(Received 17 December 2025, Accepted 29 January 2026, Scheduled 19 February 2026)

Quantitative Durability Prediction of Photovoltaic Roof Waterproof Performance Using a Heterogeneous Cumulative Grey Model

Zhi Li[1], Zhongliang Yao[1], Yongtong Li[2,*], Wenxuan Cui[2], and Yixuan Li[2]

[1] *China Railway Electrification Bureau Group Xi'an Electrification Engineering Co., Ltd., Baoji 721000, China*
[2] *School of Civil Engineering, Tangshan University, Tangshan 063000, China*

ABSTRACT: Building Integrated Photovoltaics (BIPV) technology has emerged as a significant trend in building low-carbon and energy-efficient structures. Exposure to rainwater erosion and immersion can cause its waterproofing failure, significantly shorten the service life of the roof, and considerably lower indoor comfort. To address this, we developed a construction method for waterproofing the details of photovoltaic roofs with embedded bolts. This approach optimized the joint design, enhanced the continuity of the waterproofed layer, and improved the building efficiency. We propose a Multivariate Heterogeneous Accumulation Grey Model to quantify the performance and long-term degradation of PV roof waterproofing, which can fully exploit factors such as temperature change and water flow. The application of the new model for quantitative prediction not only demonstrates the effectiveness of the new process improvements, but also provides a novel theoretical tool for future research. Prediction results indicate that the total exudate volume under the new process is less than 2,000 mL (only 1/6 that of the control group). The experiments demonstrate that key waterproofing details with embedded bolts are superior to those of traditional methods in terms of impermeability and durability. The results provide a scientific and technical solution for improving the waterproofness of photovoltaic roofs.

1. INTRODUCTION

With the development of low-carbon and energy-efficient buildings, Building-Integrated Photovoltaics (BIPV) have become popular [1]. Compared with photovoltaic systems, BIPV integrates photovoltaic modules as building envelope elements. In diverse BIPV applications, rooftop photovoltaics have advantages, such as high power generation efficiency and good solar exposure. However, the waterproofing quality of photovoltaic roofs remains an issue that limits their performance [2]. Unlike traditional roofing systems, photovoltaic roofs require frames to be mounted on the roof. The contact details between the frames and roof panels can easily degrade the waterproof layer. When they are severely eroded and exposed by rainwater, defects in these parts can deteriorate waterproof performance. Therefore, long-term waterproofing reliability is an important technical guarantee for reliable photovoltaic system development. The design and construction quality of waterproofing details are key factors for the lifespan and safety of building-integrated photovoltaic systems [3].

The performance of roof waterproofing is important for building comfort and service. Several researchers have conducted targeted research in several technical directions to improve waterproofing performance. For intelligent waterproofing, Salgado et al. [4] developed a smart system based on monitoring and driving electronic devices together with temperature-humidity sensors and heating blocks. The smart system effectively prevents the cracking of waterproofing materials with smart detection and maintenance, which improves the durability of the waterproofing layer. Feiteira et al. [5] explored the mechanical properties of several polymer systems used in waterproofing and investigated the effect of thickness and reuse on their tensile and anti-indentation performance. Ahn et al. [6] used satellite imagery to develop a building classification service platform for waterproof which provides an efficient technical solution for regional roof waterproof management.

Additionally, because of the long service life of BIPV roof waterproofing and obstructions, such as suspended ceilings, early-stage leakage is not possible, and water damage can increase and cause repairs and costs. Therefore, accurate forecasts of waterproofing durability must be sought. Popular predictors for waterproofing include support vector machines [7], linear regression [8], and neural networks [9]. These methods often require a large amount of data to be reliable; in practice, available data for the waterproofing of photovoltaic roofs do not have large samples, long time series, etc. Therefore, existing datasets for photovoltaic roof waterproofness typically have small sample sizes, small time series, and a lack of correlation among variables.

Grey prediction models can perform very well under exact data conditions. They are widely used in small-sample forecasting problems. The grey system theory has the unique advantage of solving problems with limited information and incomplete data. Differential equations and data preprocessing

* Corresponding author: Yongtong Li (lyt18632012952@163.com).

TABLE 1. Comparison of the proposed process and model with existing literature.

Related References	Literature Review
Hu et al. [21]	The Euler-Lagrange coupling method was employed to predict waterproofing performance. However, this model has relatively high computational costs and lacks stability.
Mao et al. [22]	The study analyzed the waterproofing performance of immersed tube tunnel joints by considering two primary factors: compressive load and seawater environment. It also estimated the service life of gaskets, though the results lack rigor.
Wu et al. [23]	A numerical simulation of the waterproof gasket was conducted using the Mooney-Rivlin model. This model requires a substantial amount of data.
This study	Most literatures analyze waterproofing materials, while few focus on process improvements. This study proposes a novel process (embedded bolt photovoltaic roof) and predicts its performance using HGMC $(1, N)$.

can be used to make reliable predictions using up to four data points [10]. Grey prediction models have been applied to energy consumption [11], air quality [12], and construction [13].

To improve the grey model prediction, many authors have published improvement studies, most of which focus on the following three points:

(1) Buffer operator: Buffer operators can relax a grey prediction model by reducing the amount of data fluctuation and maintaining the trend of the original sequence, making the adverse effect of abnormal data fluctuation less serious. Other studies proposed strengthening buffer operators [14], weakening buffer operators [15], intelligent buffer operators [16].

(2) Error correction: Prediction accuracy can also be improved by assessing the prediction error and adding correction. For example, we have implemented an error compensation method in the Grey Multivariate Convolution model (GMC$(1, N)$) [17], and we have also implemented triangular residual correction to remove common errors of the grey model for error correction [18].

(3) Accumulation method optimization: Accumulations method optimization is a common technical tool for improving the model prediction performance. Optimization techniques include fractional-order accumulation [19] and new information priority accumulation [20].

All the improvements proposed above are performed in grey prediction models with the aim of improving the prediction accuracy and stability. Based on the grey system theory, we propose a Heterogeneous Accumulation Grey Multivariate Convolution Model (HGMC$(1, N)$), which follows the principles of new information priority and differential information. It assigns more weight to new information and fully exploits differential information in the original data sequences. Compared with previous models, this model employs heterogeneous accumulation operators for different variable sequences and extracts key differential features from a finite information space. Because this model has been predicted with high accuracy, and it has been verified by case studies, we present this work to forecast photovoltaic roof performance. Table 1 shows a comparison with other references.

The contributions of this study are twofold: (1) experimentally evaluate the waterproofing capability of the proposed photovoltaic roof detail using embedded bolts and collecting water-seepage data; (2) develop a durability prediction model for photovoltaic roof waterproofing from the heterogeneous accumulation grey model. Our work provides a scientific solution for BIPV roof leakages, applies the grey prediction model to building durability evaluation, and is crucial for low-carbon and energy-efficient buildings.

