

Grey System Model GM (1,1) for Predicting Student Performance

Adel Najar, Naser Qamhieh, Noureddine Amrane, Saleh Mahmood

Department of Physics, College of Science, United Arab Emirates University, Al Ain, United Arab Emirates,

Corresponding author: adel.najar@uaeu.ac.ae

doi: <https://doi.org/10.37745/bjmas.2022.04153>

Published July 26, 2023

Citation: Najar A., Qamhieh N., Amrane N., and Mahmood S. (2024) Grey System Model GM (1,1) for Predicting Student Performance, *British Journal of Multidisciplinary and Advanced Studies*,5(4),8-18

ABSTRACT: *The purpose of this study is to predict student quiz assignments in the General Physics I course during one semester using the Grey system GM (1,1) model. The prediction model integrates the first-order two-variable grey differential equation model from the Grey system theory. Additionally, a MATLAB program for the GM (1,1) prediction model was developed. The predicted results were compared with the original GM (1,1) model, showing good agreement.*

KEYWORDS: Grey model, prediction, assignments

INTRODUCTION

People are highly interested in forecasting the future trend of some events, such as population growth, business growth, and investment in the stock market, which are necessary to be forecasted to make decisions or to develop some policies. Since prediction is mainly used to reduce uncertainty or risk, therefore the accuracy of the prediction is very important. How to select an appropriate method that is accurate to forecast an event in the future is a crucial problem. Consequently, a method with low cost and high accuracy of forecasting has been the goal of management decision-makers.

The traditional prediction model needs to deal with a large number of samples or normal distribution that cannot be used for short-term prediction. To overcome these limitations, in recent years several methods were developed. Artificial intelligence was introduced to modify traditional prediction methods by using artificial neural networks (Wei & Yang, 1999), the regression analysis that identifies and measures causal relationships between variables (Yen & Lin, 1997), multivariate time series models used by Gooijer et al. (de Gooijer & Klein, 1989), and Grey system theory was proposed by Deng (Julong, 1989).

Grey system model GM (m, n) has two parameters n and m, where m is the order of the difference equation and n is the number of variables. The grey model is used in a system with incomplete information, uncertain behavior, and unclear mechanisms. It's used to understand analysis, observe system developments, and make forecasting. It's a strong point that can be used with only four pieces of data. Furthermore, the distribution of the samples doesn't require several assumptions to be made. In addition, several studies have demonstrated that the Grey model (1,1) has extremely high prediction accuracy with only a small number of data (Deng, 2003; Julong, 2004).

In this paper, we developed a code on MATLAB based on the GM (1,1) model to predict the quiz grades for students in General Physics I at United Arab Emirates University (UAEU).

METHOD

The amount of data for the GM (1,1) model can be as little as four pieces. In this study, we addressed series prediction problems. Series prediction involves establishing a grey prediction model directly based on given data. Different Grey models were introduced including GM (1,1), GM (1, n), GM (2,1), GM (2,n), and GM (0,n). We focus on GM (1,1) due to its computational efficiency.

Grey Model (1,1)

The steps in the construction of the grey prediction model GM(1,1) are as follows:

If data have $n \geq 4$, $x(0) \in \mathbb{R}^+$, and

$$\sigma^{(0)}(\mathbf{i}) \in \left(\frac{-2}{e^{n+1}}, \frac{+2}{e^{n+1}} \right)$$

$$\sigma^{(0)}(\mathbf{i}) = \frac{x^{(0)}(n-1)}{x^{(0)}(\mathbf{i})} \quad (1)$$

where $\sigma^{(0)}(\mathbf{i})$ is called class ratio.

The most commonly used grey forecasting model is GM (1,1), which indicates one variable is employed in the model and the first-order differential equation is adopted to match the data generated by the Accumulation Generating Operation (AGO). The AGO reveals the hidden regular pattern in the system development. Before the algorithm of GM(1,1) is described (Zhang et al., 2012), the raw data series is assumed to be:

$$\mathbf{x}^{(0)} = (\mathbf{x}^{(0)}(\mathbf{1}), \mathbf{x}^{(0)}(\mathbf{2}), \mathbf{x}^{(0)}(\mathbf{3}), \dots, \mathbf{x}^{(0)}(\mathbf{n})) \quad (2)$$

where n is the total number of modeling data. The AGO formation of $x(1)$ is defined as:

$$\mathbf{x}^{(1)} = (\mathbf{x}^{(1)}(1), \mathbf{x}^{(1)}(2), \mathbf{x}^{(1)}(3), \dots, \mathbf{x}^{(1)}(n)) \quad (3)$$

