The Improved Conformable Fractional Grey Model Using ADMM and HOA Algorithm

Huiyan Xu [1,*,#], Derui Wang [2,#], Yizhuo Wang [3,#]

[1] Department of Mathematics, Jilin University, Changchun, China, 130012

[2] Department of Mathematics, Northeastern University, Shenyang, China, 110801

[3] Department of Mathematics, Dalian University of Technology, Dalian, China, 116024

* Corresponding Author Email: xhy1823742034@163.com

[#]These authors contributed equally.

Abstract. The grey model (GM) is a forecasting model known for its ability to provide accurate predictions for small sample data, but it struggles with complex nonlinear time series. This paper introduces a fractional calculus prediction model, utilizing Conformable Fractional Difference (CFD) and Conformable Fractional Accumulation (CFA) to calculate the new time series. The traditional least squares method for parameter estimation is transformed into a Least Absolute Shrinkage and Selection Operator (LASSO) problem and is solved using the dual form of the Alternating Direction Method of Multipliers (ADMM). Additionally, the Hiking Optimization Algorithm (HOA) is employed to optimize the fractional order parameter. As a result, the improved model demonstrates greater accuracy than the original model when dealing with small sample data.

Keywords: CFGM Model, LASSO Regression, ADMM, HOA Algorithm.

1. Introduction

GM plays a key role in the theory of the grey system pioneered by Deng Julong in 1982. The main idea of the grey prediction was originally proposed by Deng in 1983, and the first grey model was elicited by Deng in 1984 [1]. As an emerging cross-sectional discipline, GM has been greatly developed in just twenty years. It has become one of the most important methods of forecasting, decision-making, evaluating, planning, controlling, and modeling in many fields, such as society, economy [2], education, science, and technology [3]. GM is the algorithm specially designed to solve the prediction problem in the case of incomplete data information. Compared with other different prediction methods, GM does not require a large amount of raw data, instead, it is effective in generating accurate predictions with limited or unreliable data. For many practical situations to predict, traditional GM (such as GM (1,1)) works in some cases but performs less than expected when dealing with complex, non-linear time series.

To solve this problem, many scholars innovate and improve on the basis of GM, such as the multivariate grey model of dynamic background value method [4], or improvements of the traditional least squares method to a more general nonlinear regression for computational prediction [5]. It is notable that in GM, the raw data is often simply added when calculating the new time series. For this aspect, the flexibility and adaptability of the prediction model could be enhanced by introducing the fractional calculus-based prediction model. CFD [6] is a difference operation based on the conformable fractional derivative, and it is mainly used to improve the performance of gray system models. CFD allows the handling of heterogeneous and irregular changes in time series data. Compared to traditional first-order differential methods, CFD provides more flexibility to capture subtle changes in the data by introducing the concept of fractional order. CFD is also suitable for nonlinear systems and can better reflect the complexity in practical problems. CFA was introduced in the cumulative calculations after the introduction of CFD. CFA is a cumulative operation based on the conformable fractional derivative, used to describe the cumulative process of time series data. By introducing the concept of fractional order, CFA provides a flexible modeling ability for the data accumulation process to handle different types of time series. CFA is also able to capture more effectively the dynamic changes in the data, especially in the nonlinear and complex data patterns. According to the above-improved algorithm, we build a conformable fractional grey model (CFGM), which is based on the traditional GM and can be used to predict the sample data with a small sample size.

The least squares method is commonly used to estimate the parameters in the traditional CFGM model. However, in order to extract the features and get the sparse solution of the original problem, we transform it into a LASSO regression problem which adds l_1 regular term in the objective function. As an important algorithm in prediction models, regression has many research and application fields, such as linear regression, logistic regression, and ridge regression. LASSO regression can achieve a higher accuracy of model prediction and enhance the model's generalizability with the cost of a certain estimation bias [7].

There are many related algorithms for solving LASSO problems, we consider ADMM to solve this problem. The motivation for choosing ADMM is its ability to solve convex optimization problems by breaking them down into some smaller sub-problems so that each of which is easier to deal with [8].

The HOA was proposed by David H. Wolpert and William Macready [9], which was inspired by hiking. Outdoor travel is characterized by completely random terrain. A team in the movement, according to the member's familiarity with the terrain and physical condition updates the leader in real-time, to ensure that the team can reach the destination in the fastest and most effective. This algorithm adopts the idea of constantly checking the "terrain" to adjust the "leader" and "speed", in order to obtain the optimal solution. In the model constructed in this paper, the HOA algorithm is used to calculate the optimal fractional order parameter.

Some of the following key innovations are demonstrated in this paper:

(1) This paper enhances the least squares method for solving the parameters of the CFGM model by incorporating a regularization term to identify sparse solutions.

(2) In determining the order of the CFGM model, the HOA algorithm is employed as a replacement for the traditional brute force method in order to select the optimal solution, thereby improving model accuracy.

The rest of this paper is organized as follows: Section 2 gives an introduction of the definition of the CFD and CFA and presents the modeling procedures of the new CFGM. On the basis of the linear regression least square method, the LASSO problem obtained by adding l_1 regular term transformation is described in Section 3, which is solved by the dual form of ADMM. In Section 4, the HOA algorithm is introduced and the parameters of the CFGM model are optimized. In Section 5, the case studies of predicting raw data from the data of power production are shown, and the conclusions and perspectives are shown in Section 6.

