importance. It has been observed that most of the literature involving grey forecasting only includes *ex post* evaluation of forecast accuracy, and the Mean Absolute Percentage Error (MAPE) is a popular choice among scholars. The current study argues that *ex ante* evaluation of forecast accuracy is essential, but it did not receive due attention from the grey forecasting community even though the founder of the grey forecasting methodology himself emphasized its significance when he proposed the Posterior Variance Test (PVT).

The study forecasted the methane emissions (kt of CO_2 equivalent) from Australia and India using the optimized even grey forecasting model EGM (1, 1, α, θ). The *ex ante* evaluation of

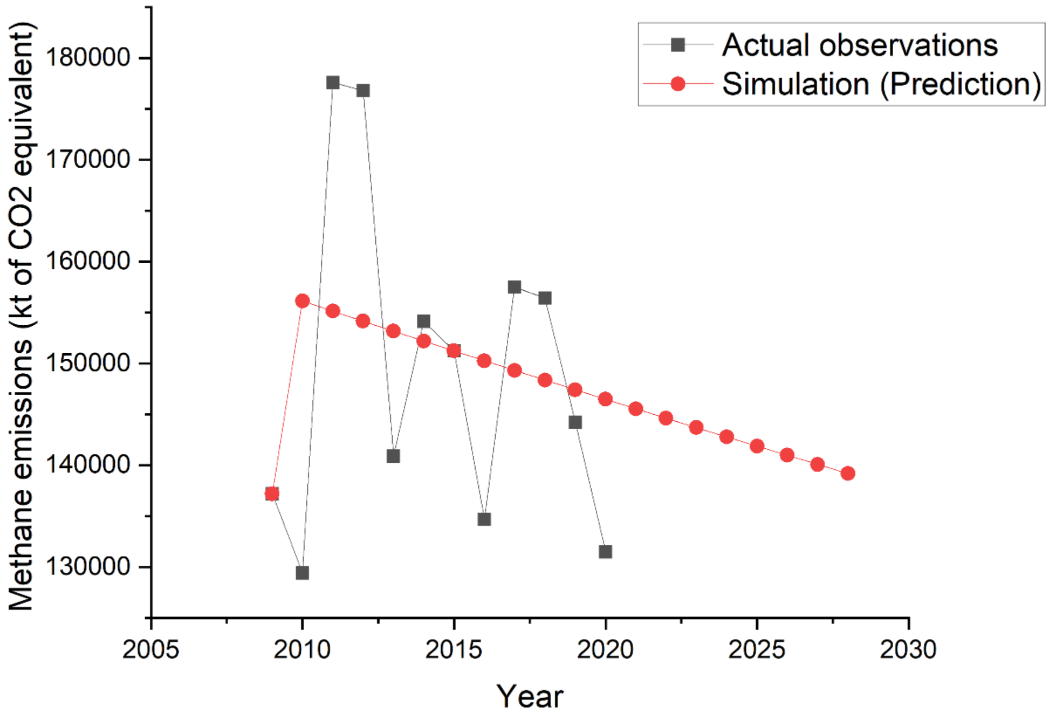

Fig 1. "Unqualified" forecast of methane emissions in Australia through EGM (1, 1, α, θ)

