Article

Multi-Point Seawall Settlement Prediction with Limited Data Volume Using an Improved Fractional-Order Grey Model

Peng Qin [1,2], Chunmei Cheng [1], Zhenzhu Meng [1], Chunmei Ding [1], Sen Zheng [2,3] and Huaizhi Su [2,4,*]

1   College of Hydraulic and Environment Engineering, Zhejiang University of Water Resources and Electric Power, Hangzhou 310018, China; qinpeng@zjweu.edu.cn (P.Q.); chengcm@zjweu.edu.cn (C.C.); mengzhzh@zjweu.edu.cn (Z.M.); dingchm@zjweu.edu.cn (C.D.)
2   College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210024, China; sen.zheng@epfl.ch
3   Laboratory of Environmental Hydraulic, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
4   The National Key Laboratory of Water Disaster Prevention, Hohai University, Nanjing 210024, China
*   Correspondence: su_huaizhi@hhu.edu.cn

Abstract: Settlement prediction based on monitoring data holds significant importance for engineering maintenance of seawalls. In practical engineering, the volume of the collected monitoring data is often limited due to the restrictions of devices and engineering budgets. Previous studies have applied the fractional-order grey model to time series prediction under the situation of limited data volume. However, the performance of the fractional-order grey model is easily affected by the inappropriate settings of fractional order. Also, the model cannot make dynamic predictions due to the characteristic of fixed step size. To solve the above problems, in this paper, the genetic algorithm with enhanced search capabilities was employed to solve the premature convergence problem. Additionally, to solve the problem of the fractional-order grey model associated with fixed step size, the real-time tracing algorithm was introduced to conduct equal-dimensionally recursive calculation. The proposed model was validated using monitoring data of four monitoring points at Haiyan seawall in Zhejiang province, China. The prediction performance of the proposed model was then compared with those of the fractional-order GM(1,1), integer-order GM(1,1), and fractal theory model. Results indicate that the proposed model significantly improves the prediction performance compared to other models.

Keywords: seawall; settlement prediction; genetic algorithm; fractional-order grey model; real-time tracing algorithm

1. Introduction

Seawalls are commonly established atop foundations of deep soft soils whose physical and mechanical features are distinctive, e.g., high compressibility, high water content, and diminished strength, as well as specific rheological behavior [1,2]. These characteristics lead to significant settlement during both the construction phase and post-construction phase. The total subsidence may reach several meters and the consolidation settlement period lasts more than 10 years [3–5]. The large vertical settlement on soft soil foundations threatens the stability of seawalls, which greatly increases the risk of structural damage and safety incidents. Therefore, predicting seawall settlement is essential to engineering security operations and risk management.

Time series monitoring data record the historical behavior of seawalls, including its structural characteristics and operational patterns [6,7] among which settlement is one of the most crucial parameters. Predicting the settlement based on time series monitoring data using mathematical algorithms forms a reliable approach for assessing the safety status of the seawall [8,9]. At present, prediction methods commonly used in seawall settlement

prediction based on monitoring data can be categorized into two groups: statistical methods and machine learning methods. Statistical methods such as the exponential curve method [10], Asaoka method [11], and Poisson curve method [12] are established based on the idea of data fitting using mathematical curves. Statistical models can not only quantify the influence of each factor on data-varying tendencies, but also effectively describe the characteristics of the data sequence including the distribution and correlation of data. These models are usually established on the basis of certain assumptions such as normality, independence, etc., which means that the accuracy and reliability of the model cannot be guaranteed in the case where the assumptions are not met [13,14]. Machine learning methods, such as wavelet theory [15], chaos theory [16], and neural networks [17], are able to process massive and multidimensional datasets to solve nonlinear and complicated problems. However, both statistical methods and machine learning methods are highly reliant on the quantity and quality of data sequences, meaning that either poor data quality or insufficient data quantity can lead to unsatisfactory prediction performance [15,18].

Insufficient data volume is a common issue in seawall settlement monitoring. In many cases, the volume of monitoring data is fairly limited, and this is not an isolated occurrence [19,20]. The reasons are multifaceted; the number of monitoring devices installed on the seawall may be insufficient to cover all critical areas, resulting in limited data availability. Even with adequate monitoring devices, the current technology may not support high-density and high-precision monitoring, thereby restricting the monitoring frequency and causing a shortage in data volume [21]. Further, due to the intricate environmental conditions in coastal areas, environmental factors during critical periods such as storm surges and floods can further impede the acquisition of monitoring data. The insufficient amount of data in seawall monitoring is a significant issue as it may limit our comprehensive evaluation of a seawall's health condition and potential risks [22,23]. This issue can be addressed in two aspects to ensure the accurate prediction of a seawall's settlement based on monitoring data. We can either increase the number of monitoring points or strengthen the ability of data processing and analysis under the situation of limited data volume.

Most previous studies developed predictive models using monitoring data at a single monitoring point. Using multi-point monitoring data prediction offers several advantages over single-point data prediction, particularly in terms of robustness and ability to capture complex relationships [24,25]. First, multi-point monitoring data provides multiple observations of the same phenomenon, reducing the impact of errors or noise from any single point. Also, different data points can capture various aspects of the system, leading to a more comprehensive understanding and improved prediction accuracy. It allows the prediction model to be less sensitive to missing or erroneous data from individual sensors. Therefore, multi-point prediction provides a solution to ensure the efficiency and accuracy of seawall settlement displacement.

In addition, in engineering cases where a lack of monitoring data is already present, employing mathematical methods to analyze and predict the existing dataset emerges as the most practical, effective, and economical approach. Grey models, especially the GM(1,1) model, exhibit significant advantages in tackling issues related to insufficient data volume [26,27]. In the realm of grey theory, the grey system comprises known information (white) and unknown or uncertain information (grey or black). Grey models use limited and incomplete data to establish differential predictive models, thereby uncovering latent patterns within the dataset [28–31]. Their core competency lies in making predictions with relatively small amounts of data. Typically, only a few data points (four or more) are required to build a grey predictive model, effectively addressing the issues of scarce historical data, incomplete series, and low reliability. In addition, grey models do not require strict regularity in the distribution of sample data, further easing the data requirements.

While grey models possess these advantages, they also have limitations. For instance, they may not be optimal for systems with clearly defined internal mechanisms, and may fall into suboptimal solutions in predicting time series data with high volatility. Wu [32] found that the GM(1,1) model exhibits considerable sensitivity to minor perturbations in data

computation using matrix perturbation theory, resulting in significant interference with the identification of model parameters. To address this issue, he proposed a fractional-order grey model on the basis of the perturbation theory of least squares [33]. The fractional-order grey model, as an extended grey prediction model, is exceptionally adept at handling data sequences that exhibit volatile variations. By introducing fractional-order calculus operators into a continuous grey model, it adeptly transforms original data sequences with fluctuations into ones with heightened regularity, thereby optimizing the traditional grey modeling mechanism and improving the modeling efficiency [34].

Although the fractional-order grey model exhibits advantages in handling complex systems, its application to the prediction of seawall settlement with limited data volume can still be constrained by system and data characteristics. As the fractional-order grey model incorporates fractional-order derivatives and polynomial functions, the determination of its parameters becomes more complex, requiring the employment of specific optimization methods to solve the coefficients of polynomial functions. For example, the variable fractional order r obtained by commonly used intelligent algorithms involves unavoidable randomness, which may increase the unitability of prediction results [35]. For example, the swarm optimization algorithm may obtain a local optimal solution instead of a global optimal solution due to its high convergence speed [36]. The artificial fish swarm algorithm is susceptible to premature convergence, that is, the algorithm may converge to a suboptimal solution without adequately exploring the searching space [37]. Further, the fractional-order grey model cannot fit the time series data dynamically. Due to the high sensitivity and compressibility characteristics of the soft soil foundation of the seawall, and influenced by external factors such as tidal scouring, storm surges, and construction disturbances, the settlement of the seawall exhibits dynamic variations at small temporal scales [38]. Seawall settlement represents a real-time evolving process, but the fractional-order grey model is limited to making predictions with a fixed step size based on existing data [8,39]; therefore, the dynamic prediction of seawall settlement still needs to be improved.

To solve the above two issues associated with the application of the fractional-order grey model in seawall settlement prediction in the case of limited data volume, this study aims to improve the fractional-order grey model by integrating other models. With this objective in mind, a genetic algorithm [40] with enhanced search capabilities was used to solve and optimize the fractional order r of the fractional-order grey model. Then, a real-time tracing algorithm [41] was used to achieve the dynamic prediction. The monitoring of the data of a seawall at the northern bank of the Qiantang River, China, was used to validate the model. The performance of the proposed model was evaluated by comparison with the fractional-order GM(1,1), integer-order GM(1,1), and fractal theory model, respectively.

This paper is organized as follows: Section 2 presents how the proposed model is developed. Section 3 exhibits the information of the case study. Section 4 shows the prediction results. Comparisons of the proposed model with other relative models are also illustrated in this section. Section 5 presents the concluding remarks.

2. Model Development

This section presents the principles of the proposed model including the definitions, modeling procedures, parameter settings, and implementation steps.