Where $\mathbf{x}^{(1)}(1) = \mathbf{x}^{(0)}(1)$ and

$$\mathbf{x}^{(1)}(k) = \sum_{i=1}^k \mathbf{x}^{(0)}(i), k = 1, 2, \dots, n \quad (4)$$

The GM(1,1) model can be constructed by establishing a first-order differential equation for $x(1)(k)$ as

$$\frac{d\mathbf{x}^{(1)}}{dt} + \mathbf{a}\mathbf{x}^{(1)} = \mathbf{b} \quad (5)$$

Then the discrete form of the GM(1,1) differential equation model is expressed as

$$\mathbf{X}(0) + \mathbf{a}\mathbf{z}(1) = \mathbf{b}$$

Where $\mathbf{z}^{(1)} = (\mathbf{z}^{(1)}(2), \mathbf{z}^{(1)}(3), \mathbf{z}^{(1)}(4), \dots, \mathbf{z}^{(1)}(n))$ are called background values of $\frac{d\mathbf{x}^{(1)}}{dt}$ and calculated by $\mathbf{z}^{(1)}(t) = 0.5(\mathbf{x}^{(1)}(t-1) + \mathbf{x}^{(1)}(t))$, $t=2,3,\dots, n$

where parameters a and b are called the developing coefficient and grey input, respectively. In practice, parameters a and b are not calculated directly from (5). Therefore, the solution of (5) can be obtained by using the least square method. That is:

$$\hat{\mathbf{x}}^{(1)}(k+1) = \left(\mathbf{x}^{(0)}(1) - \frac{\mathbf{b}}{\mathbf{a}} \right) e^{-\mathbf{a}k} + \frac{\mathbf{b}}{\mathbf{a}} \quad (6)$$

$$\text{Where } \hat{\mathbf{a}} = [\mathbf{a}, \mathbf{b}]^T = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y} \quad (7)$$

$$\text{And } \begin{bmatrix} -0.5(\mathbf{x}^{(1)}(1) + \mathbf{x}^{(1)}(2)) & \mathbf{1} \\ -0.5(\mathbf{x}^{(1)}(2) + \mathbf{x}^{(1)}(3)) & \mathbf{1} \\ \dots & \dots \\ -0.5(\mathbf{x}^{(1)}(n-1) + \mathbf{x}^{(1)}(n)) & \mathbf{1} \end{bmatrix} \quad (8)$$

$$\mathbf{Y} = [\mathbf{x}^{(0)}(2), \mathbf{x}^{(0)}(3), \dots, \mathbf{x}^{(0)}(n)]^T \quad (9)$$

The result was obtained $\hat{\mathbf{x}}^{(1)}$ from (6). Let $\hat{\mathbf{x}}^{(0)}$ be the fitted and predicted series

$$\hat{\mathbf{x}}^{(0)} = (\hat{\mathbf{x}}^{(0)}(1), \hat{\mathbf{x}}^{(0)}(2), \dots, \hat{\mathbf{x}}^{(0)}(n), \dots) \quad (10)$$

Where $\hat{\mathbf{x}}^{(0)}(1) = \mathbf{x}^{(0)}(1)$

Applying the inverse accumulated generation operation (IAGO). The predicted equation is:

$$\hat{\mathbf{x}}^{(0)}(k) = \left(\mathbf{x}^{(0)}(1) - \frac{\mathbf{b}}{\mathbf{a}} \right) (1 - e^{-\mathbf{a}}) e^{-\mathbf{a}(k-1)} + \frac{\mathbf{b}}{\mathbf{a}} \quad (11)$$

Where $k=2, 3, \dots, n$. $\hat{\mathbf{x}}^{(0)}(1), \hat{\mathbf{x}}^{(0)}(2), \dots, \hat{\mathbf{x}}^{(0)}(n)$ are called the GM(1,1) fitted sequence, while $\hat{\mathbf{x}}^{(0)}(n+1), \hat{\mathbf{x}}^{(0)}(n+2), \dots, \hat{\mathbf{x}}^{(0)}(n+2), \dots$ are called the GM(1,1) forecast values (Askari & Askari, 2011; Hongli, 2013).