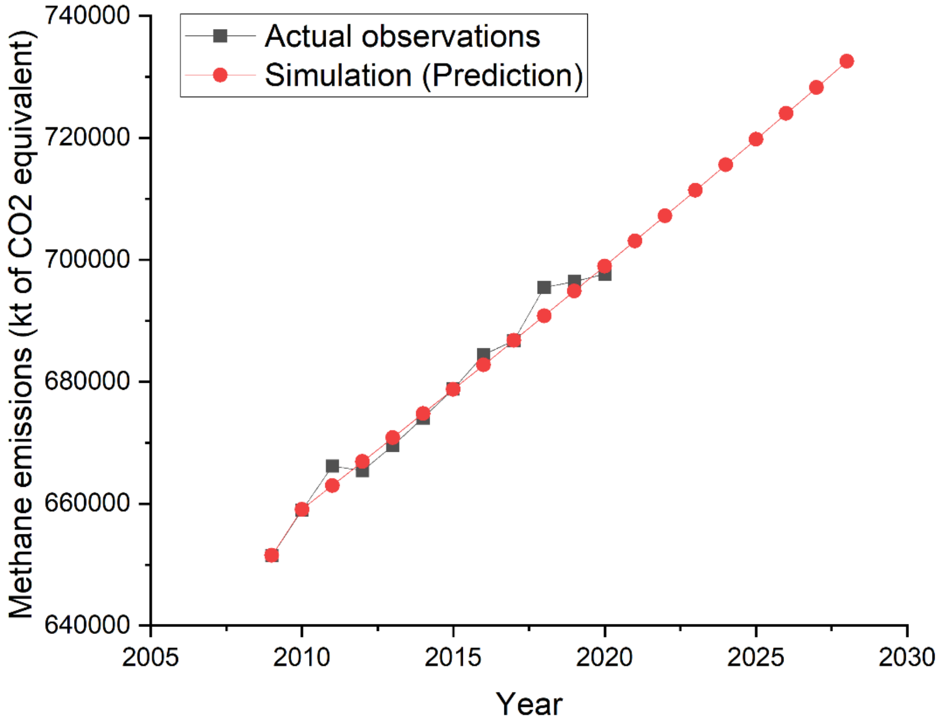

Fig 2. "Good" forecast of methane emissions in India through EGM (1, 1, α, θ)

forecast accuracy was performed through the PVT, and *ex post* evaluation was done through the MAPE. It was found that even though the MAPE certified that the model is suitable for forecasting methane emissions from both Australia and India, the PVT contested this argument (for Australia) and only certified the model's validity for India.

The results serve as a wakeup call for those who are applying the grey forecasting models without using the PVT or other *ex ante* measures. The study argues that the *ex post* evaluations, done through the MAPE or other metrics, are insufficient to establish trust in the forecast generated through mathematical models like grey forecasting models. New *ex ante* measures of grey forecasting models can be proposed in the future. Also, the validity of the PVT for different kinds of grey forecasting models is yet to be seen.

Acknowledgements

The work is supported by the Natural Science Research Foundation of Jiangsu Higher Education Institutions of China (No. 21KJB480011).

References

Ahmed, R., Sreeram, V., Mishra, Y., & Arif, M. D. (2020). A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. *Renewable and Sustainable Energy Reviews*, *124*, 109792. https://doi.org/10.1016/j.rser.2020.109792

Al-Rasheed, K. A., & Soliman, E. (2022). Difference in S-curve for different types of construction projects. *Journal of Engineering Research*, *10*(1B), 17-28. https://doi.org/10.36909/jer.10231

Ayvaz, B., & Kusakci, A. O. (2017). Electricity consumption forecasting for Turkey with non-homogeneous discrete grey model. *Energy Sources Part B – Economics Planning and Policy*, *12*(3), 260–267. https://doi.org/10.1080/15567249.2015.1089337

Bass, B., Akkur, N., Russo, J., & Zack, J. (1996). *Modelling the Biospheric Aspects of the Hydrological Cycle – Upscaling Processes and Downscaling Weather Data*. In: Regional Hydrological Respon.se to Climate Change (J. A. A. Jones *et al.* (eds.)), 39-62. https://doi.org/10.1007/978-94-011-5676-9_3

Bin, H. (2014). Using grey system theory, the port traffic volume and market share are forecasted and analyzed [运用灰色系统理论开展港口运量及占有率预测分析], *Business Culture*, *20*, 213-214.

Boamah, V. (2021). Forecasting the Demand of Oil in Ghana: A Statistical Approach. *Management Science and Business Decisions*, *1*(1), 29-43. https://doi.org/10.52812/msbd.25

Boardman, A. E., Mallery, W. L., & Vining, A. R. (1994). Learning from ex ante/ex post cost-benefit comparisons: the Coquihalla highway example. *Socio-Economic Planning Sciences*, *28*(2), 69-84.

Chen, C.-I., & Huang, S.-J. (2013). The necessary and sufficient condition for GM(1, 1) grey prediction model. *Applied Mathematics and Computation*, *219*(11), 6152-6162. https://doi.org/10.1016/j.amc.2012.12.015

Clements, M. P. (2014). Forecast uncertainty – ex ante and ex post: US inflation and output growth. *Journal of Business & Economic Statistics*, *32*(2), 206-216.

