Journal of Natural Science Review

Vol. 3, No. 4, 2025
https://kujnsr.com
e-ISSN: 3006-7804

A Comparative Study of the GM (1,1) Model and Curve Fitting Method for Forecasting Viral Hepatitis Incidence in Afghanistan up to 2030

Raz Mohammad Mohammadi✉[1], Mohammad Farooq Hakimi[2], Abdul Raqib Muslimyar[3], Wali Mohammad Azizi[4]

[1,4] Department of Mathematics, Education Faculty, Kandahar University, Kandahar, Afghanistan
[2,3] Department of Mathematics, Natural Science Faculty, Kabul Education University, Kabul, Afghanistan
✉E-mail: sodes.amin123@gmail.com (corresponding author)

ABSTRACT

This study aimed to forecast the incidence of viral hepatitis in Afghanistan using the GM(1,1) model and curve fitting methods, and to compare the predictive performance of both approaches using mean absolute error (MAE) and mean absolute percentage error (MAPE). Annual incidence data were obtained from the Hospital of Infectious Diseases through a formal request. Linear and nonlinear regression techniques, along with the first-order univariate grey prediction model (GM(1,1)), were applied to model historical trends. The model with superior predictive accuracy was used to project viral hepatitis incidence for 2015–2030. Both GM(1,1) and curve-fitting models accurately captured the incidence trends; however, GM(1,1) demonstrated superior performance (MAE = 557.95; MAPE = 6.76%) compared with exponential curve regression (MAE = 558.30; MAPE = 7.07%). Forecasts indicated 15,328.6 cases in 2025 and 39058 cases in 2030, with both models projecting a consistent upward trend, reflecting a growing public health burden. The projections highlight a growing public health burden of viral hepatitis in Afghanistan, emphasizing the urgency of effective prevention and vaccination programs. The findings can guide policymakers in resource allocation and healthcare planning, while also informing strategies to strengthen surveillance and early detection. Moreover, the study demonstrates that the GM(1,1) model is a reliable forecasting tool in contexts with limited or incomplete data, providing valuable support for evidence-based decision-making in public health. This study is the first to compare GM(1,1) and curve fitting for forecasting viral hepatitis in Afghanistan, using viral hepatitis data from Afghanistan. It provides context-specific projections through 2030 and demonstrates that GM(1,1) is a reliable tool in data-limited settings.

ARTICLE INFO

Article history:

Received: September27, 2025
Revised: October 21, 2025
Accepted: November 26, 2025
Published: December 31, 2025

Keywords:

Afghanistan; Curve Fitting Method; GM (1,1) model, Incidence; prediction; Viral Hepatitis

Cite to this article: Mohammadi, R. M., Hakimi, M. F., Muslimyar, A. R., & Azizi, W. M. (2025). A comparative study of the GM (1,1) model and curve fitting method for forecasting viral hepatitis incidence in Afghanistan up to 2030. *Journal of Natural Science Review, 3*(4), 47-62 https://doi.org/10.62810/jnsr.v3i4.324

Link to this article: https://kujnsr.com/JNSR/article/view/324

Copyright © 2025 Author(s). This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

INTRODUCTION

Viral hepatitis is a liver disease caused by infection with several types of viruses. These pathogens can lead to both acute and chronic hepatitis, and in some cases, may progress to hepatocellular carcinoma (HCC) (Razavi, 2020). The hepatitis viruses—A, B, C, E, and G—each present unique clinical profiles, posing different challenges for both patients and healthcare professionals. Fundamental research into the immune response to viral infections, along with

insights from virology, has enabled the development of vaccines for hepatitis A, B, and E, as well as highly effective antiviral treatments for hepatitis B and C (Odenwald & Paul, 2022).

Hepatitis is the second-deadliest communicable disease in 2022, after COVID-19. Also, 254 million people are living with hepatitis B, and 50 million are people living with hepatitis C worldwide. hepatitis related deaths increased from 1.1 to 1.3 million between 2019 and 2022, and every day, 3500 people are dying from viral hepatitis. Every day, approximately 3,500 individuals die due to viral hepatitis, with hepatitis B accounting for 83% and hepatitis C for 17% of these fatalities (WHO, 2024). In Afghanistan, a significant obstacle in addressing viral hepatitis is the lack of comprehensive and reliable data. A 2019 study conducted at the Kabul Infectious Disease Hospital identified 3,945 cases of viral hepatitis. Among the different types, hepatitis B (HBV) was found to be the most widespread in the country (Essar et al., 2023).