Error Analysis: the methods used for error analysis are mean absolute percentage error (MAPE) and mean absolute percentage accuracy (MAPA). The mean absolute percentage error is a good evaluation criterion often used to compare prediction performances (Ken et al., 2010).

$$\text{MAPE} = \frac{1}{n} \sum_{k=2}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| * 100\% \quad (12)$$

$$\text{MAPA} = \left(-\frac{1}{n} \sum_{k=2}^n \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right| \right) * 100\% \quad (13)$$

MATLAB Algorithm

Input $x(0)$; is the original data $x^{(0)}=(x^{(0)}(1),x^{(0)}(2),x^{(0)}(3),\dots,x^{(0)}(n))$

Testing data; represents the initial data consistent with the prediction model.

for $i=1$ to $n-1$ do

Calculation of the background values $z^{(1)}$ based on the method of accumulated generating operation

end do

for $i=1$ to $n-1$ do

calculate the background values $z^{(1)}$

end do

calculate the coefficients $\hat{a} = [a, b]^T$

for $i=1$ to $n+j$ do

calculate the fitted values and the forecast values $\hat{x}^{(0)}$.

end do

Error analysis is calculated using mean absolute percentage error (MAPE) and mean absolute percentage accuracy (MAPA).

RESULTS & DISCUSSION

In analyzing student performance in the General Physics course presented to students in their first year in science and engineering schools at United Arab Emirates University (UAEU). Conducting face-to-face assignments for students provides opportunities for collaborative problem-solving, interaction with peers, and direct instruction from the teacher. Through face-to-face discussions, demonstrations, and laboratory experiments, students can deepen their

Published by European Centre for Research Training and Development-UK
understanding of physical principles and develop critical thinking skills in a supportive classroom environment. The quizzes were conducted using Blackboard with Respondus LockDown Browser and MasteringPhysics (provided by Pearson) as tools during the face-to-face assignments in the class. A pilot study was conducted in Fall 2022. The study focused on an introductory physics course (Newtonian Mechanics) and involved 40 students across one female section. This section took the course face-to-face and completed quizzes in the class using their laptops.

Furthermore, to enhance student's understanding and learning before taking the quizzes, they were required to complete weekly MasteringPhysics homework assignments. These homework assignments included end-of-chapter discussion questions, exercises, and problems, which were due several days after the lecture. The assignments aimed to help students practice conceptual understanding, problem-solving, and critical thinking related to the basic physics concepts covered during lectures. The homework assignment was designed to tackle specific ideas and concepts, improving students' understanding and math skills. Additionally, MasteringPhysics provided hints to guide students to the correct answers. Students could also discuss their mistakes with classmates or the instructor during face-to-face office hours.

The following points outline the steps taken to prepare students for the quizzes:

- 1) After each chapter, 5-6 days were given for students to complete an online homework assignment.
- 2) After the homework assignment is done, the quiz is presented online using students' laptops in the class for 15 mins.
- 3) The quiz questions were randomized.

The course of introductory physics contains ten chapters, and nine quizzes were conducted during the semester covering the ten chapters of the course. Each quiz covers one or two chapters.

Using the GM (1,1) model developed using MATLAB code, the grades of the fourth initial quizzes were used to predict grades of quiz 5 (Q5) till quiz 9 (Q9). Table 1 presents the experimental grades of the nine quizzes during the Fall 2022 semester in General Physics I.

Table 2 presents the grades of the experimental fourth initial quizzes and the predicted grades of quiz 5 (Q5) till quiz 9 (Q9). After comparing the predicted scores with the experimental grades of the 40 students, the predicted results show that the maximum value of MAPA for the GM (1,1) model equal to 91%.

Table 1. Grades of 40 students in General Physics I during the Fall 2022 semester

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
S1	0	4.5	0	9.5	5	0	5	0	20
S2	10	4.5	15	15	0	0	0	5	15
S3	0	5	20	20	20	15	15	20	20
S4	10	14.5	10	15	5	15	10	15	20
S5	5	4.5	5	10	10	5	5	5	15
S6	5	5	5	9.5	0	0	10	5	0
S7	4	10	0	12	15	5	13	5	10
S8	5	15	5	10	5	5	10	10	20
S9	5	10	0	9.5	5	5	15	5	15
S10	9	20	4.5	14.5	15	10	15	5	15
S11	8	5	15	9.5	5	5	5	9.5	19.5
S12	0	0	10	15	5	10	10	0	20
S13	4	5	5	19.5	20	20	20	5	20
S14	18	15	5	14.5	10	0	15	10	15
S15	10	15	10	15	15	10	20	20	20
S16	9	0	5	15	5	0	5	0	19.5
S17	10	10	5	0	5	0	0	5	15
S18	10	10	10	20	20	5	20	20	20
S19	10	20	0	5	5	5	0	5	15
S20	5	5	10	14.5	5	5	5	5	19.5
S21	5	10	9.5	15	15	15	10	15	15
S22	9	20	10	15	15	15	10	20	15
S23	14	20	5	20	15	20	15	20	20