2.1. Establishment of the Fractional-Order Grey Model

The most commonly used grey model is the GM(1,1) model, which is an integer-order accumulated generating sequence. Compared to the GM(1,1) model, the fractional-order grey model exhibits greater robustness and has a better performance to cope with datasets with limited volume, and its principles include the following steps [42,43]:

Assuming that the sequence $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots x^{(0)}(n) \right)$ is the original monitoring data and $r \in R^+$, then the *r*-order accumulated generating sequence of $X^{(0)}$ can be written as $X^{(r)} = \left(x^{(r)}(1), x^{(r)}(2), \cdots x^{(r)}(n) \right)$, where

$$x^{(r)}(k) = \sum_{i=1}^{k} \frac{\Gamma(r+k-i)}{\Gamma(k-i+1)\Gamma(r)} x^{(0)}(i), \; k = 1, 2, \cdots, n. \tag{1}$$

$\Gamma(r)$ represents the Gamma function of order *r*, which is defined as

$$\Gamma(r) = \int_0^\infty e^{-t} t^{r-1} dt. \tag{2}$$

The *r*-order cumulative generating sequence of $X^{(0)}$ can be written as $X^{(-r)} = \left(x^{(-r)}(1), x^{(-r)}(2), \cdots x^{(-r)}(n) \right)$, where

$$x^{(-r)}(k) = \sum_{i=0}^{k-1} \frac{\Gamma(r+1)}{\Gamma(i+1)\Gamma(r-i+1)} x^{(0)}(k-i), \; k = 1, 2, \cdots, n \tag{3}$$

Then the first-order cumulative generating sequence of $X^{(r)}$ can be written as $X^{(r-1)} = \left(x^{(r-1)}(1), x^{(r-1)}(2), \cdots x^{(r-1)}(n) \right)$, where

$$x^{(r-1)}(k) = \left(x^{(r)} \right)^{(-1)}(k) = x^{(r)}(k) - x^{(r)}(k-1), \; k = 2, 3, \cdots, n \tag{4}$$

The sequence $Z^{(r)} = \left(z^{(r)}(2), z^{(r)}(3), \cdots z^{(r)}(n) \right)$ can be generated by the close mean value of the r-order accumulated generating sequence $X^{(r)}$, where

$$z^{(r)}(k) = \frac{x^{(r)}(k) + x^{(r)}(k-1)}{2}, \; k = 2, 3, \cdots, n \tag{5}$$

Then, the proposed model of the monitoring data is

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

The parameter $\hat{a} = [a, b]^{\mathrm{T}}$ can be estimated by the least square method:

$$\hat{a} = \left(B^{\mathrm{T}} B \right)^{-1} B^{\mathrm{T}} Y \tag{7}$$

where Y and B are

$$Y = \begin{bmatrix} x^{(r-1)}(2) \\ x^{(r-1)}(3) \\ \vdots \\ x^{(r-1)}(n) \end{bmatrix}, \; B = \begin{bmatrix} -z^{(r)}(2) - 1 \\ -z^{(r)}(3) - 1 \\ \vdots \\ -z^{(r)}(n) - 1 \end{bmatrix} \tag{8}$$

Since

$$\begin{aligned} x^{(r-1)}(k) &= x^{(r)}(k) - x^{(r)}(k-1) \\ &= \sum_{i=1}^{k} \frac{\Gamma(r+k-i)}{\Gamma(k-i+1)\Gamma(r)} x^{(0)}(i) - \sum_{i=1}^{k-1} \frac{\Gamma(r+k-i-1)}{\Gamma(k-i)\Gamma(r)} x^{(0)}(i), \; k = 2, 3, \cdots, n \end{aligned} \tag{9}$$

then

$$z^{(r)}(k) = \frac{\sum\limits_{i=1}^{k} \frac{\Gamma(r+k-i)}{\Gamma(k-i+1)\Gamma(r)} x^{(0)}(i)}{2} + \frac{\sum\limits_{i=1}^{k-1} \frac{\Gamma(r+k-i)}{\Gamma(k-i+1)\Gamma(r)} x^{(0)}(i)}{2}, \; k = 2, 3, \cdots, n \tag{10}$$

Finally, Y and B can be obtained:

$$Y = \begin{bmatrix} (r-1)x^{(0)}(1) + x^{(0)}(2) \\ \dfrac{r(r-1)}{2}x^{(0)}(1) + (r-1)x^{(0)}(2) + x^{(0)}(3) + \cdots \\ \vdots \\ \sum\limits_{i=1}^{n} \dfrac{\Gamma(r+n-i)}{\Gamma(n-i+1)\Gamma(r)}x^{(0)}(i) - \sum\limits_{i=1}^{n-1}\dfrac{\Gamma(r+n-i-1)}{\Gamma(n-i)\Gamma(r)}x^{(0)}(i) \end{bmatrix}, \tag{11}$$

$$B = \begin{bmatrix} -\dfrac{x^{(r)}(1) + x^{(r)}(2)}{2} & 1 \\ -\dfrac{x^{(r)}(2) + x^{(r)}(3)}{2} & 1 \\ \vdots & \vdots \\ -\dfrac{x^{(r)}(n-1) + x^{(r)}(n)}{2} & 1 \end{bmatrix} = \begin{bmatrix} -\dfrac{1}{2}\left[(r+1)x^{(0)}(1) + x^{(0)}(2)\right] & 1 \\ -\dfrac{1}{2}\left[\dfrac{r(r+3)}{2}x^{(0)}(1) + (r+1)x^{(0)}(2) + x^{(0)}(3)\right] & 1 \\ \vdots & \vdots \\ -\dfrac{1}{2}\left[\sum\limits_{i=1}^{n}\dfrac{\Gamma(r+n-i)}{\Gamma(n-i+1)\Gamma(r)}x^{(0)}(i) - \sum\limits_{i=1}^{n-1}\dfrac{\Gamma(r+n-i)}{\Gamma(n-i+1)\Gamma(r)}x^{(0)}(i)\right] & 1 \end{bmatrix}. \tag{12}$$

$\frac{dx^{(r)}}{dt} + ax^{(r)} = b$ is the whitenization equation of the FGM $x^{(r-1)}(k) + az^{(r)} = b$, and therefore the time response function is

$$\hat{x}^{(r)}(k) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-a(k-1)} + \frac{b}{a}, \ k = 2, 3, \cdots, n \tag{13}$$

Then, the restored monitoring data $\hat{X}^{(0)} = \left(\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \cdots \hat{x}^{(0)}(k)\right)$ can be obtained, where $\hat{x}^{(0)}(1) = x^{(0)}(1)$, and

$$\hat{x}^{(0)}(k) = \left(\hat{x}^{(r)}\right)^{-r}(k) = \sum_{i=0}^{k-1}(-1)^i \frac{\Gamma(r+1)}{\Gamma(i+1)\Gamma(r-i+1)}\hat{x}^{(r)}(k-i), \ k = 2, 3, \cdots, n \tag{14}$$

2.2. Fractional-Order Optimization Using Genetic Algorithm

The selection of the order *r* in the fractional-order grey model significantly affects the prediction accuracy. To avoid a local optimal solution problem and premature convergence problem, a genetic algorithm was used to calculate the fractional order *r*. By starting with a diverse initial population of potential solutions, the genetic algorithm helps in exploring multiple regions of the solution space simultaneously, preventing the algorithm from converging too quickly on a suboptimal solution. This diversity ensures a broad search space and increases the likelihood of finding a global optimum. In addition, the genetic algorithm selects individuals based on their fitness, which is related to the accuracy of the fractional-order grey model in settlement prediction. By favoring higher fitness solutions, the algorithm incrementally improves the model parameters [44,45].

The mean relative error (MRE) of the prediction results can be computed using Equation (15).

$$\min f(r) = \frac{1}{n-1}\sum_{k=2}^{n}\frac{\left|x^{(0)}(k) - \hat{x}^{(0)}(k)\right|}{x^{(0)}(k)}, r \in R^{+}. \tag{15}$$