One study used eleven curve estimation models to predict the trends of index transmissibility (R_0,MNI) for Hepatitis C in Xiamen city of China (Y. Wang et al., 2020). Several studies have employed different forecasting models to predict disease trends. For instance, GM (1,1), ARIMA, and curve-fitting methods were utilized to estimate the future number of patients with chronic obstructive pulmonary disease (COPD) in China (Teng et al., 2023a). In India, Long Short-Term Memory (LSTM) networks and curve-fitting techniques were applied to forecast COVID-19 case counts (Tomar & Gupta, 2020). Another study simulated the incidence trend of hepatitis C in Changsha, China, using curve fitting, the grey model GM (1,1), the ARIMA model, and the Back Propagation Neural Network (BPNN) (Chen et al., 2025). In Gombak, to assess the spread of tuberculosis (TB), researchers used multiple linear regression (MLR) and Artificial Neural Network (ANN) models (Mohidem et al., 2021). Additionally, in a study forecasting the incidence trend of pneumoconiosis in China, three models were used: the GM(1,1) grey model, the Generalized Additive Model (GAM), and the Curve Fitting Method. The GM(1,1) model was applied to predict future trends when data were limited or uncertain. The GAM and Curve Fitting Method were used to capture more complex and nonlinear relationships in the incidence data. All models showed a decreasing trend in pneumoconiosis. However, the GAM and Curve Fitting Method achieved higher accuracy, highlighting the importance of choosing models based on data characteristics and practical needs(Zhou et al., 2021). In a study, the paper proposes a novel grey prediction model to improve forecasting accuracy when different time series data patterns require different weighting of new and old information. It is applied to predict CO_2 emissions of the BRICS countries (Brazil, Russia, India, China, South Africa)(Guo et al., 2021). At the United Arab Emirates University, the grey system GM (1,1) model was employed to forecast students' quiz and homework performance in General Physics I over the course of a semester (Najar et al., 2024). The forecasting approach utilized a first-order, two-variable grey differential equation model based on grey system theory. To improve prediction accuracy, the GM(1,1) model was implemented in MATLAB (Najar et al., 2024). In another research mention that the GM (1,1) model has usage in economics, computer knowledge, and others

and GM (1,1) model were used to predict seasonal time series (Kartikasari & Hikmah, 2022) And in China Health statistic Yearbook by AIDS incidence data, they developed both canonical GM (1,1) model and an enhanced version for prediction of AIDS incidence in China the prediction exactness for both standard and advanced GM (1,1) models were evaluated as Level 1, indicating best performance (Zhao et al., 2021).

Despite this extensive global research, forecasting studies on viral hepatitis in Afghanistan remain scarce. No previous research has systematically applied or compared forecasting models for hepatitis incidence in the country. Given Afghanistan's limited health data infrastructure, identifying models that perform reliably with sparse datasets is particularly important.

In this study, the GM model and Curve Fitting Methods are employed to forecast the incidence of viral hepatitis in Afghanistan between 2025 and 2030 and to assess the accuracy and predictive performance of these two models. The purpose of this research is to find suitable models for predicting the spread of viral hepatitis in Afghanistan. And to evaluate the anticipated next dangers and load of hepatitis, in addition to inform evidence- based strategies for its prevention and control.

METHODS AND MATERIALS

In a formal request letter, the Hospital of Infectious Diseases provided annual incidence data on viral hepatitis in Afghanistan. The dataset provided the empirical foundation for model building, validation, and forecasting, and it included reported instances from 2015 to 2024. The dataset comprises officially reported cases and might not include undetected infections due to Afghanistan's inadequate monitoring infrastructure, but it offers the most trustworthy longitudinal dataset currently accessible.

Descriptive Statistical Analysis

MATLAB and SPSS (Version 27) were used for preliminary data exploration and visualization to evaluate temporal patterns, identify anomalies, and understand the distributional features of hepatitis incidence. Annual growth rates, trend visualization, and residual inspection were examples of descriptive metrics.

Modeling Approaches

Two forecasting frameworks were employed: the Grey Prediction Model GM(1,1) and Curve Fitting Regression Models (linear and nonlinear).

FINDINGS

Table 1*. Yearly data and predicted incidence number of viral hepatitis in Afghanistan from 2015 to 2024, residual, relative error: results of GM (1,1) model*

GM (1,1) Model for Viral Hepatitis Incidence in Afghanistan (2015-2024)					
Development coefficient (a) = -0.1871					
Grey input (b) = 2117.0709					
Years	Actual data	Predicted value	Residual	|Residual|	Relative Error
2015	2521	2521	0	0	0.00%
2016	2686	2846.6	-160.6	160.6	5.98%
2017	3211	3432.2	-221.2	221.2	6.89%
2018	3866	4138.2	-272.2	272.2	7.04%
2019	4991	4989.5	1.5	1.5	0.03%
2020	6315	6015.8	299.2	299.2	4.74%
2021	6900	7253.3	-353.3	353.3	5.12%
2022	8075	8745.4	-670.4	670.4	8.30%
2023	12777	10544.4	2232.6	2232.6	17.47%
2024	11345	12713.4	-1368.4	1368.4	12.06%
Model Accuracy Metrics:					
MAE = 557.95					
MAPE = 6.76%					
Posterior Variance Ratio C = 0.2573					
Small Error Probability P = 1.0000					

Table 2: Approximate R-squared, F and P values, and constant in five types of curve regression

Model	**R-square**	**F-value**	**p-value**	**constant**
First-degree polynomial	0.897	69.342	<.001	108.733
Second-degree polynomial	0.939	53.522	<.001	2218.233
Third-degree polynomial	0.941	32.032	<.001	3065.700
Exponential	0.975	307.175	<.001	1921.808
S-shaped	0.591	11.548	.009	9.064