Collins. (2023a). *Ex ante*. Collins English Dictionary. https://www.collinsdictionary.com/dictionary/english/ex-ante

Collins. (2023b). *Ex post*. Collins English Dictionary. https://www.collinsdictionary.com/dictionary/english/ex-post

Cudjoe, D., Brahim, T., & Zhu, B. (2023). Assessing the economic and ecological viability of generating electricity from oil derived from pyrolysis of plastic waste in China. *Waste Management*, *168*, 354-365. https://doi.org/10.1016/j.wasman.2023.06.015

Deng, J. (1984). The Differential Grey Model (GM) and its Implement in Long Period Forecasting of Grain [灰色动态模型(GM)及在粮食长期预测中的应用], *Exploration of Nature Magazine*, *3*, 37-43.

Deng, J. (1996). *Basic Methods of Grey Systems* [灰色系统基本方法](4th Ed.). Huazhong University of Science and Technology Press.

EPA. (2023). *Importance of Methane*. The United States Environmental Protection Agency. https://www.epa.gov/gmi/importance-methane

Gowrisankar, A., Priyanka, T. M. C., Saha, A., Rondoni, L., Kamrul Hassan, M., & Banerjee, S. (2022). Greenhouse gas emissions: A rapid submerge of the world. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *32*(6), 061104. https://doi.org/10.1063/5.0091843

Guo, R., & Guo, D. (2009). Random fuzzy variable foundation for grey differential equation modeling. *Soft Computing*, *13*, 185-201. https://doi.org/10.1007/s00500-008-0301-4

Javed S. A, Zhu, B., & Liu S. (2020). Forecast of Biofuel Production and Consumption in Top CO2 Emitting Countries using a novel grey model. *Journal of Cleaner Production*, *276*, 123977. https://doi.org/10.1016/j.jclepro.2020.123997

Javed, S. A., & Cudjoe, D. (2022). A novel Grey Forecasting of Greenhouse Gas Emissions from four Industries of China and India. *Sustainable Production and Consumption*, *29*, 777-790. https://doi.org/10.1016/j.spc.2021.11.017

Javed, S. A., Gunasekaran, A., & Mahmoudi, A. (2022). DGRA: Multi-sourcing and Supplier Classification through Dynamic Grey Relational Analysis method. *Computers & Industrial Engineering*, *173*, 108674. https://doi.org/10.1016/j.cie.2022.108674

Khan, A. M., & Osińska, M. (2023). Comparing forecasting accuracy of selected grey and time series models based on energy consumption in Brazil and India. *Expert Systems with Applications*, *212*, 118840. https://doi.org/10.1016/j.eswa.2022.118840

Kırbaş, İ., Sözen, A., Tuncer, A. D., & Kazancıoğlu, F. Ş. (2020). Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches. *Chaos, Solitons & Fractals*, *138*, 110015. https://doi.org/10.1016/j.chaos.2020.110015

Klimberg, R. K., Sillup, G. P., Boyle, K. J., & Tavva, V. (2015). *Forecasting performance measures – what are their practical meaning?*. In: Advances in Business and Management Forecasting. http://dx.doi.org/10.1108/S1477-4070(2010)0000007012

Lin, Y., Chen, M. Y., & Liu, S. (2004). Theory of grey systems: capturing uncertainties of grey information. *Kybernetes*, *33*(2), 196-218. https://doi.org/10.1108/03684920410514139

Liu, A., Lin, V. S., Li, G., & Song, H. (2022). Ex ante tourism forecasting assessment. *Journal of Travel Research*, *61*(1), 64-75. https://doi.org/10.1177/0047287520974456

Liu, S., Yang, Y, Wu, L., *et al.* (2014). Grey system theory and its application [灰色系统理论及其应用] (7th Ed.). Beijing: Science Press.

Liu, S., Yang, Y., & Forrest, J. Y.-L. (2022). *Grey Systems Analysis – Methods, Models and Applications*. Singapore: Springer.