The minimum of the average relative error is solved using the genetic algorithm to determine the optimal fractional order *r*. Figure 1 shows the working principles of the genetic algorithm. Detailed procedures of genetic algorithm refer to [44–47].
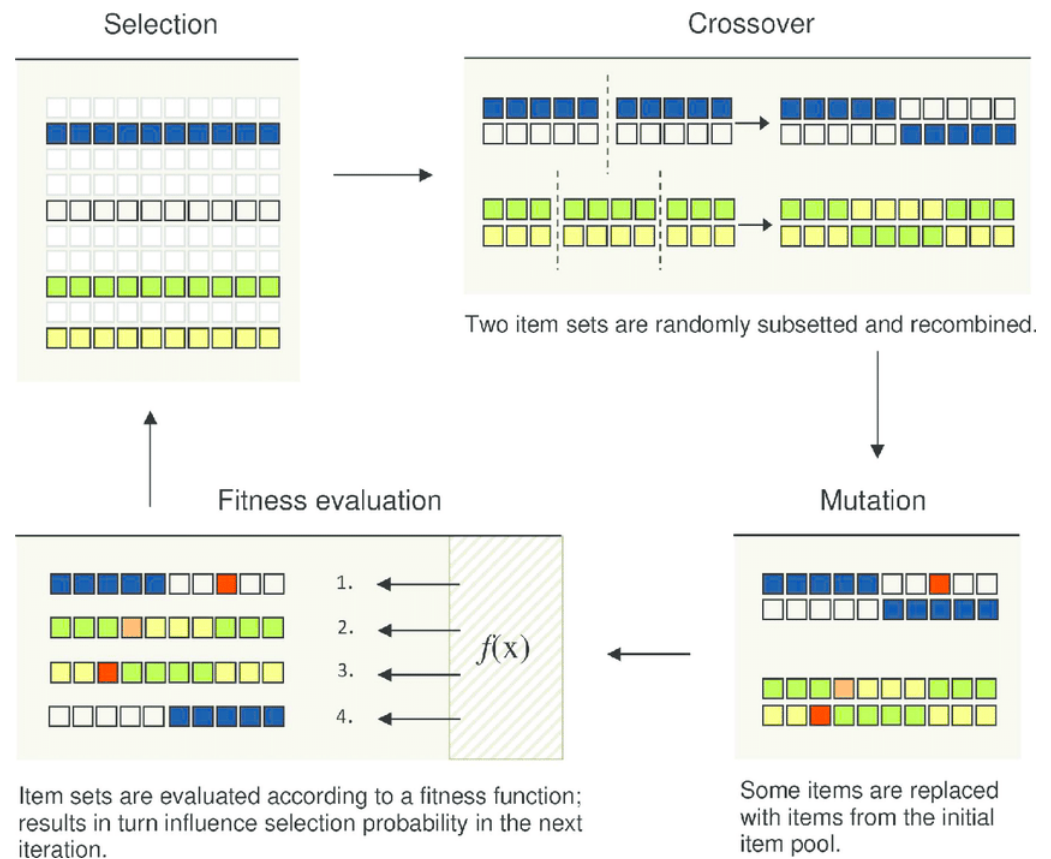
Figure 1. Working principle of genetic algorithm-based optimization.

2.3. Equal-Dimensionally Recursive Calculation Using the Real-Time Tracing Algorithm

To achieve a dynamic prediction of seawall settlement, the real-time tracing algorithm is used to complete the above-introduced fractional grey model based on a genetic algorithm. The real-time tracing algorithm plays a crucial role in overcoming the limitations of the fractional-order grey model related to fixed step size by dynamically adjusting the model parameters in response to new data. The fractional-order grey model relies on a fixed step size that can limit its accuracy and responsiveness to changes in the data; the real-time tracing algorithm adapts the step size based on the latest data points, and thus updates the fractional order and other parameters in real-time. This allows the model to quickly adapt to new patterns and trends in the data and ensures that the model remains relevant and accurate over time.

The real-time tracing algorithm is applied in equal-dimensionally recursive calculation. Suppose that the time series is $\{x_i | i = 1 : n\}$, and the samples of $\{x_i | i = 1 : n\}$ after sample n $\{x_{n+1}, x_{n+2}, \ldots, x_{n+m}\}$ are the values to be predicted using the measured monitoring values $\{x_1, x_2, \ldots, x_n\}$. Assuming that the prediction step length is m, the first m samples of the time series $\{x_{m+1}, x_{m+2}, \ldots, x_{n+m}\}$ are discarded after the last m values are achieved, and then the newly obtained m measured data are added to the original time series to form a new time series. Finally, the above calculations can be repeated until the dimension of the data columns remains unchanged. To fully exploit the valuable information from the seawall monitoring data, this study adopted a single-step prediction approach, where m is set to 1. The schematic of the real-time tracing algorithm is illustrated in Figure 2.
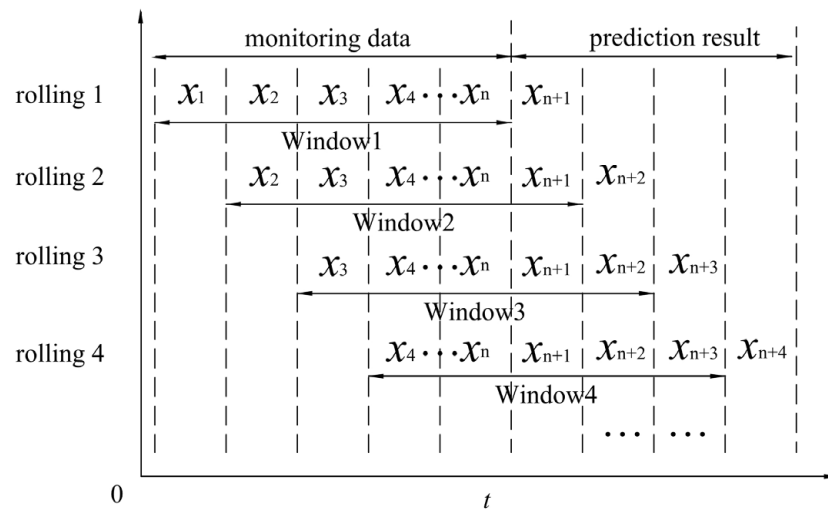
Figure 2. Schematic representation of the real-time tracing algorithm.

Figure 3 shows the flowchart of the proposed improved fractional-order grey model using the genetic algorithm and real-time tracing algorithm.
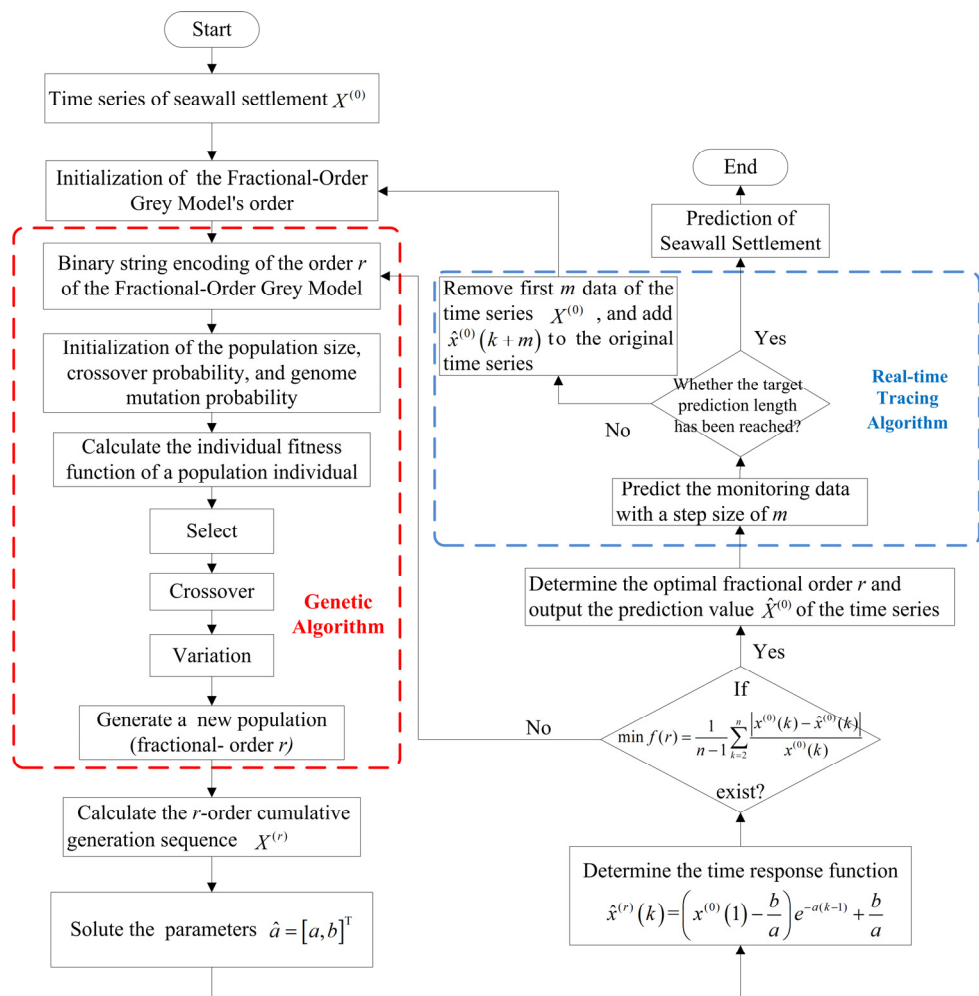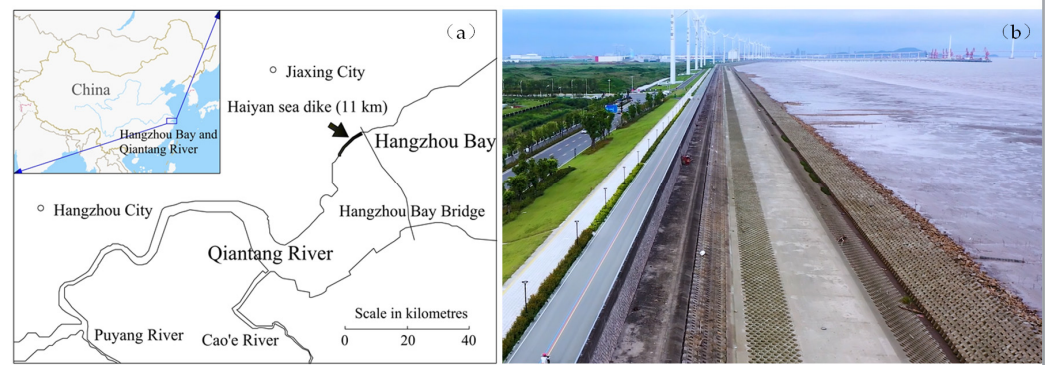

**Figure 3.** Flowchart of the proposed model.

3. Case Study

To validate the proposed model, the Haiyan seawall, which is located on the north bank of the Qiantang River in Zhejiang province, China, was used as a case study. The total length of the seawall is around 11 km. Figure 4a,b show the geographic location and the air-view of the Haiyan seawall, respectively.