Table 3*. Yearly data and fitted incidence of viral hepatitis in Afghanistan from 2015 to 2030, along with the difference between actual and predicted values, involving the results and MAE, MAPE for both the GM(1,1) model and the exponential curve regression model*

		GM (1,1) Model		**Exponential Curve Regression Model**	
Years	**Actual Cases**	**Fitted/Predicted* Values**	**Error**	**Fitted/Predicted* Values**	**Error**
2015	2521	2521	0	2319.6	201.4
2016	2686	2846.6	-160.6	2799.6	-113.6
2017	3211	3432.2	-221.2	3379.1	-168.1
2018	3866	4138.2	-272.2	4078.4	-212.4
2019	4991	4989.5	1.5	4922.5	68.5
2020	6315	6015.8	299.2	5941.3	373.7
2021	6900	7253.3	-353.3	7171.0	-271.0
2022	8075	8745.4	-670.4	8655.2	-580.2
2023	12777	10544.4	2232.6	10446.5	2330.5
2024	11345	12713.4	-1368.4	12608.6	-1263.6
2025	--	**15328.6**	--	**15168**	--

2026	--	**18481.8**	--	**18251**	--
2027	--	**22284**	--	**21961**	--
2028	--	**26868**	--	**26425**	--
2029	--	**32394**	--	**31796**	--
2030		**39058**		**38259**	
MAE	--	557,95	--	558.30	--
MAPE (%)	--	6.76%	--	7.07%	--

Note: **Predicted values: The bold numbers correspond to the predicted values from each model.*

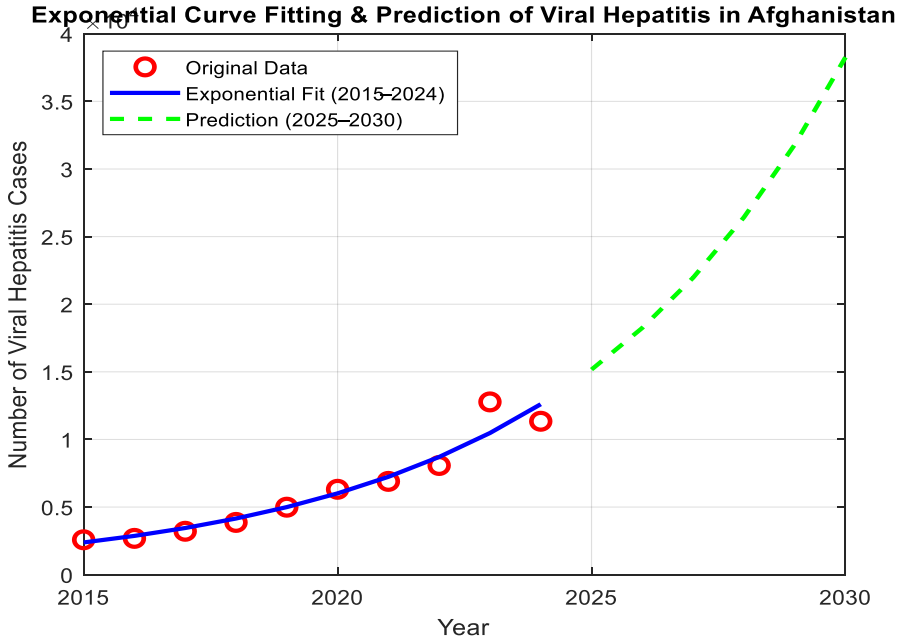

Figure 1. *Number of incidence with Viral Hepatitis in Afghanistan from 2015 to 2030*

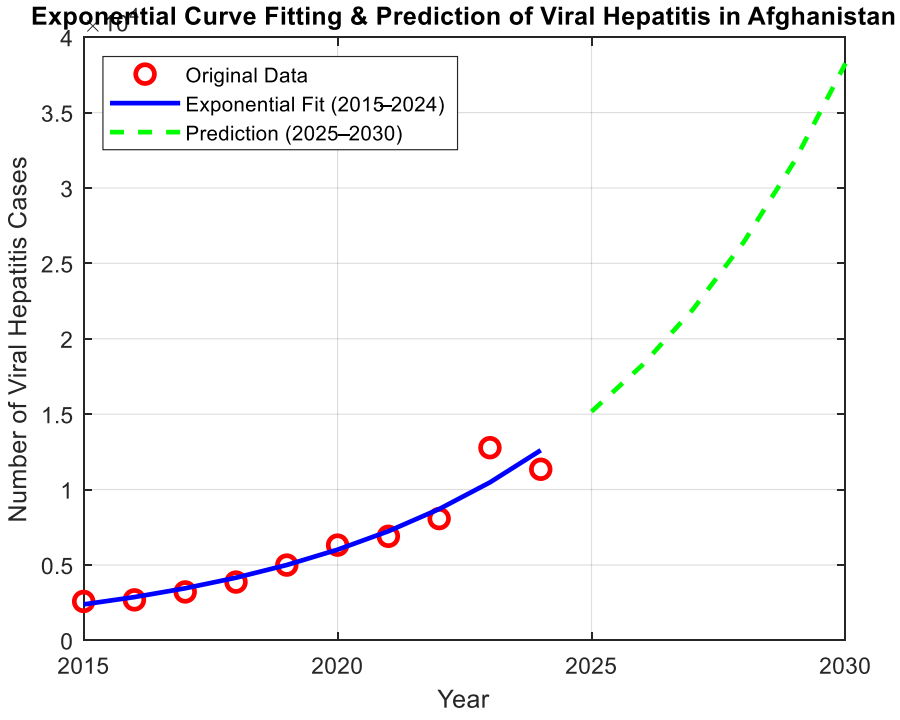

Figure 2. *Number of incidence with Viral Hepatitis in Afghanistan from 2015 to 2030*