2. PHOTOVOLTAIC ROOF EXPERIMENT

2.1. Waterproofing Experiment for Photovoltaic Roofs

In this paper, we propose a waterproof detail for photovoltaic roofs with embedded bolts (Figure 1). Our design consists of "embedding and fixing — membrane sealing — joint encapsulation — system integration". We propose to build the details as follows: (1) Support bolts on the steel frame of the light roof panels to ensure safe and stable mounting. (2) During the laying of a waterproof membrane, openings are treated at a bolt penetration point to improve the membrane continuity. (3) Photovoltaics supports are secured by bolt compression, and the entire assembly is secured by the waterproof membrane and combined to form a continuous and unified waterproof layer.

The proposed waterproofing detail with embedded bolts is suitable for photovoltaic support installation on lightweight roof panels with steel frames. It can be widely applied in the photovoltaic retrofitting of large-span buildings, such as urban metro depots, industrial plants, and parking structures. To verify its waterproofing effectiveness and collect the test data required for the grey model, a prototype of the photovoltaic roof detail with embedded bolts was fabricated. A control group, using the traditional connection method (welding), was also prepared for comparison.

2.2. Data Acquisition

This study employed both a control group and an experimental group, comprising 12 datasets. The experimental details for each group are as follows.

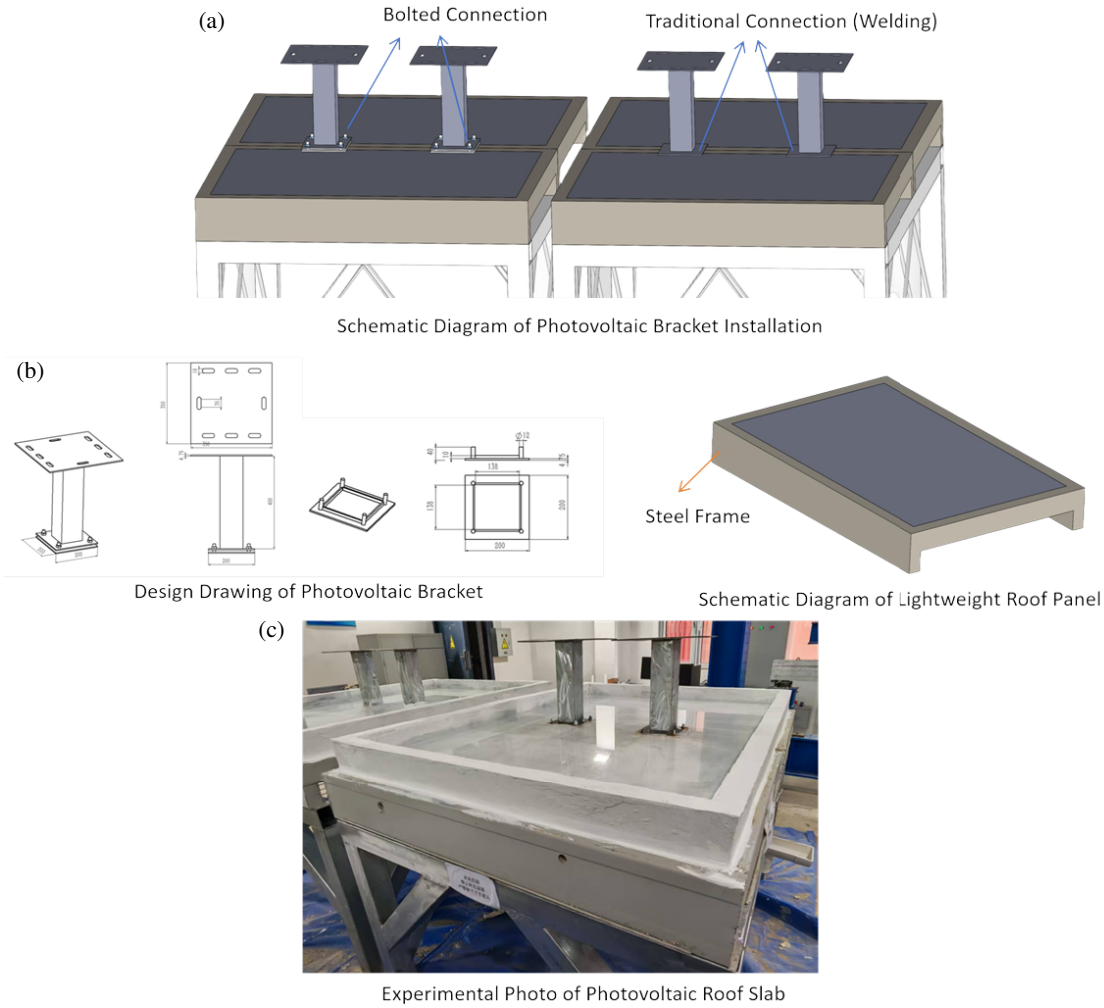

FIGURE 1. Waterproofing experiment for photovoltaic roofs.

2.2.1. Control Group Setup

The Control Group (G1–G6) imitated the traditional connection (welding) scenario of standard on-site photovoltaic support installed, which is a common idea in early BIPV projects and is used as a benchmark for waterproofing performance. The control groups were as follows.

1) The photovoltaic support was welded into the light roof.
2) A waterproof coating was applied, and a 1.5 mm thick PVC membrane with a 100 mm lap width was laid.

2.2.2. Experimental Group Setup

The E1–E6 experiment groups were constructed as follows. Based on the photovoltaic roof, the waterproofing detail (including the embedded bolts proposed) was as follows:

1) The prefabricated base was positioned on the roof panel and welded to the main rib using perimeter welding.
2) The sleeves were laid over the bolts, waterproof coated, and a 1.5 mm thick PVC membrane was laid with a lap width of 100 mm.
3) The bars were lifted off the sleeves, and the photovoltaic supports were aligned with the embedded bolt holes.

The control and experimental teams used water ponding. Roof loading by jacks was used to push the water up and gather the experimental data required for the grey model. The data used were the total amount of water seepage per 10 min window (independent variable) and the temperature of the surrounding area (dependent variable). The data showed different peaks and significant variability.

To implement the accumulation strategy of the model, adjacent accumulation is used for the local temperature data, as the temperature data of adjacent time periods have the highest effect on the total seepage volume. The total seepage volume data are used for new information priority accumulation, which gives greater weight to new data than to past data, highlighting that recent seepage information has a much higher impact on the prediction results.

3. MODELS

We aim to empirically measure the performance of roof waterproofing at different temperatures with cumulative water seepage to provide theoretical aid for designing waterproofing materials and their climate adaptation benefits. Owing to the limited sample size, traditional statistical learning or deep learning-based prediction models often cannot adequately perform accurately and sometimes lead to overfitting or high prediction volatility. To address these data characteristics, this study employs a grey model with a differentiated accumulation strategy for analysis, which can more fully utilize the effective information in the available data.