Figure 4. The (**a**) geographic location and (**b**) air-view of the Haiyan seawall.

The Haiyan seawall is a vertical dike with a soil–rock mixed structure with revetments on both sides. The study area is mainly intertidal with flat terrain. The studied seawall is built on silty soft soil with an average thickness of 25.0 m, and it is reinforced by plastic drainage boards. Figure 5 exhibits the distribution of the soil layers and monitoring points. The foundation boarding is arranged in a plum blossom pattern, with a horizontal range of 36.0 m, a depth of 15.0 m, and an interval of plastic drainage boards of 1.2 m. It can be seen from the figure that the foundation includes five layers: silt layer, mucky clay layer, mucky silty clay layer, clay layer, and silty clay layer. The settlement monitoring data used in this study was collected at the maximum settlement section Z7+000, which includes four monitoring points SS5, SS6, SS7, and SS8. As the four monitoring points are selected, the prediction involves using various data points collected from multiple sources or sensors over time to predict future values. Multipoint data can capture complex interactions and dependencies between different variables, leading to more accurate and insightful predictions. These monitoring points are located at 1 m, 5.5 m, 9.4 m, and 14 m under the foundation, respectively. We used the JG86 type electromagnetic induction instrument to monitor the settlement of the seawall, which consists of a settlement measuring tube and magnetic rings. A settlement magnetic ring is placed every 3 to 4 m along the settlement measuring tube. During the construction loading period, monitoring was conducted once every 3 days.
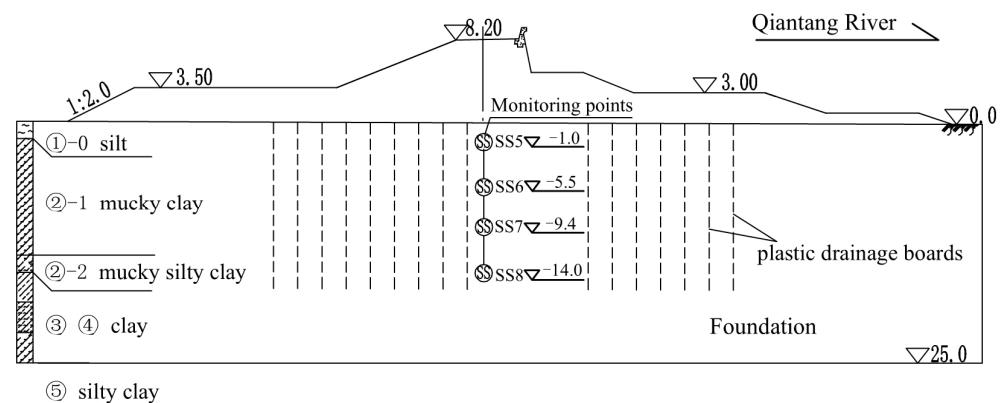
Figure 5. The distribution of the soil layers and monitoring points at the Haiyan seawall.

Figure 6 shows the time variation in settlement at the selected four monitoring points at section Z7+000. Overall, the monitoring data of each of these four selected points exhibit a nonlinear upward dynamic increasing trend. Furthermore, since seawall settlement primarily stems from the consolidation of the foundation soil, the settlement curve displays rapid growth at its initial stages followed by a slower progression. As the depth of the monitoring points increases, the change in settlement gradually increases. The settlement of monitoring point SS8 ranged within 40 mm; however, the settlement of monitoring point SS5 reached up to 200 mm. This is because the soil properties change with the increase in depth, for example, the transition from looser surface soil to denser subsurface soil. This change may increase the compressibility of the deeper soil layers, resulting in greater settlement under the same external force. Additionally, the deeper soil layers are subjected to a greater weight from the overlying soil layers (i.e., vertical stress). The increased stress can cause the voids between soil particles to be compressed, thereby making the settlement more serious. Furthermore, settlement is a process that develops over time. In some cases, the settlement of deeper soil layers may gradually increase due to time-dependent effects such as soil consolidation and creep.
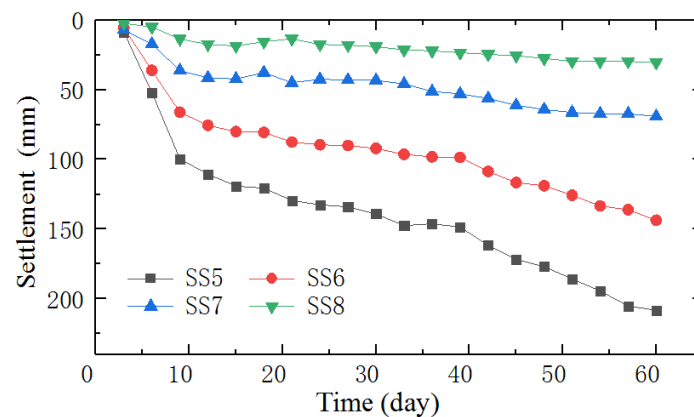
Figure 6. Time variation in the monitored settlement at monitoring points SS5, SS6, SS7, and SS8.

4. Results and Discussions

This section presents the prediction results and intermediate calculation of the proposed model. In addition, the prediction performance of the proposed model is compared to other related methods including the fractional-order GM(1,1) model, the integer-order GM(1,1) model, and fractal theory. The proposed model does not require high computer specifications. A basic configuration, such as Intel Core i5 CPU, 8 GB RAM, and 256 GB SSD, is sufficient.

4.1. Intermediate Calculation and Prediction Results

Since the modelling processes for data at each monitoring point are the same, settlement at monitoring point SS5 was selected as an example to exhibit the intermediate calculation of the proposed model. It was a data sequence containing 20 numbers, wherein 15 of them were selected as fitting data, while the remaining 5 served as testing data.

The initial order was set to 1, to make the fractional-order GM(1,1) model equivalent to the integer-order GM(1,1) model. The initial order 1 was encoded as a binary string to generate the initial population. The genetic algorithm was employed with the control parameters including a population size M of 21, crossover probability P_c of 0.75, and mutation probability P_m of 0.01, and then the evolution was carried out for 50 generations.

The evolution curve with respect to the genetic generations is depicted in Figure 7. The results show that the average relative error reached its minimum (i.e., 5.934) at the 38th generation, corresponding to the optimal fractional order of 0.641.
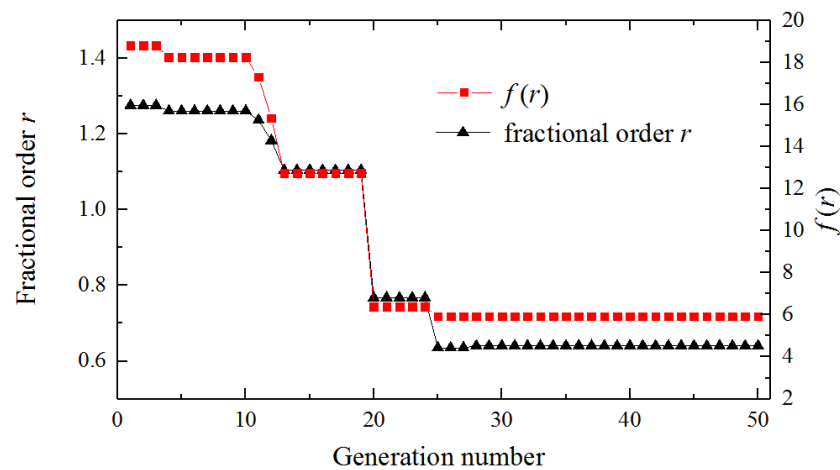
Figure 7. Evolution of fractional order r and f (r).

Calculated by the fractional-order grey model, the time response function can be obtained:

$$\hat{x}^{(0.641)}(k) = -5283.07e^{-0.012(k-1)} + 5291.67.$$

The optimal fractional orders were incorporated into the fractional-order grey model using the real-time tracing algorithm, and then the settlement of the seawall can be predicted. The monitored and predicted settlement at monitoring point SS5 as well as their residual are shown in Table 1. In general, the relative residual between the monitored data and predicted data is fairly small.

Table 1. The monitored and predicted settlement at monitoring point SS5.