The annual reported cases of viral hepatitis in Afghanistan from 2015 to 2024 are shown in Table 3, together with the fitted values from the exponential curve regression model and the GM(1,1) model. The table allows direct comparison of model performance by displaying the difference (error) between actual and predicted values for each year. Although the GM(1,1) model often yields fewer errors in most years, both models track the rising trend in hepatitis incidence. While the exponential model also works very well but exhibits slightly higher variances, the GM(1,1) model matches the early years (2015–2019) exceptionally well, with minor deviations. For all models, there are greater disparities in 2023 and 2024, when real cases increased dramatically, leading to bigger residuals. Forecasts for 2024 through 2030 are included in the table, and both models indicate that hepatitis incidence will continue to rise. By 2030, the exponential model forecasts 38,259 cases, while the GM(1,1) model forecasts 39,058 cases. GM(1,1) (MAE = 557.95; MAPE = 6.76%) outperforms the exponential model (MAE = 558.30; MAPE = 7.07%), as indicated by the model accuracy indicators at the bottom. Overall, the table shows that both models are trustworthy; nevertheless, GM(1,1) offers somewhat better forecast accuracy for the hepatitis trend in Afghanistan.

Grey Model $(\boldsymbol{GM})$

conditions, for example, uncertainty, multi-data input, discrete data, and the poverty of the data and analysis model the $GM\ (1.1)$ model is the simplest and most commonly used model in grey theory. While it became popular in the 1980s. The model has been widely used since its promotion (Deng, 1982). Unlimited research has been done in real-world implementation of the $GM\ (1.1)$ forecasting model in the last 40 years, and this research continues up to now. One kind of GM model is the grey model, which is most active and has extensive usage. Dang and Liu (2004) proposed the first-order univariate grey model $GM\ (1.1)$ model using $x(n)$ as the initial value. Later, in 2014, Liu and colleagues developed four fundamental variations of $GM\ (1.1)$ model. namely EGM, DGM, EDGM, and ODGM, by conducting simulated examinations. Furthermore, different types of forms of $Gm\ (1.1)$ power model were found by Wang (2013). Moreover, he studied their time-response function properties (Liu, 2025). The cumulative generation process plays a role in "strengthening" the randomness of the numbers defined by (Madhi & Mohamed, 2022). An article introduces a generalized $GM\ (1.1)$ model along with its modelling approach and develops a grey prediction model to address real-world problem (M. Cheng & Xiang, 2017). There are three types of grey prediction models: white, grey, and black systems. The principal purpose of grey systems is to define the relations among conditions, for example, variability in multi-input datasets, discrete data, data sparseness, and analysis model structure. The famous is theis $GM\ (1.1)$ model. Grey prediction uses a limited dataset to approximate the behavior of a known system; grey models predict the future values of time series. Furthermore, assumed that all values from the data used in grey system analysis must be positive and that the time-series frequency is invariant (Lotfalipour et al., 2013).

The Four Fundamental states of GM (1,1)

Four fundamental states of the GM (1,1) model are: Even Grey Model (EGM), Original Difference Grey Model (ODGM), Even Difference Grey Model (EDGM), and Discrete Grey Model (DGM).

Definition Let $X^{(0)}=\left(x^{(0)}(1),x^{(0)}(2),x^{(0)}(3),...,x^{(0)}(n)\right),x^{(0)}(k)\geq 0$ and $X^{(1)}$ be the accumulation generating operation sequence of $X^{(0)}$; that is.

$$X^{(1)}=\left[x^{(1)}(1),x^{(1)}(2),x^{(1)}(3),...,x^{(1)}(n)\right] \quad (1)$$

where $x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i),k=1,2,...,n$ Then

$$x^{(0)}(k)+ax^{(1)}(k)=b \quad (2)$$

Referred to as the original form of model GM (1,1), which is a difference equation. By using the least squares method, we can find an approximate value of the vector that holds the equality.

$$\hat{a}=\left(B^{T}B\right)^{-1}B^{T}Y \quad (3)$$

Where

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

Definition Let $X^{(0)},X^{(1)}$ and, just like definition, let.

$$Z^{(1)}=\left(Z^{(1)}(2),Z^{(1)}(3),...,Z^{(1)}(n)\right), \quad (4)$$

Where $Z^{(1)}(k)=\frac{1}{2}\left(x^{1}(k)+x^{(1)}(k-1)\right),$ then

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

In this case, matrix B is:

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

Curve Fitting Method

Curve-fitting methods, which involve regression, are used to analyze and model relationships between variables. Regression is a statistical technique for examining and modeling these relationships. (Kowsher et al., 2020). Regression plays a vital role in prediction and curve fitting. Both linear and nonlinear regression are widely used across many fields, including the natural sciences and psychology. Furthermore, their parameters have a linear relationship. For instance, polynomial regression models a nonlinear relationship between the response variable and predictors by incorporating higher-degree terms of the predictors. However, this approach appears complex in terms of parameter estimation. Nonlinear regression offers interpretability, miserliness, and predictive power, and nonlinear models are suitable for use, miserly, and explainable because the parameters can be associated with interpretable and meaningful factors (Huang & He, 2024).