3.1. Multivariate Heterogeneous Cumulative Grey Model

Conventional grey models typically apply a uniform accumulated generating operation (AGO) to both the feature series (cumulative seepage) and influencingfactor series (temperature). This homogeneous accumulation weakens intervariable heterogeneity, induces loss of informative differences, and degrades the predictive accuracy and stability [24]. A heterogeneous accumulation grey model (HGMC$(1,N)$) is proposed. For the seepage series, newinformationpriority accumulation (NIP-AGO) is adopted to increase the weight of recent observations and downweight obsolete data. For the temperature series, adjacent accumulation (A-AGO) is employed to capture instantaneous driving and thermalinertia effects and mitigate local-information loss caused by oversmoothing in the conventional 1AGO. The modeling procedure is as follows:

(1) Given the cumulative seepage data $S_1^{(0)}$, NIP-AGO is applied to obtain the corresponding accumulated sequence as follows:

$$s_1^{(\varpi)}(k)=\sum_{i=1}^{k}\varpi^{(k-i)}s_1^{(0)}(i),\varpi\in(0,1),\ k=1,2,\ldots,m. \quad (1)$$

$$S_1^{(\varpi)}=\left\{s_1^{(\varpi)}(1),s_1^{(\varpi)}(2),\ldots,s_1^{(\varpi)}(m)\right\}, \quad (2)$$

(2) Given the temperature data $S_j^{(0)}$, A-AGO is applied to obtain the corresponding influencingfactor sequence as follows:

$$s_j^{(\chi_j)}(k)=\gamma_j s_j^{(0)}(k-1)+s_j^{(0)}(k),\chi_j\in(-1,1),\\ j=2,\ldots,N,k=1,2,\ldots,m. \quad (3)$$

$$S_j^{(\chi_j)}=\left\{s_j^{(\chi_j)}(1),s_j^{(\chi_j)}(2),\ldots,s_j^{(\chi_j)}(m)\right\}, \quad (4)$$

(3) The basic form of the grey prediction model (discrete form) is represented as follows:

$$s_1^{(0)}(t)+c_1 g_1^{(\varpi)}(t)=\sum_{j=2}^{N}c_j h_j^{(\chi_j)}(t)+\mu,$$

$$g_1^{(\varpi)}(t)=\frac{s_1^{(\varpi)}(t)+s_1^{(\varpi)}(t-1)}{2},$$

$$h_j^{(\chi_j)}(t)=\frac{s_j^{(\chi_j)}(t)+s_j^{(\chi_j)}(t-1)}{2}. \quad (5)$$

(4) The whitening equation represents the dynamic equilibrium of the system. It converts the accumulated sequence into a continuous-time system using a whitening differential equation.

$$\frac{ds_1^{(\varpi)}(t)}{dt}+c_1 s_1^{(\varpi)}(t)=c_2 s_2^{(\chi_2)}(t)+c_3 s_3^{(\chi_3)}(t)+\ldots\\ +c_N s_N^{(\chi_N)}(t)+\mu,\ t=1,2,\ldots,m, \quad (6)$$

(5) The parameter values are estimated using the least squares method as follows:

$$[\hat{c}_1,\hat{c}_2,\ldots,\hat{c}_N,\hat{u}]^T=\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{T}, \quad (7)$$

where

$$\mathbf{A}=\begin{bmatrix} -\frac{s_1^{(\varpi)}(1)+s_1^{(\varpi)}(2)}{2} & \frac{s_2^{(\chi_2)}(1)+s_2^{(\chi_2)}(2)}{2} & \cdots & \frac{s_N^{(\chi_N)}(1)+s_N^{(\chi_N)}(2)}{2} & 1\\ -\frac{s_1^{(\varpi)}(2)+s_1^{(\varpi)}(3)}{2} & \frac{s_2^{(\chi_2)}(2)+s_2^{(\chi_2)}(3)}{2} & \cdots & \frac{s_N^{(\chi_N)}(2)+s_N^{(\chi_N)}(3)}{2} & 1\\ & & & & \\ -\frac{s_1^{(\varpi)}(m-1)+s_1^{(\varpi)}(m)}{2} & \frac{s_2^{(\chi_2)}(m-1)+s_2^{(\chi_2)}(m)}{2} & \cdots & \frac{s_N^{(\chi_N)}(m-1)+s_N^{(\chi_N)}(m)}{2} & 1 \end{bmatrix},$$

$$\mathbf{T}=\begin{bmatrix} s_1^{(\varpi)}(2)-s_1^{(\varpi)}(1)\\ s_1^{(\varpi)}(3)-s_1^{(\varpi)}(2)\\ \vdots\\ s_1^{(\varpi)}(m)-s_1^{(\varpi)}(m-1) \end{bmatrix},$$

(6) From Eq. (6), the time series solution is obtained as follows:

$$\hat{s}_1^{(1)}(t)=s_1^{(0)}(1)e^{-\hat{c}_1(t-1)}\\ +\sum_{\tau=2}^{t}\left\{e^{-\hat{c}_1(t-\tau+0.5)}\frac{f(\tau)+f(\tau-1)}{2}\right\} \quad (8)$$

where

$$f(t)=\hat{c}_2 s_2^{(\chi_2)}(t)+\hat{c}_3 s_3^{(\chi_3)}(t)+\ldots+\hat{c}_N s_N^{(\chi_N)}(t)+u. \quad (9)$$

(7) Because the sequence is previously subjected to an accumulation operation to smooth the data, an inverse accumulation procedure should be performed during the solution process to recover the original data characteristics.

$$\hat{s}_1^{(0)}(1)=s_1^{(0)}(1),\quad \hat{s}_1^{(0)}(k)=\hat{s}_1^{(\varpi)}(k)-\varpi\hat{s}_1^{(\varpi)}(k-1),\\ k=2,3,\ldots,m. \quad (10)$$

The HGMC$(1,N)$ model contains multiple unknown parameters. Therefore, in this study, a particle swarm optimization algorithm (PSO) is employed to optimize these parameters. We take the mean absolute percentage error (MAPE) as a fitness function of parameter optimization algorithm and convert the parameter search into an answer to the following nonlinear objective function:

$$\min \text{MAPE}=\frac{1}{m}\sum_{k=1}^{m}\left|\frac{s_1^{(0)}(k)-\hat{s}_1^{(0)}(k)}{s_1^{(0)}(k)}\right| \quad (11)$$

TABLE 2. Prediction results of HGMC$(1, N)$ model with different initial value perturbations.

True value	HGMC$(1, N)$ model prediction value			
	$(0, 0, 0)$	$(-0.1, 0, 0)$	$(-0.1, 0.1, 0)$	$(-0.1, 0.1, 0.1)$
24047.30	24047.30	21642.57	21642.57	21642.57
27004.00	26973.79	26982.75	26982.93	26989.70
34684.00	34636.46	34592.51	34636.77	34657.75
36620.00	36620.00	36653.67	36585.61	36595.65
34915.70	34947.62	35016.67	34915.72	34915.72
36466.00	36158.54	36214.27	36465.97	36465.96
36499.80	36578.76	36541.04	36428.16	36456.21

To enhance the prediction accuracy and validate the effectiveness of the model, the following tasks are carried out in this study: (1) the key parameters of the proposed model are optimized using a dedicated optimization algorithm; (2) the rationality of the initial values for the new model is verified in Section 3.2; (3) the stability validation of the new model is conducted in Section 3.3; and (4) the new model is applied to a practical case of water seepage in photovoltaic roofs in Section 3.4, to confirm its engineering applicability.