Time	Monitored Data	Predicted Data	Residual	Relative Residual
3	8.6	8.6	0	0
6	52.1	65.3	−13.24	−0.254
9	99.6	89.7	9.87	0.099
12	110.5	99.7	10.84	0.098
15	119.1	107.9	11.16	0.094
18	120.6	115.4	5.20	0.043
21	129.5	126.4	3.09	0.024
24	132.5	129.1	3.36	0.025
27	134.0	135.7	−1.73	−0.013
30	139.1	142.2	−3.13	−0.023
33	147.4	148.7	−1.32	−0.009
36	146.1	150.2	−4.12	−0.028
39	148.5	151.8	−3.26	−0.022
42	161.6	163.4	−1.78	−0.011
45	171.6	165.1	6.52	0.038
48	176.9	166.2	10.67	0.060
51	185.7	187.8	−2.12	−0.011
54	194.3	197.9	−3.59	−0.018
57	205.1	203.1	1.99	0.010
60	208.1	210.5	−2.38	−0.011

Figure 8 illustrates the time evolution of the predicted and monitored settlement at each monitoring point. The red points denote the predicted data, and the black points denote the monitored data. This comparison provides valuable insights into the accuracy and reliability of the prediction model. In detail, the data sequence at each monitoring point comprises 20 data points. Among these, the first 15 data points were chosen for fitting the model, meaning they were used to calibrate the model's parameters and ensure it accurately reflects the monitored settlement. The remaining 5 data points, set aside as testing data,

served to validate the model's predictive capabilities on unseen data. It can be seen from the figure that the monitored data aligns remarkably well with the predicted data, not only for the fitting dataset but also for the testing dataset. This strong correlation indicates that the prediction model possesses a high degree of accuracy and is capable of generalizing well to unknown data. By comparing the predicted values with the actual monitored values, we can evaluate the accuracy of our prediction model and understand the relation between different monitoring points. The monitored settlement at monitoring points SS5 and SS6 over time shows a smooth curve. The monitored settlement at monitoring points SS7 and SS8 displays different characteristics compared to SS5 and SS6, where the fluctuations during day 5 to day 20 are more obvious. The specific reasons are analyzed as follows: The primary reason for the settlement of the seawall situated on soft soil foundations is the step-by-step loading during the construction of the embankment. The loading causes the gradual expulsion of water and air from between the soil particles, reducing the voids in the soft soil. As a result, the entire foundation undergoes systematic deformation. The surface layer of the seawall foundation consists of about 50 cm of silt. Unlike the underlying soft soil, the silt has a very high water content and low structural strength. During the initial half-month period of the project, the silt would have been squeezed out to the sides along with the construction loading process, resulting in significant instantaneous settlement. Due to construction constraints, the monitoring instruments took samples with an interval of 3 days. This resulted in poor trend consistency in the collected settlement data, making it difficult for current models to accurately fit the data from day 5 to day 20.
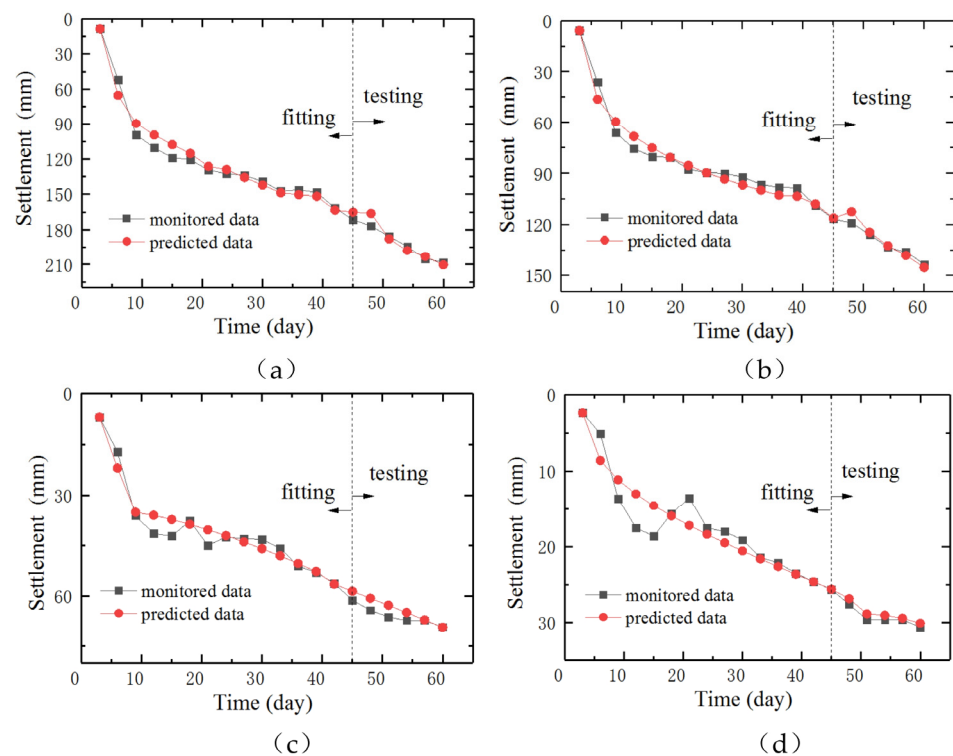
Figure 8. Time evolution of the predicted and monitored settlement at the selected monitoring points: (**a**) SS5, (**b**) SS6, (**c**) SS7, (**d**) SS8.

Figure 9 illustrates the time evolution of the relative residual of the predicted and monitored settlement at each monitoring point. The relative residual is below 0.1 for all four monitoring points, which suggests that the model is capable of capturing the general trends and behaviors of the settlement process, regardless of specific local conditions. This global accuracy is crucial for providing a comprehensive understanding of the settlement phenomenon and for making informed decisions regarding potential risks and mitigation measures. In addition, the relative residual exhibits some fluctuations at the initial stage.

The fluctuations in the relative residual during the initial stage likely indicates that the model is still adjusting to the initial conditions and data. This could be due to the inherent uncertainty in the initial measurements or the model's parameters being set to approximate values. Over time, as more data is collected and the model is able to learn from these observations, the fluctuations tend to decrease, indicating the model's ability to adapt and improve its predictions. Even though the evolution curve of the relative residual of these four monitoring points are similar globally, the fluctuations of monitoring points SS5 and SS6 are more obvious than those of monitoring points SS7 and SS8. The relative residual globally reflects the prediction accuracy. Prediction accuracy for seawall settlement can vary at different monitoring points due to several factors related to both the characteristics of the site and the limitations of the prediction model. Different monitoring points are located on varying soil types, each with unique properties such as compressibility and permeability. These properties significantly influence settlement behavior. In addition, the distribution of loads on the seawall, including static and dynamic loads from waves, tides, and human activity, may differ across monitoring points, affecting settlement rates. Thus, the historical data used for training the model vary in quality and quantity across different monitoring points, affecting the model's ability to learn accurate patterns. For the model's prediction ability itself, as the interactions between different influencing factors (e.g., soil properties, environmental conditions) are complex and non-linear, the present model may not capture these interactions accurately.
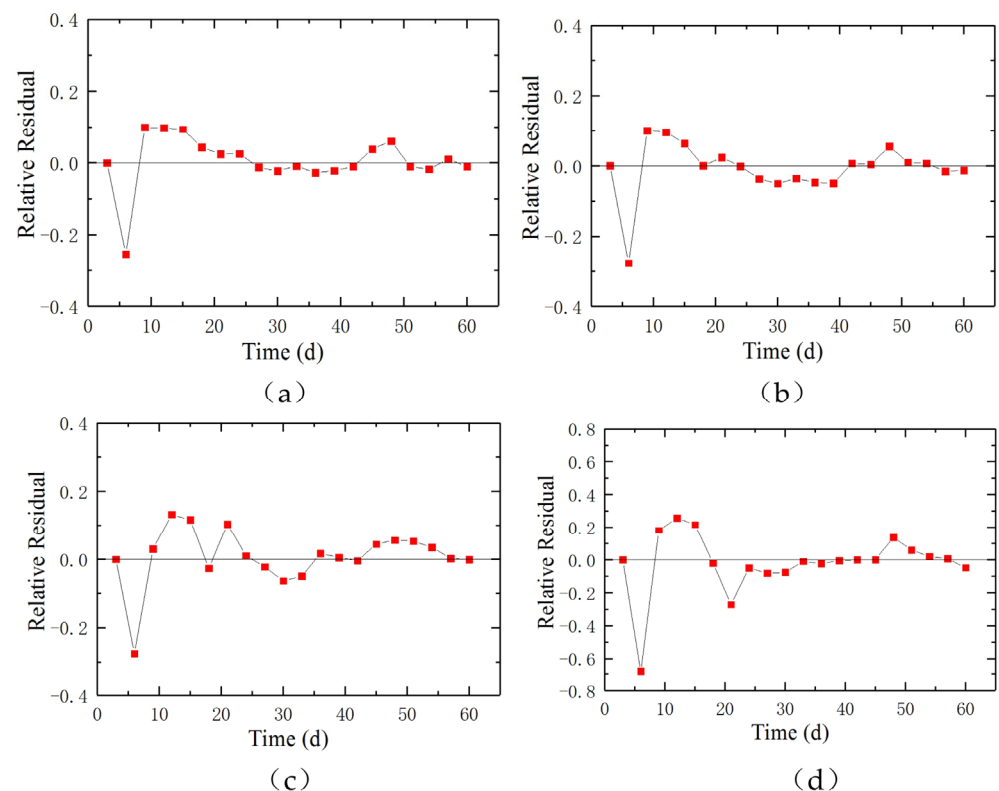
Figure 9. Time evolution of the relative residual of the predicted and monitored settlement at each monitoring points: (**a**) SS5, (**b**) SS6, (**c**) SS7, (**d**) SS8.