Linear Regression

One of the simplest applications of least-squares approximation involves fitting a linear model to paired observational data. The equation of the straight line is:

$$y = a_0 + a_1 x + e \tag{6}$$

The values of a_0 and a_1 are

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^{\,2} - \left(\sum x_i\right)^2}$$
$$a_0 = \bar{y} - a_1 \bar{x} \tag{7}$$

Polynomial Regression

Although specific engineering datasets exhibit clear patterns, they are not adequately captured by linear models; in such cases, a curved model provides a more appropriate representation. One possible option for fitting the data is a polynomial function. This is called polynomial regression, and we can extend the least-squares method to fit higher-order polynomials to the data. Assume a Quadratic polynomial we want to fit:

$$y = a_0 + a_1 x + a_2 x^2 + e \tag{8}$$

The total of the squared residuals represents

$$s_r = \sum \left(y_i - a_0 - a_1 x_i - a_2 x_i \right)^2 \tag{9}$$

Taking derivative with respect to unknown coefficients:

$$s_r = \sum \left(y_i - a_0 - a_1 x_i - a_2 x_i \right)^2 \tag{10}$$

After derivation with respect to a_0. a_1 and a_2 The following system will be archived.

$$\begin{aligned} & n a_0 + \left(\sum x_i \right) a_1 + \left(\sum x_i^2 \right) a_2 = \sum y_i \\ & \left(\sum x_i \right) a_0 + \left(\sum x_i^2 \right) a_1 + \left(\sum x_i^3 \right) a_2 = \sum x_i y_i \\ & \left(\sum x_i^2 \right) a_0 + \left(\sum x_i^3 \right) a_1 + \left(\sum x_i^4 \right) a_2 == \sum x_i^2 y_i \end{aligned} \tag{11}$$

This analysis can be easily generalized to a higher-order polynomial.

Model Evaluation

Model evaluation metrics serve as essential tools for objectively and intuitively assessing the performance of prediction models. By comparing the values of different evaluation metrics across models, it is possible to analyses their predictive strengths and weaknesses systematically. A prediction model that consistently outperforms competitors across multiple metrics is generally considered more robust and reliable, particularly for forecasting sensitive areas such as disease trends.

In this study, three different prediction methods were evaluated and compared using widely accepted model evaluation indicators. The quality of model fitting was assessed through the coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE). These metrics collectively provide insights into the degree of correlation between observed and predicted values and the magnitude of prediction errors. Furthermore, the predictive validity of the models was assessed using the relative prediction error, which measures forecast accuracy on real-world data.

Overall, integrating multiple evaluation metrics not only provides a comprehensive understanding of model performance but also facilitates the selection of the most suitable prediction approach for practical decision-making and policy planning in disease monitoring and control (Teng et al., 2023b).

$$\begin{cases} R^2 = 1 - \dfrac{ss_r}{ss_t} \\ ss_r = \sum (y_i - f_i)^2 \\ ss_t = \sum (y_i - \bar{y})^2 \\ \bar{y} = \dfrac{1}{n}\sum y_i \end{cases} \quad (11)$$

Here, ss_r represents the sum of squared residuals, while ss_t Denotes the total sum of squares. The symbol y_i Refers to the observed (actual) value, fi indicates the predicted (fitted) value, and $\bar{y}$ Stands for the mean of the observed values. The coefficient of determination (R^2) ranges from 0 to 1, with values approaching 1 indicating a stronger goodness-of-fit for the model.

DISCUSSION

In relation to our results, both the GM (1,1) model and curve-fitting methods can be used to forecast next year's viral hepatitis incidence in Afghanistan. Each model performs well in terms of fitting and prediction. The GM (1,1) model showed better prediction accuracy than curve-fitting methods.