3.2. Initial Value Validity

In traditional grey prediction models, the initial value is always the first observation of the accumulated sequence that does not fully incorporate the initial information. In the proposed HGMC$(1, N)$ model, both the accumulation order and weighting coefficients are different, and their initial values can be adjusted depending on the data structure and accumulation procedure by studying the influence of slight perturbations on the model results. However, as the order of the new model varies dynamically, the model output also changes, rendering the theoretical proof of this behavior difficult. Therefore, a case study was conducted for empirical validation.

3.2.1. Case 1 Analysis of Coal Consumption in Inner Mongolia

The data for this analysis were derived from [25], with coal consumption in Inner Mongolia serving as the system characteristic sequence S_1, while population and regional Gross Domestic Product (GDP) were considered as the related factor sequences U_2, U_3. Let $\delta_1, \delta_2, \delta_3$ denote the various disturbances in the original sequences $S_1^{(0)}(1), U_2^{(0)}(1), U_3^{(0)}(1)$. The HGMC$(1, N)$ model was applied to predict under different initial-value perturbations to examine the effect of initial-value variation on prediction performance, as shown in Table 2.

(0,0,0) indicates that $S_1^{(0)}(1)$, $U_2^{(0)}(1)$, $U_3^{(0)}(1)$ remain unchanged, and $(-0.1, 0, 0)$ signifies a 10% decrease in $S_1^{(0)}(1)$, whereas the other two variables remain constant. $(-0.1, 0.1, 0)$ represents a 10% decrease in $S_1^{(0)}(1)$, a 10% increase in $U_2^{(0)}(1)$, and the third variable remains unchanged. Similarly, $(-0.1, 0.1, 0.1)$ reflects simultaneous changes in all the three variables. As observed from the results, when $S_1^{(0)}(1)$ decreased by 10%, the overall prediction results changed. When $S_1^{(0)}(1)$ decreased by 10%, $U_2^{(0)}(1)$ increased by 10%, and $U_3^{(0)}(1)$ increased by 10%. The prediction results corresponded to the fifth column in the table, where it is evident that the initial values underwent substantial changes. In conclusion, when the initial values in the HGMC$(1, N)$ model were altered, the predicted outcomes exhibited significant changes, whether the disturbances were applied simultaneously to all three variables at -10%, $+10\%$, and $+10\%$, or whether individual disturbances were applied to $S_1^{(0)}(1)$ (-10%), $U_2^{(0)}(1)$ $(+10\%)$, and $U_3^{(0)}(1)$ $(+10\%)$.

The results show that the HGMC$(1, N)$ model is capable of transmitting the implied system information by varying the initial value. This confirms that the model accurately captures the effects of new information using a dynamic internal parameter adaptation. This feature corresponds to a precise capture of the dynamic evolution laws of the model and is consistent with the minimum information of the grey prediction theory. Thus, the proposed model is more sensitive to the initial information, more adapted to small-sample prediction scenarios, and produces more reliable prediction results.

3.3. Solution Stability

Model stability in the grey system theory is required for reliable predictions. Parameters in traditional grey models are estimated using the least squares method and processed through accumulation, leading to a high prediction bias, which is why stability analysis of the model solution is important. As such, HGMC$(1, N)$ model is highly complex and can result in the accumulation of a large number of numerical disturbances in the parameter estimation. To address this issue, we present the matrix perturbation theory and error bound derivation to quantitatively describe the influence of small perturbations on the model and its stability. This section analyzes the disturbance bounds of the HGMC$(1, N)$ model solution, with the following conclusions:

Lemma 1 [26] Assume that $D \in \mathbf{C}^{m\times n}, \mathbf{Y} \in \mathbf{C}^m$. Let the solutions to the least squares problems $\|Ds - y\|_2 = \min$ and

$\|Ys-c\|_2=\min$ be denoted by s and $s+h$, respectively. $D^\dagger$ is the generalized inverse matrix of D.

If $\text{rank}(D)=\text{rank}(Y)=n$ and $\|D^\dagger\|_2\|\mathbf{E}\|_2<1$ are valid, then

$$\|h\|\le\frac{\kappa_\dagger}{\gamma_\dagger}\left(\frac{\|\mathbf{E}\|_2}{\|D\|}\|x\|+\frac{\|k\|}{\|D\|}+\frac{\kappa_\dagger}{\gamma_\dagger}\frac{\|\mathbf{E}\|_2}{\|D\|}\frac{\|r_x\|}{\|D\|}\right),\quad(12)$$

where $\gamma_\dagger=1-\|D^\dagger\|_2\|\mathbf{E}\|_2$, $r_x=y-Ds$, $\kappa_\dagger=\|D^\dagger\|_2\|D\|$.

Theorem 1 Let s be the solution of the HGMC$(1,N)$ model, and $[\hat{c}_1,\hat{c}_2,\ldots,\hat{c}_N,\hat{u}]=(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{T}$ satisfy **Lemma 1**. Assuming that the original data $\{S_1^{(0)}(k),S_2^{(0)}(k),S_3^{(0)}(k),\ldots,S_N^{(0)}(k)\},k=2,3,\ldots,m-1$ experiences a disturbance $\{\delta_1,\delta_2,\delta_3,\ldots,\delta_N\}$ and under this perturbation, the model solution can be expressed as follows:

$$L\left[s_j^{(0)}(k)\right]=\frac{\kappa_\dagger\sqrt{\delta_1^2\left(1+\sum_{i=1}^{m-k}(\varpi^i+\varpi^{i-1})^2\right)+\sum_{i=2}^{N}\left(1+(\chi_i+1)^2+\chi_i^2\right)\delta_i^2}}{2\gamma_\dagger\|\mathbf{A}\|}\left(\left(\|s\|+\frac{\kappa_\dagger\|r_x\|}{\gamma_\dagger\|\mathbf{A}\|}\right)+\frac{\kappa_\dagger\sqrt{1+\sum_{i=1}^{m-k}(\varpi^i-\varpi^{i-1})^2}|\delta_1|}{\gamma_\dagger\|\mathbf{A}\|}\right.$$

Proof:

If $\{\delta_1,\delta_2,\delta_3,\ldots,\delta_N\}$ disturbance occurs in $\{S_1^{(0)}(k),S_2^{(0)}(k),S_3^{(0)}(k),\ldots,S_N^{(0)}(k)\},k=2,3,\ldots,m-1$, then

$$\hat{\mathbf{A}}=\mathbf{A}+\Delta\mathbf{A}=\mathbf{A}+\begin{bmatrix}0&0&\ldots&0&0\\0&0&\ldots&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots\\-\frac{1}{2}\delta_1&\frac{1}{2}\delta_2&\ldots&\frac{1}{2}\delta_N&0\\-\frac{\varpi+1}{2}\delta_1&\frac{1+\chi_2}{2}\delta_2&\ldots&\frac{1+\chi_N}{2}\delta_N&0\\-\frac{\varpi^2+\varpi}{2}\delta_1&\frac{\chi_2}{2}&\ldots&\frac{\chi_N}{2}&0\\\vdots&\vdots&\vdots&\vdots&\vdots\\-\frac{\varpi^{m-k}+\varpi^{m-k-1}}{2}\delta_1&0&0&0&0\end{bmatrix}$$

$$\hat{\mathbf{T}}=\mathbf{T}+\Delta\mathbf{T}=\mathbf{T}+\begin{bmatrix}0\\0\\\vdots\\\delta_1\\(\varpi-1)\,\delta_1\\\left(\varpi^2-\varpi\right)\delta_1\\\vdots\\\left(\varpi^{m-k}-\varpi^{m-k-1}\right)\delta_1\end{bmatrix}$$

From this, we can conclude that

$$\Delta\mathbf{A}^T\Delta\mathbf{A}=\begin{bmatrix}\frac{\delta_1^2}{4}(1+\sum_{i=1}^{m-k}(\varpi^i+\varpi^{i-1})^2)&\ldots&-\frac{\delta_1\delta_N}{4}\left((1+\varpi)^2\chi_N+\varpi+2\right)&0\\-\frac{\delta_1\delta_2}{4}\left((1+\varpi)^2\chi_2+\varpi+2\right)&\ldots&\frac{\delta_2\delta_N}{4}(1+(1+\chi_2)(1+\chi_N)+\chi_2\chi_N)&0\\\vdots&\ddots&\vdots&\vdots\\-\frac{\delta_1\delta_N}{4}((1+\varpi)^2\chi_N+\varpi+2)&\ldots&\frac{\delta_N^2}{4}\left(1+(1+\chi_N)^2+\chi_N^2\right)&0\\0&\ldots&0&0\end{bmatrix}$$

Therefore

$$\|\Delta\mathbf{A}\|_2=\sqrt{\lambda_{\max}(\Delta\mathbf{A}^T\Delta\mathbf{A})}<\sqrt{\begin{array}{l}\frac{\delta_1^2}{4}\left(1+\sum_{i=1}^{m-k}(\varpi^i+\varpi^{i-1})^2\right)\\+\frac{\delta_2^2}{4}\left(1+(\chi_2+1)^2+\chi_2^2\right)\\\ldots\\+\frac{\delta_N^2}{4}\left(1+(\chi_N+1)^2+\chi_N^2\right)\end{array}}$$

$$=\frac{\sqrt{\delta_1^2\left(1+\sum_{i=1}^{m-k}(\varpi^i+\varpi^{i-1})^2\right)+\sum_{i=2}^{N}\left(1+(\chi_i+1)^2+\chi_i^2\right)\delta_i^2}}{2}$$

$$\|\Delta\mathbf{T}\|_2=\sqrt{\begin{array}{l}1+(\varpi-1)^2+\left(\varpi^2-\varpi\right)^2\\+\ldots\left(\varpi^{m-k}-\varpi^{m-k-1}\right)^2\end{array}}\,|\delta_1|$$

$$=\sqrt{1+\sum_{i=1}^{m-k}\left(\varpi^i-\varpi^{i-1}\right)^2}\,|\delta_1|$$

Then the disturbance boundary is

$$L\left[s_j^{(0)}(k)\right]=\frac{\kappa_\dagger\sqrt{\delta_1^2\left(1+\sum_{i=1}^{m-k}(\varpi^i+\varpi^{i-1})^2\right)+\sum_{i=2}^{N}\left(1+(\chi_i+1)^2+\chi_i^2\right)\delta_i^2}}{2\gamma_\dagger\|\mathbf{A}\|}\left(\left(\|s\|+\frac{\kappa_\dagger\|r_x\|}{\gamma_\dagger\|\mathbf{A}\|}\right)+\frac{\kappa_\dagger\sqrt{1+\sum_{i=1}^{m-k}(\varpi^i-\varpi^{i-1})^2}|\delta_1|}{\gamma_\dagger\|\mathbf{A}\|}\right.$$

Based on perturbation boundary analysis, it can be concluded that as the size of the original sequence data increases, the perturbation of the solution also increases. Although a larger perturbation boundary does not necessarily imply a greater disturbance, an increase in the data volume can lead to instability in the prediction outcomes of the model. Therefore, this model is more appropriate for situations with limited data. When $-0.5<\chi_i<1$, the perturbation boundary of the solution increases as the order of χ_i increases. In general, a larger order leads to a larger perturbation boundary, thereby causing instability in the predictions of the model. Furthermore, the perturbation boundary is also related to the number of related factors. The more related the factors are, the larger the perturbation boundary is. Consequently, this model is not suitable for problems involving a large number of related factors.

3.4. Case Study

In the preceding sections, the characteristics of the proposed HGMC$(1,N)$ model, namely, the heterogeneous accumulation mechanism, parameter optimization strategy, and stability under perturbation, were theoretically verified. However, a theoretical analysis alone is insufficient to demonstrate the model's practical effectiveness and superiority. Therefore, the proposed model was applied to a representative case study from the literature for comparative analysis with existing grey prediction models and other conventional approaches. Two evaluation metrics — the Mean Absolute Simulation Percentage Error (MASPE) and Mean Absolute Prediction Percentage Error

TABLE 3. Fitting results of China's clean energy consumption under different models.

Raw data	GM(1, N)	OBGM(1, N)	GMC(1, N)	NARX	GOMC(1, N)	HGMC(1, N)
32511.61	32511.61	32511.61	32511.61	28844.89	32511.61	32511.61
39007.39	34854.67	38785.49	38302.11	30859.66	37978.83	38984.63
42525.13	52621.57	43081.23	42576.52	39520.92	42474.17	42525.16
48116.08	54480.64	47557.65	46900.68	46717.28	47203.99	47241.75
52018.50	55361.71	52438.78	51415.26	51088.18	52309.93	52708.46
57963.93	58190.50	57523.32	56092.70	56326.59	57849.53	57864.02
61897.00	63714.18	62180.06	60440.68	61013.22	63527.15	61873.38
66352.00	69674.74	66263.59	63937.05	64416.26	69026.36	66668.41
MASPE	8.91	0.74	2.12	6.39	1.72	**0.56**
74358.00	74660.71	69923.32	66478.69	61547.88	74388.79	73867.42
MAPPE	0.41	5.96	10.60	17.23	0.04	**0.61**

(MAPPE) — were introduced to assess model errors during the fitting and prediction phases. The calculation formulas are as follows:

$$\mathrm{APE}_k = \left| \frac{\hat{s}_1^{(0)}(k) - s_1^{(0)}(k)}{s_1^{(0)}(k)} \right| \times 100\% \, (k = 1, 2, \ldots m,)$$

$$\mathrm{MASPE} = \frac{1}{mf} \sum_{k=1}^{mf} \mathrm{APE}_k$$

$$\mathrm{MAPPE} = \frac{1}{m - mf} \sum_{k=mf+1}^{m} \mathrm{APE}_k$$

Based on the above evaluation metrics, the case study is analyzed as follows.