4.2. The Evaluation Criteria of Prediction Performance

The prediction performance of the prediction model is evaluated by calculating the coefficient of determination (R^2), the mean relative error (MRE), the mean squared error (MSE), and the root mean squared error (RMSE) of the predicted results. These indexes express the extent of prediction error by measuring the deviation between the predicted

data α_i and the monitored data β_i. The expression of MRE is shown in equation 15, and the expressions of R^2, MSE, and RMSE are as follows:

$$R^2 = 1 - \frac{\sum\limits_{i=1}^{n} (\alpha_i - \beta_i)}{\sum\limits_{i=1}^{n} (\alpha_i - \overline{\beta_i})}, \tag{16}$$

$$MSE = \frac{1}{n}\sum\limits_{i=1}^{n} (\alpha_i - \beta_i)^2, \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n}\sum\limits_{i=1}^{n} (\alpha_i - \beta_i)^2} \tag{18}$$

Table 2 exhibits the evaluation criteria including the R^2, MRE, MSE, and RMSE of the proposed model at each monitoring point. The R^2, which is a measure of how much the data points are spread out from their mean, is greater than 0.8 for both the testing dataset and the fitting dataset, which implies the model is robust and can be validated. The high R^2 suggests that the model explains a large portion of the variation in the dependent variable, indicating a good fit. The low values of MRE, MSE, and RMSE further confirm the model's accuracy in predicting the target variable. These evaluation criteria collectively demonstrate the model's reliability and effectiveness.

Table 2. The evaluation criteria of the proposed model at each monitoring point.

Monitoring Points	Fitting Data				Testing Data			
	R^2	MRE	MSE	RMSE	R^2	MRE	MSE	RMSE
SS5	0.974	5.577	46.452	6.585	0.820	0.779	2.791	4.999
SS6	0.971	5.671	22.731	4.606	0.853	1.140	2.603	3.261
SS7	0.977	0.091	12.755	3.570	0.853	0.759	0.501	0.952
SS8	0.878	13.432	5.292	2.222	0.818	1.730	0.309	0.606

4.3. Comparison of the Proposed Model with Other Relevant Models

To better evaluate the performance of the proposed model, we compared its prediction accuracy with the fractional-order GM(1,1), integer-order GM(1,1), and fractal theory model. The integer-order GM(1,1) is a widely used grey model which relies on the principle of differential equations and exponential smoothing to generate predictions. The fractal theory model provides a framework for modeling complex processes that exhibit self-similarity across different scales, so as to capture the inherent hierarchical patterns in the data sequence. The fractional-order GM(1,1) is a generalization of the traditional integer-order grey model, which introduces the concept of fractional-order calculus to improve the flexibility and accuracy of the model. By allowing the order of differentiation and integration to be non-integer, the fractional-order model can better capture complex and non-linear dynamics in data.

Figure 10 shows the time evolution of fitting and predicting results of seawall settlement at the selected monitoring points using these four different models. Firstly, it is noteworthy that all these four models demonstrate a certain level of accuracy in fitting the historical data and predicting future seawall settlement trends, the curve trend of the predicted value aligns precisely with the actual curve. This indicates that the predicted curve generated through each model is rational and logical, and each model has captured essential elements of the system dynamics.
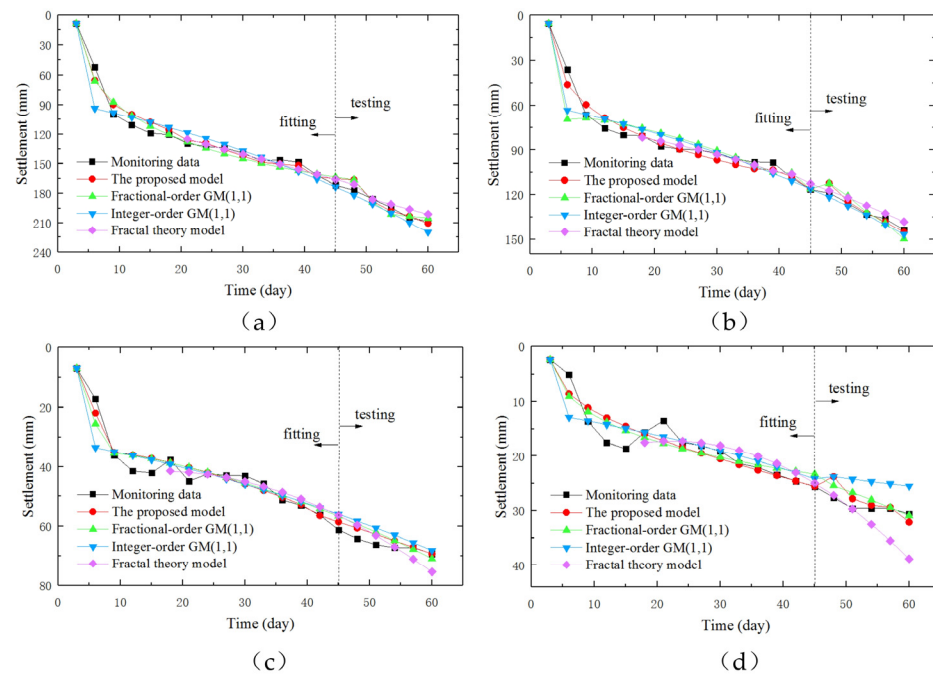
Figure 10. Time evolution of fitting and predicting results of seawall settlement at the selected monitoring points using the four selected models: (**a**) SS5, (**b**) SS6, (**c**) SS7, (**d**) SS8.

Figure 11 shows the time evolution of the relative residual of the four models for each monitoring point. The relative residuals of all four models are very small. Most are located within the range of −0.1 to 0.1. It can be seen that the proposed model shows a tighter fit to the monitoring data compared with the other three models, indicating a higher degree of accuracy. This can be evidenced by a narrower confidence interval or a high coefficient of determination between the predicted and monitored data. In addition, for all four prediction models, the relative residual exhibits some fluctuations at the initial stage. The relative residual refers to the difference between the predicted value and the actual value, normalized or scaled in a way that makes it easier to compare across different data points or time periods. These fluctuations in the relative residual at the initial stage may result from several reasons: First, the initial period does not have enough data points to accurately capture the trends, which may lead to higher errors and fluctuations in the relative residuals. Second, the models need initialization before they can start making accurate predictions. During this initial phase, the models might not be fully calibrated. Third, there might be external factors, such as measurement errors, which introduce noise and uncertainty into the predictions.

When discussing the performance of prediction models, it is important to consider not only how the models behave at different stages but also the overall accuracy. Figure 12 compares the coefficients of determination R^2 of the fitting and testing data for these four models, which represent an overall evaluation of the prediction performance. Using the datasets of all the monitoring points, the proposed model outperforms the other three models. The comparisons demonstrate that the proposed model exhibits superior predictive power, especially in capturing the nonlinearities and complexities of seawall settlement over time under the situation of limited data volume. Seawall settlement is affected by many factors including soil properties and hydrological conditions. A model that can accurately predict such complex processes is valuable for risk management. In addition, the R^2 calculated using fitting data is higher than that calculated using testing data for all these four models. It is a common observation that when the model is evaluated on testing data, which is independent of the training data, its performance tends to be lower. This is because the testing data contains additional uncertainty and variability that the model has not encountered during training.
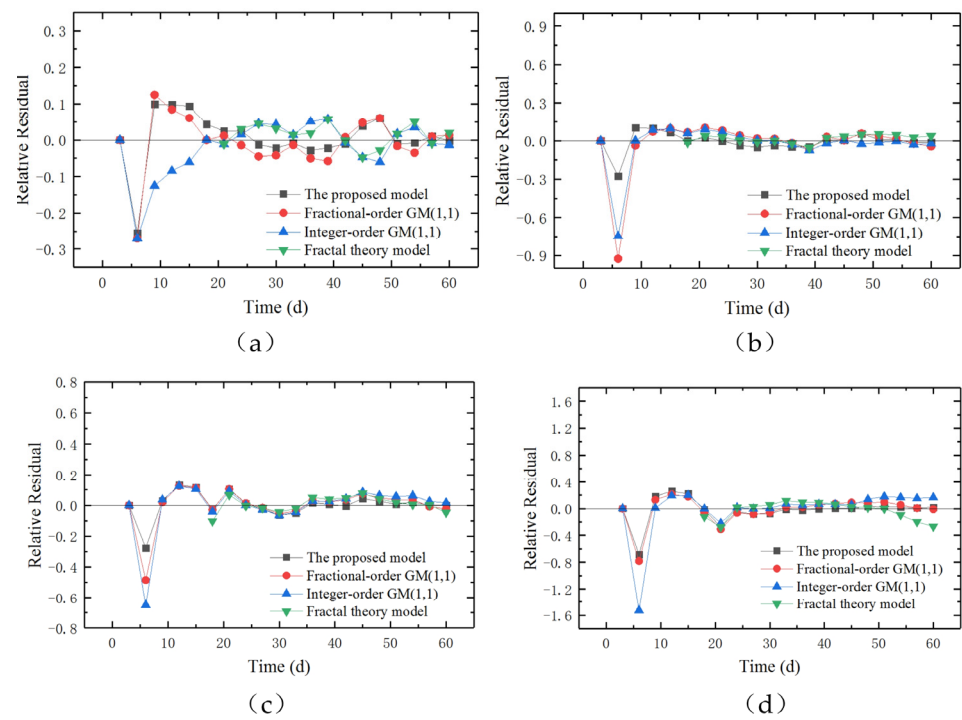
Figure 11. Time evolution of the relative residual of seawall settlement at the selected monitoring points using the four selected models: (**a**) SS5, (**b**) SS6, (**c**) SS7, (**d**) SS8.
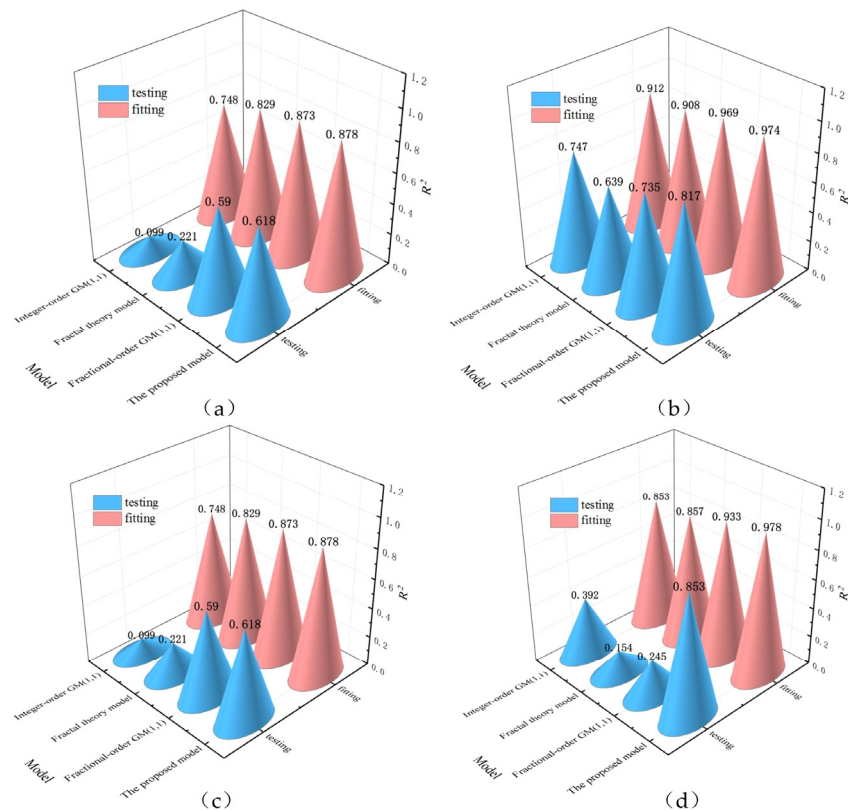
Figure 12. Comparison of the R^2 of the fitting and testing data at each monitoring point: (**a**) SS5, (**b**) SS6, (**c**) SS7, (**d**) SS8 using the four selected models.