In a study forecasting the incidence trend of pneumoconiosis in China, three models were used: the GM(1,1) grey model, the Generalized Additive Model (GAM), and the Curve Fitting Method. The GM(1,1) model was applied to predict future trends when data were limited or uncertain. The GAM and Curve Fitting Method were used to capture more complex and nonlinear relationships in the incidence data. All models showed a decreasing trend in pneumoconiosis. However, the GAM and Curve Fitting Method achieved higher accuracy, highlighting the importance of choosing models based on data characteristics and practical needs (Zhou et al., 2021). In another comparative study of GM (1,1) and ARIMA models for hepatitis B forecasting, the MAE, MAPE, and RMSE metrics from the ARIMA model were not as high as those from the GM (1,1) model, both for fitting and forecasting. So the ARIMA (3,1,1) (0,1,2)12 model can better predict than GM (1,1) (Y. Wang et al., 2018). In another research paper, three prediction models were used to predict the prevalence of COPD in China: the GM (1,1) model, the ARIMA (0,1,2) model, and curve fitting. Among these three models, the optimal was GM, the curve-fitting method was third, and the ARIMA(0,1,2) was between them (Teng et al., 2023a). In a 2019 study, GM(1,1), ARIMA, and negative binomial regression models were used to forecast hand-foot-mouth disease cases in Yiwu City, Zhejiang province, China, and all three models' predictions and actual values were in good agreement. The GM model had the best-fitting results (Fu et al., 2019). According to the number of diagnosed pneumoconiosis cases in China from 2000 to 2019, an investigation found three prediction models for the incidence trend of pneumoconiosis, including GAM, a curve-fitting method, and the GM(1,1) model. The result showed that, among the three

forecasting models, the Generalized Additive Model was optimal. It attained the highest R-squared value, the lowest average fitting relative error, and the smallest prediction relative error (Zhou et al., 2021b). In one research paper, the ARIMA model was applied to forecast the epidemic trends of several infectious and chronic non-communicable diseases, illustrating its strong predictive ability. For example, the ARIMA model was applied to estimate authentic cases and forecast the epidemic trajectory of COVID-19 from April 24 to July 7, 2020, across the 15 countries most affected by the pandemic in terms of confirmed cases, deaths, and recoveries (Singh et al., 2020). We can apply curve-fitting methods to forecast future COPD incidence. Among the three methods evaluated, it showed the poorest performance in terms of fitting efficiency, trustworthiness, and forecasting efficiency. This result aligns with separate research on COVID-19 prediction, and the curve-fitting methods had the lowest prediction accuracy when comparing various algorithms among the techniques assessed (Achterberg et al., 2022).

Computation of curve-fitting methods is simple, but they have limited applicability to curve types. This makes it challenging to ensure the chosen curve yields the best maximum fit. We observed that the exponential curve provided the best fit; however, it performed poorly on test data (Hu, 2017). The GM (1,1) is suitable for sparse or minimal datasets and can achieve strong predictive performance, especially when the data exhibit specific trends (Yao et al., 2023). It is widely applied in many areas, including population growth (L. Y. Wang, 2014). It is widely used for predicting population (Y. Cheng et al., 2014). Has application in real estate prices, logistics, and related fields (Li & Lu, 2015). A study finds that the GM(1,1) model has strong predictive ability for the incidence trends of typhoid and paratyphoid fever cases in Wuhan City (Yang et al., 2018).

This study has several limitations. First, the analysis relied on reported incidence data, which in Afghanistan are subject to underreporting and incomplete surveillance, potentially affecting forecast accuracy. Second, the models used were based solely on historical case numbers and did not incorporate socioeconomic, behavioural, or environmental factors that may influence transmission dynamics. Third, the forecasting period was limited to 2025–2030, and longer-term predictions may require more sophisticated models or hybrid approaches. Despite these limitations, the study provides an important baseline for understanding hepatitis trends in Afghanistan and demonstrates the applicability of GM(1,1) in data-limited settings.

China Viral hepatitis poses a significant threat to public health, and ranks among the leading causes of death globally (Usuda et al., 2024). Viral hepatitis imposes a fundamental epidemiological and economic burden, implicating accurate forecasting to inform public health strategies. This study used two modeling methods, specifically the curve-fitting method and the GM(1,1) model, to model the reported incidence rate of viral hepatitis. A comparative analysis selected the curve-fitting method and the GM(1,1) model, which achieves excellent simulation accuracy for incidence trends in Afghanistan. Furthermore, GM's prediction was better than the exponential ones. Forecasting models are statistical or

computational tools that estimate future outcomes based on input data. Therefore, selecting an appropriate model for a specific disease is important for accurate

CONCLUSION

To predict future viral hepatitis incidence in Afghanistan, this study used and contrasted two forecasting techniques: curve-fitting regression methods and the GM(1,1) gray prediction model. The increasing trend in reported cases from 2015 to 2024 was well represented by both models; however, the GM(1,1) model showed better predictive accuracy, as evidenced by its superior posterior error characteristics and lower MAE and MAPE values. This demonstrates that grey system theory is appropriate for health forecasting in settings with limited data, such as Afghanistan.

Both models' predictions show that the incidence of hepatitis will continue to climb significantly between 2025 and 2030. According to the GM(1,1) model, there would be more than 38,000 cases by 2030, indicating an increasing epidemiological burden. This anticipated increase underscores the need to bolster national hepatitis control initiatives, including increased vaccine coverage, early detection, community-based awareness campaigns, and enhanced clinical management.

The work reveals a more general methodological discovery that goes beyond generating quantitative forecasts: gray models are instrumental in low-resource settings because they can produce accurate predictions even when health surveillance data is lacking. Policymakers, public health experts, and epidemiologists looking for computationally efficient methods for planning and resource allocation can thus benefit from the findings.