Published by European Centre for Research Training and Development-UK

S24	15	0	15	10	0	0	5	5	20
S25	10	15	15	20	15	0	20	19.5	20
S26	8	19.5	0	20	19	0	5	20	20
S27	5	19.5	15	15	20	15	0	0	20
S28	14	5	5	5	10	10	15	5	15
S29	5	5	20	10	0	10	10	15	15
S30	13	4.5	10	10	10	0	0	10	14.5
S31	10	0	5	9.5	5	5	5	5	14.5
S32	5	10	0	15	5	15	0	15	20
S33	20	15	20	20	17.5	20	20	20	20
S34	20	15	10	10	17.5	15	20	20	20
S35	0	5	5	5	5	0	5	0	15
S36	5	15	20	20	20	15	0	0	0
S37	13	5	5	14.5	5	0	5	0	20
S38	5	5	0	14.5	0	0	5	10	0
S39	4	10	5	5	4	0	0	0	9.5
S40	4	5	5	10	5	5	10	0	0

Many researchers have noted that the predictive accuracy of the grey model is often unsatisfactory (Li et al., 2007; Wen & Chang, 2005). Specifically, the coefficients used in the GM (1,1) prediction model are not optimal, leading to lower accuracy. To address this issue, research has focused on several areas: altering raw data sequences to enhance their smoothness; improving methods for calculating the parameters of GM(1,1) and optimizing the conformation of model background values; optimizing the initial values of GM(1,1); correcting residual errors of the model; developing and proposing expanded models of GM(1,1); examining the conditions for modeling; and creating combined prediction models using other technologies (Zhou, 2013).

Table 2. The prediction results using GM (1,1) of quizzes 5 to 9

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
S1	0	4.5	0	9.5	6	1	4	0.5	19
S2	10	4.5	15	15	1	0	0.5	4.5	16
S3	0	5	20	20	20	16	14.5	19	19.5
S4	10	14.5	10	15	5.5	16	10.5	16	19
S5	5	4.5	5	10	10.5	5.5	5.5	8	15.5
S6	5	5	5	9.5	0.5	0.5	11	4.5	2
S7	4	10	0	12	16	6	12.5	5.5	10.5
S8	5	15	5	10	6	5.5	10.5	11	18
S9	5	10	0	9.5	5.5	5.5	15.5	7	14
S10	9	20	4.5	14.5	14.5	11	15.5	6	14.5
S11	8	5	15	9.5	5.5	6	6.5	9	17
S12	0	0	10	15	5.5	11	12	4	17
S13	4	5	5	19.5	19	19.5	20	8	18.5
S14	18	15	5	14.5	10.5	3	14	11	14.5
S15	10	15	10	15	15.5	10.5	18	19	19.5
S16	9	0	5	15	5.5	3	6	3	17
S17	10	10	5	0	5	3	2	4	10
S18	10	10	10	20	20	7	17	18	19.5
S19	10	20	0	5	5.5	6	3	5	13
S20	5	5	10	14.5	5.5	5.5	6	6.5	17
S21	5	10	9.5	15	14	15	10.5	14.5	15.5
S22	9	20	10	15	15.5	15.5	12.5	18	16
S23	14	20	5	20	15.5	19	16	19.5	20
S24	15	0	15	10	5	0.5	6	6.5	17
S25	10	15	15	20	15.5	4	17	19	20
S26	8	19.5	0	20	19.5	5	6.5	17	19
S27	5	19.5	15	15	19	116	4	5	17
S28	14	5	5	5	10.5	10.5	16	6	14

Published by European Centre for Research Training and Development-UK

S29	5	5	20	10	0.5	11	11	14.5	16
S30	13	4.5	10	10	10.5	3	4	9	15
S31	10	0	5	9.5	6	5.5	6	7	13
S32	5	10	0	15	5.5	14	5	13	17
S33	20	15	20	20	18	19	20	19.5	20
S34	20	15	10	10	17	16	18.5	19.5	20
S35	0	5	5	5	5	0.5	4.5	1	10
S36	5	15	20	20	19.5	16	5	3	0.5
S37	13	5	5	14.5	6.5	2	5.5	0.5	15
S38	5	5	0	14.5	0.5	0.5	6	11	3
S39	4	10	5	5	4.5	0.5	0.5	0	7
S40	4	5	5	10	5.5	5	9.5	2	0.5

Figure 1 shows the experimental grade distribution for the nine quizzes and the predicted grades for student S10 during the semester. The Mean Absolute Percentage Error (MAPE) is 1.4% and the Mean Absolute Percentage Accuracy (MAPA) is 91%. To improve this method, a combination of GM (1,1) and the Taylor approximation method can be proposed as an alternative. The Taylor approximation method, which combines Taylor series expansion with the least squares method, is an iterative calculation method used to obtain the optimal parameters, minimizing the convergence error. This prediction model can be adjusted repeatedly until it reaches the optimal values, thereby minimizing the prediction error.