Figure 13 compares the MRE of the proposed model with the fractional-order GM(1,1), integer-order GM(1,1), and fractal theory model. A lower MRE indicates higher accuracy

and better performance of the model. The proposed model exhibits a lower MRE than the other models for all monitoring points. This suggests that it is able to more accurately capture the underlying dynamics in the data sequence, leading to more reliable and accurate predictions.
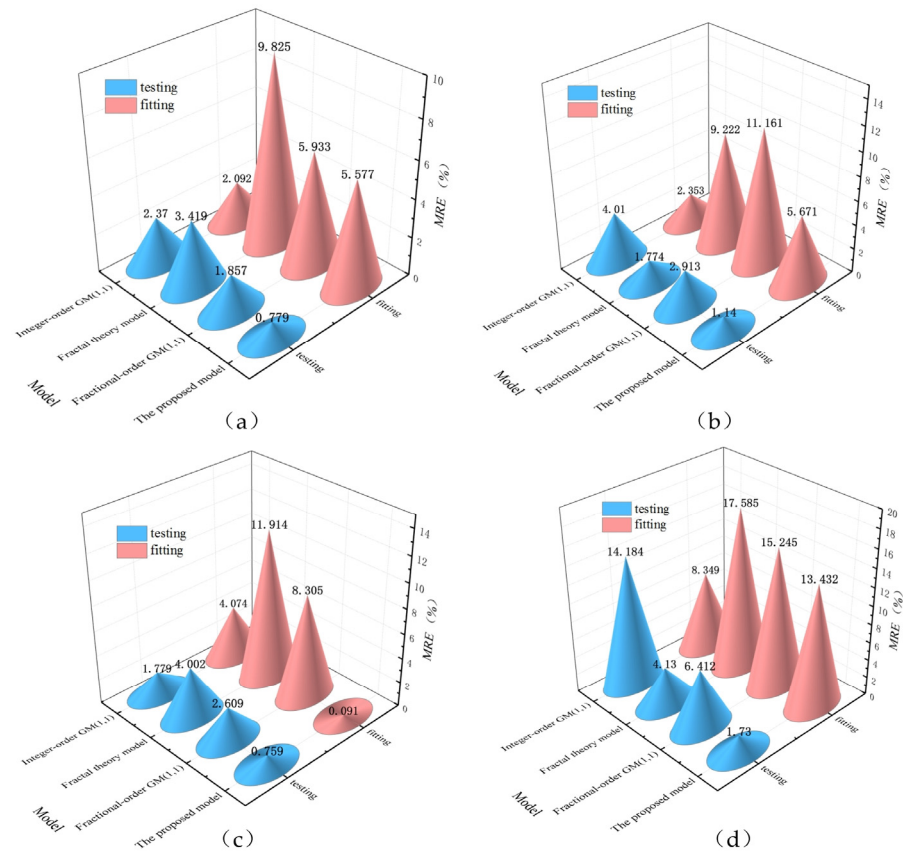
Figure 13. Comparison of the MRE of the fitting and testing data at each monitoring point: (**a**) SS5, (**b**) SS6, (**c**) SS7, (**d**) SS8 using the four selected models.

5. Conclusions

As the performance of the fractional-order grey model is easily affected by the inappropriate setting of the fractional order, and considering the limited data volume and the nonlinear characteristics of seawall settlement monitoring data, this paper proposed an improved fractional-order grey model for seawall settlement prediction by integrating a genetic algorithm and the real-time tracing algorithm. First, the genetic algorithm with enhanced search capabilities was employed to solve the premature convergence problem. Then, to solve the problem of the fractional-order grey model associated with fixed step sizes, the real-time tracing algorithm was introduced to conduct equal-dimensionally recursive calculation. The proposed model was validated using monitoring data of the Haiyan seawall in Zhejiang province, China. The prediction performance of the proposed model was compared with those of the fractional-order GM(1,1), integer-order GM(1,1), and fractal theory model, demonstrating its superior performance compared to other methods such as the fractional-order GM(1,1) model, the integer-order GM(1,1) model, and fractal theory model.

The proposed model effectively predicts seawall settlement trends and can be extended to address prediction challenges in other fields with limited data volume. The significance of this finding lies in its practical implications. A reliable settlement prediction model is crucial in various engineering applications, such as construction projects, where precise forecasting of ground movements is essential for safety and stability. The high degree of fit between the predicted and monitored data suggests that the model can be trusted to

provide accurate settlement predictions, thereby enhancing the safety and efficiency of related projects.

While the current model shows promising results, there are still opportunities for improvement. For example, its relatively high time complexity poses some disadvantages in terms of computation time (computational complexity of the proposed model with other models are shown in Table 3). Additionally, the model could be further refined by incorporating additional influencing factors, such as soil properties, groundwater conditions, or changes in loading conditions. Also, this paper took multi-point monitoring data into account in the prediction model, which helps to provide a comprehensive understanding of the trends of seawall settlement comparing to single-point prediction. However, one disadvantage of the current study is that the correlation among different monitoring points is not entered. Considering the correlation among each monitoring point would be interesting for further studies.

Table 3. Computational complexity of the proposed model with other models.

Models	The Proposed Model	The fractional-Order GM(1,1)	Integer-Order GM(1,1))	Fractal Theory
Variable number of parameters	6 (including fractional-order r, population size M, crossover probability P_c, mutation probability P_m, typical number of iterations, predicted length m)	1 (fractional-order r)	0	1 (including fractal dimension D)
Calculating time	12 s	3 s	1 s	5 s

Author Contributions: Conceptualization, P.Q.; data curation, Z.M.; formal analysis, C.C. and H.S.; funding acquisition, P.Q. and H.S.; investigation, C.C.; methodology, S.Z. and P.Q.; project administration, C.D.; resources, Z.M.; software, H.S.; supervision, P.Q.; validation, P.Q.; visualization, S.Z.; writing—original draft, P.Q.; writing—review and editing, P.Q. and Z.M. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Su, H.; Qin, P.; Qin, Z. A method for evaluating sea dike safety. *Water Resour. Manag.* **2013**, *27*, 5157–5170. [CrossRef]
2. Qin, P.; Jiang, Y.; Cheng, C. How Do Structural Safety, Ecological Functions and Social Development Influence Construction of Ecological Seawalls for Coastal Protection and Sustainability? *Water Resour. Manag.* **2024**, *38*, 1807–1824. [CrossRef]
3. Feng, W.Q.; Yin, J.H.; Chen, W.B.; Tan, D.Y.; Wu, P.C. A new simplified method for calculating consolidation settlement of multi-layer soft soils with creep under multi-stage ramp loading. *Eng. Geol.* **2020**, *264*, 105322. [CrossRef]
4. Bayesteh, H.; Mansouriboroujeni, R. Mechanisms of settlement of a rubble mound breakwater on a soft soil in tidal flats. *Mar. Georesour. Geotechnol.* **2020**, *38*, 1163–1176. [CrossRef]
5. Wu, X.T.; Liu, J.N.; Xie, Z.M. Deformation and Strength Characteristics of Marine Soft Soil Treated by Prefabricated Vertical Drain-Assisted Staged Riprap under Seawall Construction. *Buildings* **2023**, *13*, 2322. [CrossRef]
6. Meng, Z.; Wang, Y.; Zheng, S.; Wang, X.; Liu, D.; Zhang, J.; Shao, Y. Abnormal Monitoring Data Detection Based on Matrix Manipulation and the Cuckoo Search Algorithm. *Mathematics* **2024**, *12*, 1345. [CrossRef]
7. Blázquez-García, A.; Conde, A.; Mori, U.; Lozano, J.A. A review on outlier/anomaly detection in time series data. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 1–33. [CrossRef]
8. Ciaburro, G.; Iannace, G. Machine learning-based algorithms to knowledge extraction from time series data: A review. *Data* **2021**, *6*, 55. [CrossRef]
9. Zheng, S.; Gu, C.; Shao, C.; Hu, Y.; Xu, Y.; Huang, X. A novel prediction model for seawall deformation based on CPSO-WNN-LSTM. *Mathematics* **2023**, *11*, 3752. [CrossRef]

10. Ying, H.W.; Huang, Z.J.; Ge, H.B.; Shen, H.W.; Wang, K.L.; Gong, X.N. Curve-fitting method for settlement based on staged loading condition and its engineering application. *J. Southeast Univ. (Nat. Sci. Ed.)* **2021**, *51*, 300–305.