The study admits certain limitations despite its advantages. Due to underreporting, reported incidence data may not accurately reflect the actual frequency of viral hepatitis, and the employed prediction models did not account for environmental, behavioral, or socioeconomic factors. To improve long-term prediction accuracy, future research should consider hybrid forecasting models, such as GM-ARIMA combinations, machine learning techniques, and multivariate grey models.

Finally, the study shows that the GM(1,1) model is a reliable and valuable tool for predicting future illness trends and gives the first comparative forecasting analysis of hepatitis incidence in Afghanistan. The findings highlight the need for prompt legislative action and well-informed public health initiatives to reduce the rising incidence of viral hepatitis over the next 10 years.

AUTHORS CONTRIBUTIONS

- Raz Mohammad Mohammadi (R.M.M.) conceptualized the study, performed data analysis, and applied the forecasting models.
- Mohammad Farooq Hakimi (M.F.H) acquired the data, contributed to methodology design, and interpreted the results.

- Abdul Raqib Muslimyar (A.R.M.) conceptualized the study, drafted the manuscript, conducted the literature review, and critically revised the work.
- Wali Mohammad Azizi (W.M.A) conceptualized the study, drafted the manuscript, conducted the literature review, and critically revised the work.
- All authors reviewed and approved the final version of the manuscript.

ACKNOWLEDGMENT

We sincerely thank all healthcare professionals for their ongoing dedication and hard work in combating viral hepatitis.

FUNDING INFORMATION

No funding is available for the manuscript.

DATA AVAILABILITY STATEMENT

The viral hepatitis incidence data used in this study were obtained from the Hospital of Infectious Diseases via an official request letter. These data are not publicly available due to institutional confidentiality agreements, but may be made available upon reasonable request and with the Hospital of Infectious Diseases' permission. Data analysis and modeling were conducted using MATLAB and SPSS 25.0.

CONFLICT OF INTEREST STATEMENT

The authors report no conflicts of interest in this work.

REFERENCES

Achterberg, M. A., Prasse, B., Ma, L., Trajanovski, S., Kitsak, M., & Van Mieghem, P. (2022). Comparing the accuracy of several network-based COVID-19 prediction algorithms. *International Journal of Forecasting*, *38(2), 489–504*. https://doi.org/10.1016/j.ijforecast.2020.10.001

Chen, S., Zeng, P., Yang, K., Zhang, X., Yang, X., & Zhou, Y. (2025). Trends in the hepatitis C incidence rate in Changsha, China, 2005 to 2023, and projections to 2025. **The Journal of Infection in Developing Countries*, *19(1), 140–145.* https://doi.org/10.1016/j.ijforecast.2020.10.001

Cheng, M., & Xiang, M. (2017). Generalized GM(1,1) model and its application. *Journal of Grey System*, *29(3).*

Cheng, Y., Yun, S., Xu, X., & Lv, P. (2014). A study on teaching and application of grey prediction model—Taking commercial housing prices of Huangshi City as an example. *Energy Education Science and Technology Part A: Energy Science and Research*, *32(3), 1821–1828*. https://doi.org/10.3855/jidc.19790

Essar, M. Y., Siddiqui, A., & Head, M. G. (2023). Infectious diseases in Afghanistan:

Strategies for health system improvement. *Health Science Reports*, *6(12), e1775*. https://doi.org/10.1002/hsr2.1775

Fu, T., Chen, T., Dong, Z.-B., Luo, S.-Y., Miao, Z., Song, X.-P., Huang, R.-T., & Sun, J.-M. (2019). Development and comparison of forecast models of hand-foot-mouth disease with meteorological factors. *Scientific Reports*, *9(1), 15691*. https://doi.org/10.1038/s41598-019-52044-5

Guo, J., Liu, W., Tu, L., & Chen, Y. (2021). Forecasting carbon dioxide emissions in BRICS countries by exponential cumulative grey model. *Energy Reports, 7, 7238–7250*. https://doi.org/10.1016/j.egyr.2021.10.075

Hu, Y.-C. (2017). A genetic-algorithm-based remnant grey prediction model for energy demand forecasting. *PLoS ONE, 12(10), e0185478*. https://doi.org/10.1371/journal.pone.0185478

Huang, H.-H., & He, Q. (2024). Nonlinear regression analysis. *arXiv preprint arXiv:2402.05342.* https://doi.org/10.48550/arXiv.2402.05342

Kartikasari, M. D., & Hikmah, N. (2022). Decomposition method with application of grey model GM(1,1) for forecasting seasonal time series. *Pakistan Journal of Statistics and Operation Research*, 411–416. https://doi.org/10.18187/pjsor.v18i2.3533

Kowsher, M., Uddin, M. J., Moheuddin, M. M., & Turaba, M. Y. (2020). Two new regression and curve fitting techniques using numerical methods. In *Algorithms for Intelligent Systems.* Springer.: https://ssrn.com/abstract=3590089

Li, Y., & Lu, S. (2015). Research on prediction of regional grain logistics demand based on the grey-regression model. In *2015 IEEE International Conference on Grey Systems and Intelligent Services (GSIS)(pp. 97–101). IEEE*.