Loh, T. P., Cooke, B. R., Markus, C., Zakaria, R., Tran, M. T. C., Ho, C. S., ... & IFCC Working Group on Method Evaluation Protocols. (2023). Method evaluation in the clinical laboratory. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *61*(5), 751-758. https://doi.org/10.1515/cclm-2022-0878

Ma, X., Wu, W., Zeng, B., Wang, Y., & Wu, X. (2020). The conformable fractional grey system model. *ISA Transactions*, *96*, 255-271. https://doi.org/10.1016/j.isatra.2019.07.009

Moody, D. L. (2003). The method evaluation model: a theoretical model for validating information systems design methods. *ECIS 2003 Proceedings*, *79*. http://aisel.aisnet.org/ecis2003/79

Podrecca, M., & Sartor, M. (2023). Forecasting the diffusion of ISO/IEC 27001: a Grey model approach. *The TQM Journal*, *35*(9), 123-151. https://doi.org/10.1108/TQM-07-2022-0220

Rastrigin, L. (1984). *This Chancy, Chancy, Chancy World*. Moscow: Mir Publishers.

Savić, D. (2019). When is 'grey' too 'grey'? – A case of grey data. *The Grey Journal*, *15*(2), 71-76.

Singh, P. K., Pandey, A. K., & Bose, S. C. (2022). A new grey system approach to forecast closing price of Bitcoin, Bionic, Cardano, Dogecoin, Ethereum, XRP Cryptocurrencies. *Quality & Quantity*, *57*, 2429–2446. https://doi.org/10.1007/s11135-022-01463-0

Steece, B. (1986). Ex-ante measures for evaluating forecasting models. *Engineering Costs and Production Economics*, *10*(1), 25-34. https://doi.org/10.1016/0167-188X(86)90017-0

Steiger, O. (2018). *Ex ante and Ex post*. In: The New Palgrave Dictionary of Economics. London: Palgrave Macmillan. https://doi.org/10.1057/978-1-349-95189-5_315

Talafuse, T. P., & Pohl, E. A. (2017). Small sample reliability growth modeling using a grey systems model. *Quality Engineering*, *29*(3), 455-467. https://doi.org/10.1080/08982112.2017.1318920

Tan, X., Deng, J., & Chen, X. (2007). *Generalized grey relational grade and grey relational order test*. In: 2007 IEEE International Conference on Systems, Man and Cybernetics (pp. 3928-3931). IEEE. https://doi.org/10.1109/ICSMC.2007.4414141

Tian, X., Wu, W., Ma, X., & Zhang, P. (2021). A new information priority accumulated grey model with hyperbolic sinusoidal term and its applications. *International Journal of Grey Systems*, *1*(2), 5-19. https://doi.org/10.52812/ijgs.27

Tulkinov, S. (2023). Grey forecast of electricity production from coal and renewable sources in the USA, Japan and China. *Grey Systems: Theory and Application*, *13*(3), 517-543. https://doi.org/10.1108/GS-10-2022-0107

UNEP. (2021). *Methane emissions are driving climate change. Here's how to reduce them*. The United Nations Environment Programme. https://www.unep.org/news-and-stories/story/methane-emissions-are-driving-climate-change-heres-how-reduce-them

Wei, B., & Xie, N. (2022). On unified framework for continuous-time grey models: An integral matching perspective. *Applied Mathematical Modelling*, *101*, 432-452. https://doi.org/10.1016/j.apm.2021.09.008

Wu, W., Ma, X., Zhang, H., Tian, X., Zhang, G., & Zhang, P. (2022). A Conformable Fractional Discrete Grey Model CFDGM (1,1) and its Application. *International Journal of Grey Systems*, *2*(1), 5-15. https://doi.org/10.52812/ijgs.36

Xie, N. (2022). A summary of grey forecasting models. *Grey Systems: Theory and Application*, *12*(4), 703-722. https://doi.org/10.1108/GS-06-2022-0066

Yu, W., Xia, L., & Cao, Q. (2023). Forecasting digital economy of China using an Adaptive Lasso and grey model optimized by particle swarm optimization algorithm. *Journal of Intelligent & Fuzzy Systems*, *44*(2), 2543-2560. https://doi.org/10.3233/jifs-222520