11. Spross, J.; Larsson, S. Probabilistic observational method for design of surcharges on vertical drains. *Géotechnique* **2021**, *71*, 226–238. [CrossRef]

12. Fan, H.; Chen, Z.; Shen, J.; Cheng, J.; Chen, D.; Jiao, P. Buckling of steel tanks under measured settlement based on Poisson curve prediction model. *Thin-Walled Struct.* **2016**, *106*, 284–293. [CrossRef]

13. Sitharam, T.G.; Kolathayar, S. Geotechnical considerations for coastal reservoirs. *Sustain. Water Resour. Dev. Using Coast. Reserv.* **2020**, 61–83. [CrossRef]

14. Zhu, M.; Li, S.; Wei, X.; Wang, P. Prediction and stability assessment of soft foundation settlement of the fishbone-shaped dike near the estuary of the Yangtze River using machine learning methods. *Sustainability* **2021**, *13*, 3744. [CrossRef]

15. Pinsky, M.A. *Introduction to Fourier Analysis and Wavelets*; American Mathematical Society: Providence, RI, USA, 2023; Volume 102.

16. Khodadadi, N.; Abualigah, L.; Al-Tashi, Q.; Mirjalili, S. Multi-objective chaos game optimization. *Neural Comput. Appl.* **2023**, *35*, 14973–15004. [CrossRef]

17. Taye, M.M. Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. *Computation* **2023**, *11*, 52. [CrossRef]

18. Chen, H.; Huang, S.; Xu, Y.P.; Teegavarapu, R.S.V.; Guo, Y.; Nie, H.; Xie, H. Using baseflow ensembles for hydrologic hysteresis characterization in humid basins of Southeastern China. *Water Resour. Res.* **2024**, *60*, e2023WR036195. [CrossRef]

19. Liao, K.W.; Hsiung, H.J. Reliability-based safety analysis of a seawall with incomplete information on tsunamis. *Coast. Eng. J.* **2017**, *59*, 1750014. [CrossRef]

20. Huang, J.X.; Qu, K.; Li, X.H.; Lan, G.Y. Performance evaluation of seawalls in mitigating a real-world tsunami wave using a nonhydrostatic numerical wave model. *J. Mar. Sci. Eng.* **2022**, *10*, 796. [CrossRef]

21. Nunn, P.D.; Klöck, C.; Duvat, V. Seawalls as maladaptations along island coasts. *Ocean Coast. Manag.* **2021**, *205*, 105554. [CrossRef]

22. Trimble, S.; Houser, C. Seawalls and signage: How beach access management affects rip current safety. *Beach Manag. Tools-Concepts Methodol. Case Stud.* **2018**, *24*, 497–524.

23. Amini, E.; Marsooli, R.; Ayyub, B.M. Assessing Beach–Seawall Hybrid Systems: A Novel Metric-Based Approach for Robustness and Serviceability. *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.* **2024**, *10*, 04023062. [CrossRef]

24. Li, X.; Liu, X.; Li, C.Z.; Hu, Z.; Shen, G.Q.; Huang, Z. Foundation pit displacement monitoring and prediction using least squares support vector machines based on multi-point measurement. *Struct. Health Monit.* **2019**, *18*, 715–724. [CrossRef]

25. Cao, W.; Wen, Z.; Feng, Y.; Zhang, S.; Su, H. A Multi-Point Joint Prediction Model for High-Arch Dam Deformation Considering Spatial and Temporal Correlation. *Water* **2024**, *16*, 1388. [CrossRef]

26. Deng, J. Introduction to grey system theory. *J. Grey Syst.* **1989**, *1*, 1–24.

27. Liu, S.; Forrest, J.Y.L. *Grey systems: Theory and Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2021.

28. Javanmardi, E.; Liu, S. Exploring grey systems theory-based methods and applications in analyzing socio-economic systems. *Sustainability* **2019**, *11*, 4192. [CrossRef]

29. Shan, X.; Cao, Y. Forecasting guangdong's marine science and technology, marine economy, and employed persons by coastal regions—Based on rolling grey MGM (1, m) model. *Water* **2022**, *14*, 824. [CrossRef]

30. Tao, L.; Liang, A.; Xie, N.; Liu, S. Grey system theory in engineering: A bibliometrics and visualization analysis. *Grey Syst. Theory Appl.* **2022**, *12*, 723–743. [CrossRef]

31. Hu, M.; Liu, W. Grey system theory in sustainable development research—A literature review (2011–2021). *Grey Syst. Theory Appl.* **2022**, *12*, 785–803. [CrossRef]

32. Wu, L.F.; Liu, S.F.; Cui, W.; Liu, D.L.; Yao, T.X. Non-homogenous discrete grey model with fractional-order accumulation. *Neural Comput. Appl.* **2014**, *25*, 1215–1221. [CrossRef]

33. Wu, L.; Liu, S.; Fang, Z.; Xu, H. Properties of the GM (1, 1) with fractional order accumulation. *Appl. Math. Comput.* **2015**, *252*, 287–293. [CrossRef]

34. Liu, J.; Liu, S.; Fang, Z. Fractional-order reverse accumulation generation gm (1, 1) model and its applications. *J. Grey Syst.* **2015**, *27*, 52.

35. Zeng, B.; Liu, S. A self-adaptive intelligence gray prediction model with the optimal fractional order accumulating operator and its application. *Math. Methods Appl. Sci.* **2017**, *40*, 7843–7857. [CrossRef]

36. Xia, L.; Ren, Y.; Wang, Y. Forecasting China's total renewable energy capacity using a novel dynamic fractional order discrete grey model. *Expert Syst. Appl.* **2024**, *239*, 122019. [CrossRef]

37. Yuan, Y.; Zhao, H.; Yuan, X.; Chen, L.; Lei, X. Application of fractional order-based grey power model in water consumption prediction. *Environ. Earth Sci.* **2019**, *78*, 266. [CrossRef]

38. Qin, P.; Cheng, C. Prediction of seawall settlement based on a combined LS-ARIMA model. *Math. Probl. Eng.* **2017**, *2017*, 7840569. [CrossRef]

39. Meng, W.; Zeng, B.; Li, S. A novel fractional-order grey prediction model and its modeling error analysis. *Information* **2019**, *10*, 167. [CrossRef]

40. Katoch, S.; Chauhan, S.S.; Kumar, V. A review on genetic algorithm: Past, present, and future. *Multimed. Tools Appl.* **2021**, *80*, 8091–8126. [CrossRef]

41. Qin, P.; Su, H.Z.; Shen, Y.J. Prediction of Seawall Settlement Based on ARIMA-RTA model. *Hydro-Sci. Eng.* **2013**, *11*, 66–70.

42. Zhang, J.; Qin, Y.; Zhang, X.; Che, G.; Sun, X.; Duo, H. Application of non-equidistant GM (1, 1) model based on the fractional-order accumulation in building settlement monitoring. *J. Intell. Fuzzy Syst.* **2022**, *42*, 1559–1573. [CrossRef]
43. Yang, H.; Gao, M.; Xiao, Q. A novel fractional-order accumulation grey power model and its application. *Soft Comput.* **2023**, *27*, 1347–1365. [CrossRef]
44. Mirjalili, S. *Evolutionary Algorithms and Neural Networks: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 780.
45. Hamdia, K.M.; Zhuang, X.; Rabczuk, T. An efficient optimization approach for designing machine learning models based on genetic algorithm. *Neural Comput. Appl.* **2021**, *33*, 1923–1933. [CrossRef]
46. Sohail, A. Genetic algorithms in the fields of artificial intelligence and data sciences. *Ann. Data Sci.* **2023**, *10*, 1007–1018. [CrossRef]
47. Avazzadeh, Z.; Nikan, O.; Nguyen, A.T. A localized hybrid kernel meshless technique for solving the fractional Rayleigh–Stokes problem for an edge in a viscoelastic fluid. *Eng. Anal. Bound. Elem.* **2023**, *146*, 695–705. [CrossRef]