Liu, S. (2025). *Grey systems analysis: Methods, models and applications. * Springer Nature.

Lotfalipour, M. R., Falahi, M. A., & Bastam, M. (2013). Prediction of CO_2 emissions in Iran using grey and ARIMA models. *International Journal of Energy Economics and Policy, 3(3), 229–237.*

Madhi, M., & Mohamed, N. (2022). Improving GM(1,1) model performance accuracy based on the combination of optimized initial and background values in time series forecasting. *Open Access Library Journal, 9(4), 1–17*. https://doi.org/10.4236/oalib.1108416

Mohidem, N. A., Osman, M., Muharam, F. M., Elias, S. M., Shaharudin, R., & Hashim, Z. (2021). Prediction of tuberculosis cases based on sociodemographic and environmental factors in Gombak, Selangor, Malaysia: A comparative assessment of multiple linear regression and artificial neural network models. *The International Journal of Mycobacteriology, 10(4), 442–456*. DOI: 10.4103/ijmy.ijmy_182_21

Najar, A., Qamhieh, N., Amrane, N., & Mahmood, S. (2024). Grey system model GM(1,1) for predicting student performance. *British Journal of Multidisciplinary and Advanced Studies, 5(4), 8–18*. https://doi.org/10.37745/bjmas.2022.04153

Odenwald, M. A., & Paul, S. (2022). Viral hepatitis: Past, present, and future. *World Journal of Gastroenterology, 28*(14), 1405*. DOI: https://dx.doi.org/10.3748/wjg.v28.i14.1405

Razavi, H. (2020). Global epidemiology of viral hepatitis. *Gastroenterology Clinics of North America, 49(2), 179–189*. doi: 10.1016/j.gtc.2020.01.001.

Singh, R. K., Rani, M., Bhagavathula, A. S., Sah, R., Rodriguez-Morales, A. J., Kalita, H., Nanda, C., Sharma, S., Sharma, Y. D., Rabaan, A. A., et al. (2020). Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model. *JMIR Public Health and Surveillance, 6(2), e19115*. doi:10.2196/19115

Teng, Y., Jian, Y., Chen, X., Li, Y., Han, B., & Wang, L. (2023). Comparison of three prediction models for predicting chronic obstructive pulmonary disease in China. *International Journal of Chronic Obstructive Pulmonary Disease, 2961–2969.* https://doi.org/10.2147/COPD.S431115

Tomar, A., & Gupta, N. (2020). Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *Science of the Total Environment, 728, 138762*. https://doi.org/10.1016/j.scitotenv.2020.138762

Usuda, D., Kaneoka, Y., Ono, R., Kato, M., Sugawara, Y., Shimizu, R., Inami, T., Nakajima, E., Tsuge, S., Sakurai, R., et al. (2024). Current perspectives of viral hepatitis. *World Journal of Gastroenterology, 30(18), 2402*. doi: 10.3748/wjg.v30.i18.2402

Wang, L. Y. (2014). A forecasting research on population size of Jilin Province based on grey GM(1,1) model. *Applied Mechanics and Materials, 667, 273–276*. https://doi.org/10.4028/www.scientific.net/AMM.667.273

Wang, Y., Shen, Z., & Jiang, Y. (2018). Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China. *PLoS ONE, 13*(9), e0201987*. https://doi.org/10.1371/journal.pone.0201987

Wang, Y., Zhao, Z., Wang, M., Hannah, M. N., Hu, Q., Rui, J., Liu, X., Zhu, Y., Xu, J., Yang, M., et al. (2020). The transmissibility of hepatitis C virus: A modelling study in Xiamen City, China. *Epidemiology & Infection, 148, e291*. DOI: https://doi.org/10.1017/S0950268820002885

World Health Organization. (2024). Global hepatitis report 2024. *World Health Organization*. https://www.who.int/publications/i/item/9789240091672

Yang, X., Zou, J., Kong, D., & Jiang, G. (2018). The analysis of GM(1,1) grey model to predict the incidence trend of typhoid and paratyphoid fevers in Wuhan City, China. *Medicine, 97(34), e11787*. DOI: 10.1097/MD.0000000000011787

Yao, J., Fan, M., Xiao, Y., Liang, X., Cai, C., & Wang, Y. (2023). Spatial–temporal characteristics of corrected total phosphorus pollution loads from agricultural non-point sources in Tuojiang River watershed, Sichuan Province of southwestern China. *Environmental Science and Pollution Research, 30(14), 42192–42213*. https://doi.org/10.1007/s11356-023-25244-w

Zhao, D., Zhang, H., & Zhang, R. (2021). Comparison of GM(1,1) and improved GM(1,1) models for prediction of human immunodeficiency virus and acquired immune deficiency syndrome incidence in PR China. *The Southeast Asian Journal of Tropical Medicine and Public Health, 52(6), 733–741.*

Zhou, D., Zhu, D., Li, N., & Han, B. (2021). Exploration of three incidence trend prediction models based on the number of diagnosed pneumoconiosis cases in China from 2000 to 2019. *Journal of Occupational and Environmental Medicine, 63(7), e440–e444*. DOI: 10.1097/JOM.0000000000002258