3.4.1. Case 2: Prediction of Clean Energy Consumption in China

The method presented in this case study was used to compare the HGMC(1, N) model proposed in this paper. The data are sourced from [27], where the clean energy consumption from 2011 to 2019 (S_1) is considered the system characteristic sequence, while the total population from 2011 to 2019 (S_2) and the Gross Domestic Product (GDP) (S_3) are treated as the related factor sequences. Ref. [27] provides the prediction results for various models, including the Grey Multivariable Model (GM(1, N)), Grey Multivariable Convolution Model (GMC(1, N)), NGM(1, N) model, Nonlinear Autoregressive Model with External Input (NARX) neural network, and GOMC(1, N) model. This study compares the HGMC(1, N) model with the models presented in the literature. The fitting results are presented in Table 3.

By comparing the fitting results of the six models, it is evident that the GOMC(1, N), OBGM(1, N), and HGMC(1, N) models demonstrate relatively superior fitting performances. The HGMC(1, N) model demonstrates superior performance in both fitting and prediction. The new model achieves higher prediction accuracy than other models. This study demonstrates that HGMC(1, N) captures the data advantage in case study applications. As illustrated, other models show significant fluctuations in their prediction results, and the HGMC(1, N) model maintains stable and consistent fitting performance across annual data when addressing small-sample problems.

4. PREDICTION RESULTS ANALYSIS

4.1. Analysis of Control Group Prediction Results

Based on the HGMC(1, N) model, a comparative experiment involving control and experimental groups was conducted to predict and analyze the waterproofing performance of the photovoltaic roof with embedded bolts. This paper collected data every ten minutes and aggregated the data. The interval of 0–120 minutes served as the fitting data, while the interval of 120–160 minutes served as the prediction data. As shown in Figure 2, the experimental seepage volume increased rapidly in the initial stage and peaked at approximately 40–60 minutes, with the maximum exceeding 12,000 ml. It then gradually decreased in a nonlinear manner, dropping to approximately 2,000 ml in 120 min.

The model prediction results indicated that the slope during the rising phase of seepage was fitted with overall accuracy. All six control groups performed well. However, with an increase in the rate of the specimens, the estimated slope was only somewhat mixed. In the fitting range, the model fits well with the measured data, and the deviations were mainly in the peak forecast. We estimated the model parameters based on the fitting data; if the fit range slowed down and then grew quickly, we estimated the parameters that missed the instantaneous growth rate in the sudden period, which gave us better estimates of the peak predictions as an extrapolation.

From the comparison between the measured and predicted results, it is clear that the predicted curves for all six samples

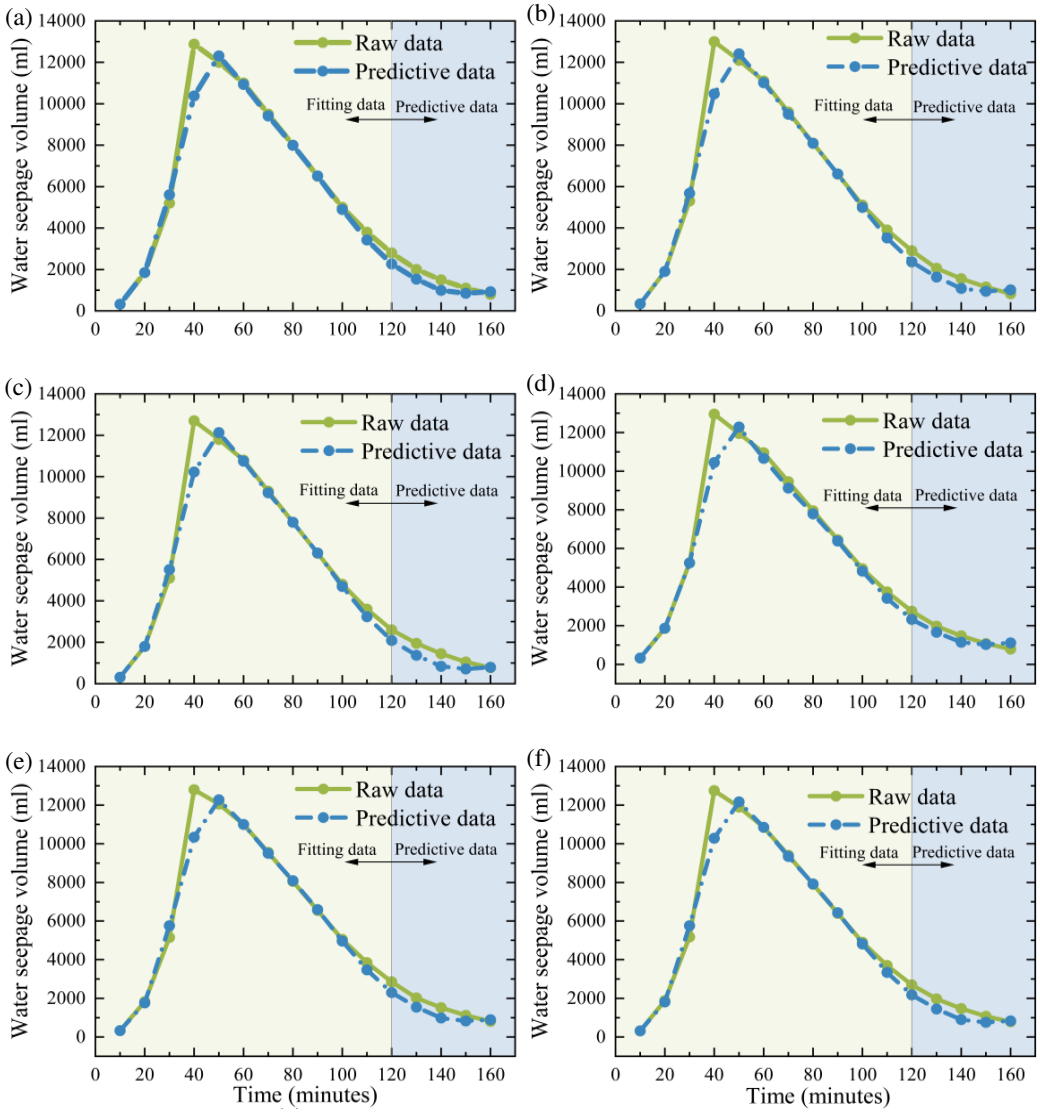

FIGURE 2. Prediction results of the six control groups.

typically coincide with the original curves. For the control group models, the median absolute percentage error (MAPE) was approximately 5%, indicating that the new model is accurate for fitting and prediction. However, during the second prediction, there were some differences between the model predictions and measured values, such as mischaracterizing residual seepage in some samples or overestimating the decline rate in tail seepages. It has been suggested that this is due to the high randomness of roof leakage caused by local crack propagation and water pressure.