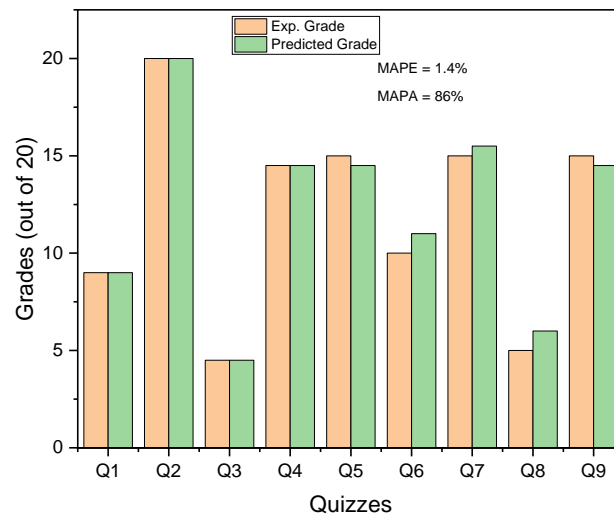


Fig. 1. Experimental grades distribution for the nine quizzes and the predicted grades for student S10 during the Fall 2022 semester.

CONCLUSION

In conclusion, this prediction model is highly useful for addressing uncertainties in systems with limited data, where traditional statistical methods fall short. The development of a MATLAB program enhances the speed and accuracy of data processing. Additionally, this model can predict student assignment evaluations, aiding in improving student performance across different courses. The GM(1,1) model specifically contributes to this improvement. However, to address the accuracy limitations of GM(1,1), a combination of GM(1,1) and the Taylor approximation method is proposed as a viable alternative.

ACKNOWLEDGEMENTS

The authors acknowledge the support by UAE University under grant numbers 12R232, and 31S464.

REFERENCES

- Askari, M., & Askari, H. (2011). Time series grey system prediction-based models: Gold price forecasting. *Trends in Applied Sciences Research*, 6(11), 1287.
- de Gooijer, J. G., & Klein, A. (1989). Forecasting the Antwerp maritime steel traffic flow: a case study. *Journal of Forecasting*, 8(4), 381-398.
- Deng, J. L. (2003). Literalizing GRA axioms. *J Grey Syst*, 4, 399-400.

- Hongli, W. (2013). Simulation and control of system dynamic of water pollutions based on modeling of differential equations using inverse GM. *International Journal of Control and Automation*, 6(4), 305-320.
- Julong, D. (1989). Introduction to grey system theory. *The Journal of grey system*, 1(1), 1-24.
- Julong, D. (2004). Grey Management: Grey Situation Decision Making in Management Sciences. *Journal of Grey System*, 16(2).
- Ken, M. L., Chen, W. C., & Lee, Y. B. (2010). Grey prediction in the study of tourist's number. *Journal of Grey System*, 13(4), 139-144.
- Li, G.-D., Yamaguchi, D., Mizutani, K., & Nagai, M. (2007). New proposal and accuracy evaluation of grey prediction GM. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 90(6), 1188-1197.
- Wei, C. H., & Yang, Y. C. (1999). A study on transit containers forecast in Kaohsiung port: applying artificial neural networks to evaluating input variables. *Journal of the Chinese Institute Transportation*, 11(3), 1-20.
- Wen, K.-L., & Chang, T.-C. (2005). The research and development of completed GM (1, 1) model toolbox using Matlab. *International Journal of Computational Cognition*, 3(3).
- Yen, J. R., & Lin, J. S. (1997). The research of ROC liner freight. *Journal of the Chinese Institute Transportation*, 10(4), 97-112.
- Zhang, C. P., Zhou, Q. Q., & Nie, J. (2012). The prediction of China CO2 emission in 2015. *International Journal of Energy Science*, 2(2), 47-50.
- Zhou, D. (2013). A new hybrid grey neural network based on grey verhulst model and BP neural network for time series forecasting. *International Journal of Information Technology and Computer Science*, 5(10), 114-120.