4.2. Analysis of Experimental Group Prediction Results

Compared to the control group, the experimental group using photovoltaic roof waterproofing detail with embedded bolts performed much better in terms of waterproofing: the total seepage volume was less than 2,000 mL (only 1/6 that of the control group). In addition, the seepage curve appears smoother without noticeable abrupt changes. The original data and prediction results are shown in Figure 3. The HGMC$(1, N)$ fits the experimental group data well, with errors ranging from 2% to 5%, because the seeps in the experimental case are more stable, and therefore the interference due to data fluctuations in the group is lessened. In addition, this model fits roughly equal to the measured data in terms of the size and decay rate of the seeping curve and nearly the same as the measured data in terms of peak timing and decay, indicating that the model's prediction accuracy increases when the seepage size is small and therefore can be used for early prediction and warning of seepages.

These results demonstrate that the photovoltaic roof waterproofing detail with embedded bolts reduces both the leakage volume and random seepage. Moreover, the seepaging decay in the experimental group followed a smooth, approximately exponential trend, where the model fitting deviation remained within 5%. This results in scientific validation of the long-term predictability of waterproofing performance in engineering applications.

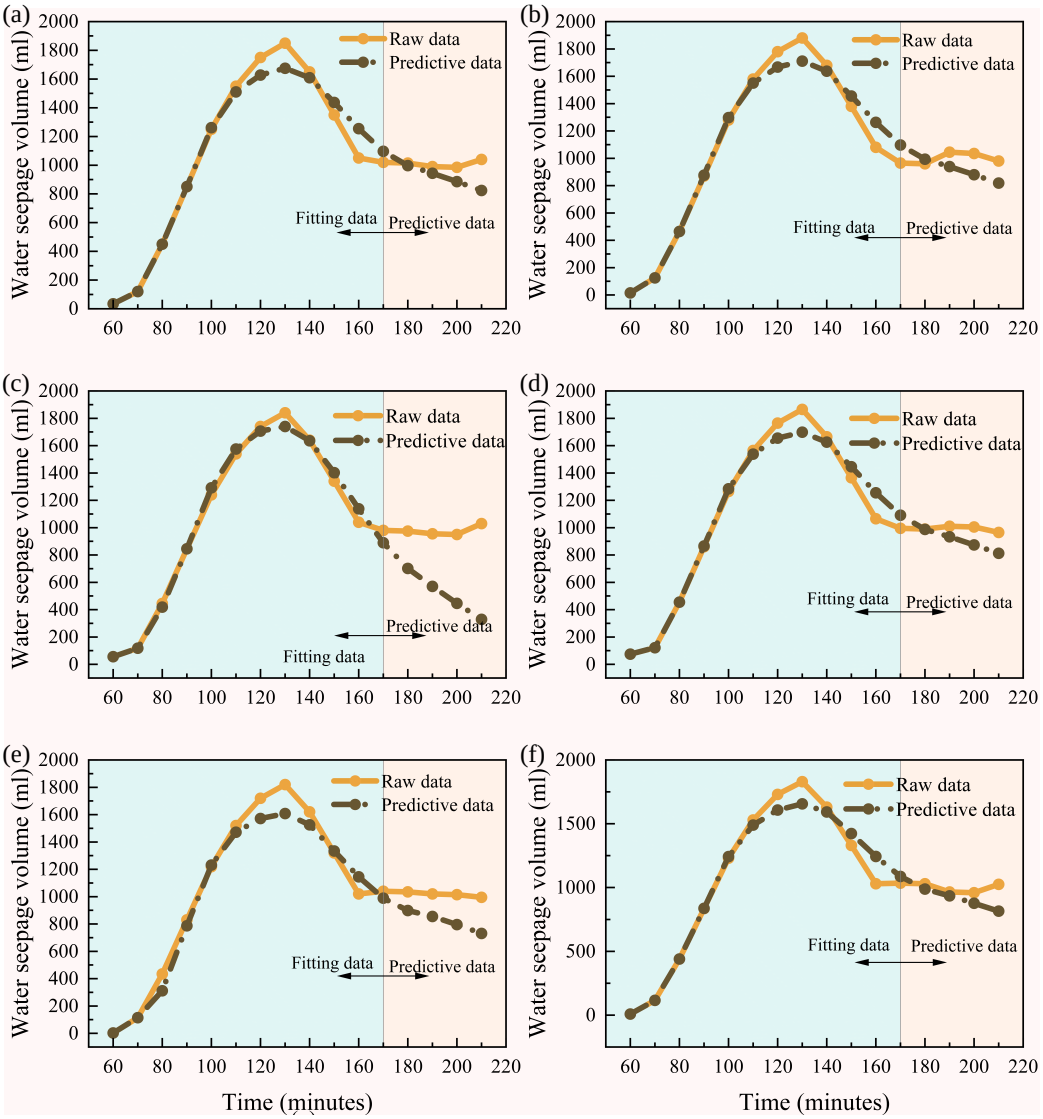

FIGURE 3. Prediction results of the six experimental groups.

5. CONCLUSION

To address the problem of weak waterproofness and durability at the connection joints of photovoltaic supports for BIPV roofs, this study considers two dimensions: joint structural optimization and predictive modeling. We have shown experimental results and theoretically modeled how to improve waterproofing and predict the durability of photovoltaic roofs. The main conclusions are as follows:

(1) An embedded-bolt photovoltaic roof waterproofing detail based on the integrated construction sequence of "embedding and fixing — membrane sealing — joint encapsulation — system integration" is proposed, which significantly improves the overall integrity of roof waterproofing. Compared with the traditional welded joint approach, the new method substantially reduces the risk of leakage and markedly enhances the durability of photovoltaic roof waterproofing.

(2) The Multivariate Heterogeneous Accumulation Grey Prediction Model (HGMC$(1, N)$) is constructed based on the principles of new-information priority and differential information. It overcomes the limitation of homogeneous accumulation in traditional grey models by applying new-information-priority accumulation to cumulative seepage volume and neighborhood accumulation to temperature data, with parameter optimization performed using the Particle Swarm Optimization algorithm.

(3) In the fitting and prediction stages for photovoltaic roof seepage, the HGMC$(1, N)$ maintained mean absolute percentage errors within 5%. It addresses the shortcomings of insufficient utilization of the initial value information in traditional models and can be effectively applied to predict the waterproofing performance of photovoltaic roofs. This study provides a scientifically sound solution for waterproofing prediction in building integrated photovoltaics (BIPV) roofs.

ACKNOWLEDGEMENT

This work is Funded by Science Research Project of Hebei Education Department (BJ2025196) and the Science and Technology Project of Tangshan (No. 24110220d).

REFERENCES

[1] Zhu, G., Y. Xin, Y. Yu, S. Wang, X. Ma, and H. Zhong, "Photovoltaics development pathway and environmental benefits dynamics from technological evolution in china," *Environmental Impact Assessment Review*, Vol. 117, 108157, 2026.

[2] Tseng, Y.-C., L.-H. Yeh, H.-L. Yip, and C.-C. Chueh, "Challenges and future prospects of organic photovoltaics for underwater applications," *Small*, Vol. 21, No. 14, 2412507, 2025.

[3] Salvalai, G. and F. Zhao, "Active prefabricated façade with building-integrated photovoltaic (APF-BIPV) technologies for high energy efficient building renovation: A systematic review of recent research advancements," *Energy and Buildings*, Vol. 348, 116440, 2025.

[4] Salgado, J., C. Ferreira, M. Midão, P. Ferreira, P. Sequeira, S. Ventura, A. Afonso, N. Simões, I. Simões, A. Silva, and L. Costa, "Integration of monitoring and actuation electronics in a smart roof waterproofing system," in *2023 Smart Systems Integration Conference and Exhibition (SSI)*, 1–6, Brugge, Belgium, 2023.

[5] Feiteira, J., J. Lopes, and J. de Brito, "Mechanical performance of liquid-applied roof waterproofing systems," *Journal of Performance of Constructed Facilities*, Vol. 27, No. 3, 244–251, 2013.

[6] Ahn, S., S. Kim, J. Chae, J. Park, and J. Kang, "Automating the classification of roof waterproofing buildings using satellite images," in *2021 International Conference on Information and Communication Technology Convergence (ICTC)*, 309–311, Jeju Island, Republic of Korea, 2021.

[7] Guan, X., M. Z. Li, J. Qin, and C. Wang, "Short-term high-speed rail passenger flow forecasting integrated extended empirical mode decomposition with multivariate and bidirectional support vector machine," *Expert Systems with Applications*, Vol. 298, 129870, 2026.

[8] Carlson, R. K., "Simple linear regression models for prediction of ionization energies, electron affinities, and fundamental gaps of atoms and molecules," *Journal of Chemical Theory and Computation*, Vol. 21, No. 7, 3382–3393, 2025.

[9] Cui, Y., X. Zhou, C. Fan, H. Lin, and S. Deng, "Application of physics-informed neural networks-based hybrid prediction models in campus thermal environment prediction," *Building and Environment*, Vol. 287, 113880, 2026.

[10] Liu, L., Y. Chen, and L. Wu, "The damping accumulated grey model and its application," *Communications in Nonlinear Science and Numerical Simulation*, Vol. 95, 105665, 2021.

[11] Zhao, K. and L. Wu, "Heterogeneity grey model and its prediction of energy consumption under the shared socioeconomic pathways," *Energy*, Vol. 319, 134851, 2025.

[12] Cai, Y., J. Wang, Y. An, Y. Dang, and L. Ye, "Research on accumulative time-delay effects between economic development and air pollution based on a novel grey relational analysis model," *Journal of Cleaner Production*, Vol. 497, 145128, 2025.

[13] Qian, W., J. Chen, and C. Ji, "A novel grey model driven by policy shifts and technological progress and its application in China's wind power supply prediction," *Energy*, Vol. 335, 138154, 2025.

[14] He, L.-Y., L.-L. Pei, and Y.-H. Yang, "An optimised grey buffer operator for forecasting the production and sales of new energy vehicles in china," *Science of the Total Environment*, Vol. 704, 135321, 2020.

[15] Qian, Y., Z. Zhu, X. Niu, L. Zhang, K. Wang, and J. Wang, "Environmental policy-driven electricity consumption prediction: A novel buffer-corrected hausdorff fractional grey model informed by two-stage enhanced multi-objective optimization," *Journal of Environmental Management*, Vol. 377, 124540, 2025.

[16] Zeng, B., F. Yin, J. Wang, and W. Meng, "Prediction and rationality analysis of new energy vehicle sales in China with a novel intelligent buffer operator," *Engineering Applications of Artificial Intelligence*, Vol. 143, 110030, 2025.

[17] Tien, T.-L., "The indirect measurement of tensile strength by the new model FGMC (1,n)," *Measurement*, Vol. 44, No. 10, 1884–1897, 2011.

[18] Zhou, D., A. Al-Durra, K. Zhang, A. Ravey, and F. Gao, "A robust prognostic indicator for renewable energy technologies: A novel error correction grey prediction model," *IEEE Transactions on Industrial Electronics*, Vol. 66, No. 12, 9312–9325, 2019.

[19] Weng, M., H. Duan, and D. Xie, "Forecasting power generation: A novel two-dimensional logistic fractional grey model," *Engineering Applications of Artificial Intelligence*, Vol. 161, 112097, 2025.

[20] Li, Y., Y. Chen, and Y. Wang, "Grey forecasting the impact of population and GDP on the carbon emission in a Chinese region," *Journal of Cleaner Production*, Vol. 425, 139025, 2023.

[21] Hu, Z., H. Mao, K. Ma, S. Xu, Y. Luo, Y. Li, and H. Wang, "Time-varying constitutive model establishment and waterproofing capability prediction of GINA gaskets in submarine immersed tunnel," *Tunnelling and Underground Space Technology*, Vol. 170, 107368, 2026.

[22] Mao, H., Z. Hu, Y. Du, H. Pan, X. Feng, K. Ma, Y. Luo, and H. Wang, "Aging mechanism and life prediction of GINA gasket on immersed tunnel under the combined action of compression load and seawater environment: A multi-scale experimental analysis," *Tunnelling and Underground Space Technology*, Vol. 169, 107252, 2026.

[23] Wu, R., H.-N. Wu, H.-Y. Li, X.-X. Yang, D.-L. Feng, H. Cao, and R.-P. Chen, "Prediction on the waterproof performance of EPDM gaskets in shield tunnels considering the stress state of compression and tension," *Tunnelling and Underground Space Technology*, Vol. 167, 107090, 2026.

[24] Li, Y., Y. Wang, T. Sun, and C. Yan, "Heterogeneous spatiotemporal forecasting the impact of carbon policies on air quality," *Air Quality, Atmosphere & Health*, Vol. 18, No. 10, 3213–3235, 2025.

[25] Duan, H. and X. Luo, "A novel multivariable grey prediction model and its application in forecasting coal consumption," *ISA Transactions*, Vol. 120, 110–127, 2022.

[26] Stewart, G. W., "On the perturbation of pseudo-inverses, projections and linear least squares problems," *SIAM Review*, Vol. 19, No. 4, 634–662, 1977.

[27] Xu, K., X. Luo, and X. Pang, "A new multivariable grey model and its application to energy consumption in China," *Journal of Intelligent & Fuzzy Systems*, Vol. 42, No. 4, 3153–3168, 2022.