2. The Conformable Fractional Grey Model (CFGM)

2.1. CFD and CFA

According to the traditional GM, the definitions of two conformable fractional orders, CFD and CFA can be introduced into the GM. Firstly, there is the raw data $\mathrm{X}^{(0)} = (x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n))$, where n is the number of data points contained in the original data. According to the definition, the traditional data accumulation process is improved to use the proposed summing formula using CFD and CFA at order α to obtain the accumulation data $X^{(\alpha)}$. According to the differential equation of CFGM, the parameters $\hat{a}$ and $\hat{b}$ can be calculated by the least squares method. After getting the results of the parameters, the whitening equation passes through the response function, and we can get the predicted accumulation data $\widehat{X^{(\alpha)}}$. Thus, the prediction of the raw data $\widehat{X^{(0)}}$ can be obtained.

Definition 1 [3] If we have the $\boldsymbol{X}^{(\mathbf{0})} = (x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n))$, where n is the number of data points contained in the original data. By using CFA to get the formulation of accurate α - order sequence data $\boldsymbol{X}^{(\boldsymbol{\alpha})} = (x^{(\alpha)}(1), x^{(\alpha)}(2), ..., x^{(\alpha)}(n))$, the cumulative formula can be

$$x^{(\alpha)}(k) = \nabla^{\alpha}\mathrm{x}^{(0)}(k) = \begin{cases} \sum_{i=1}^{k}\frac{1}{i^{1-\alpha}}x^{(0)}(i) & 0<\alpha\leq 1 \\ \sum_{i=1}^{k}x^{(\alpha-1)}(i) & \alpha>1 \end{cases} \tag{1}$$

2.2. CFGM Model

The Response function of CFGM [3] is

$$x^{(\alpha)}(t) = (x^{(0)}(1) - \frac{b}{a})e^{-a(t-1)} + \frac{b}{a} . \tag{2}$$

The Differential equation for the CFGM [3] is

$$\frac{dx^{(\alpha)}}{dt} + ax^{(\alpha)} = b. \tag{3}$$

Its discrete form is

$$\left(x^{(\alpha)}(k) - x^{(\alpha)}(k-1)\right) + \frac{a}{2}\left(x^{(\alpha)}(k) - x^{(\alpha)}(k-1)\right) = b \tag{4}$$

2.2.1. Data preparation and calculation

Let the raw data be $\boldsymbol{X}^{(\mathbf{0})}= (x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n))$, where n is the number of data points contained in the original data. By using CFA to get the formulation of accurate α - order sequence data $\boldsymbol{X}^{(\boldsymbol{\alpha})}= (x^{(\alpha)}(1), x^{(\alpha)}(2), ..., x^{(\alpha)}(n))$, the cumulative formula can be

$$x^{(\alpha)}(t) = \nabla^{\alpha}x^{(0)}(t) = \begin{cases} \sum_{i=1}^{t}\frac{1}{i^{1-\alpha}}x^{(0)}(i) & 0<\alpha\leq 1 \\ \sum_{i=1}^{t}x^{(\alpha-1)}(i) & \alpha>1 \end{cases} . \tag{5}$$

2.2.2. Differential equation for CFGM

Based on the α - order sequence data $X^{(\alpha)}$, the differential equation for the CFGM can be built as the differential equation.

2.2.3. Estimation of the parameters

The least square method is estimated to solve the parameters. The parameters $\hat{a}$ and $\hat{b}$ can be easily calculated by

$$[\hat{a},\hat{b}]^{T} = (B^{T}B)^{-1}B^{T}Y, \tag{6}$$

Where

$$B=\begin{bmatrix} -\frac{1}{2}(x^{\alpha}(2)-x^{\alpha}(1)) & 1 \\ -\frac{1}{2}(x^{\alpha}(3)-x^{\alpha}(2)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x^{\alpha}(n)-x^{\alpha}(n-1)) & 1 \end{bmatrix} \text{ and } \quad Y=\begin{bmatrix} x^{\alpha}(2)-x^{\alpha}(1) \\ x^{\alpha}(3)-x^{\alpha}(2) \\ \vdots \\ x^{\alpha}(n)-x^{\alpha}(n-1) \end{bmatrix}.$$

2.2.4. Return of the predicted data

According to Definition 4, the predicted α - order sequence data $\widehat{x_{\alpha}}(t)$ can be expressed as

$$\widehat{x^{(\alpha)}}(t) = (x^{(0)}(1) - \frac{\hat{b}}{\hat{a}})e^{-\hat{a}(t-1)} + \frac{\hat{b}}{\hat{a}} . \tag{7}$$

So the predicted value of the raw data $\widehat{x^{(0)}}(k) = \Delta^{\alpha}\widehat{x^{(\alpha)}}(k)$, when $\alpha \in (0,1]$, $\widehat{x^{(0)}}(k) = k^{1-\alpha}(\widehat{x^{(\alpha)}}(k) - \widehat{x^{(\alpha)}}(k-1))$.

3. The proposed CFGM-L1 Model

In this section, we first consider the LASSO regression problem and its dual form based on the Conformable Fractional Grey Model and obtain its optimal solution using the Alternating Direction Method of Multipliers.

3.1. Linear regression and Lasso regression

Considering the linear equation

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}, E(\boldsymbol{\varepsilon}) = 0, Cov(\boldsymbol{\varepsilon}) = \sigma^2 \boldsymbol{I}, \tag{8}$$

Where y is a $n \times 1$ observation vector, X is a $n \times p$ matrix, β is a $p \times 1$ unknown parameter vector, ε is a random error, σ^2 is the variance of the error. In order to estimate the parameter β, we usually use the least square method, that is to minimize the error vector $\boldsymbol{\varepsilon}$, i.e.

$$min\, Q(\beta) = min\, ||y - X\beta||^2. \tag{9}$$

We can get the solution

$$\hat{\beta} = (X'X)^{-1}X'y. \tag{10}$$

In the high dimensional space, the traditional least square regression has some problems, such as difficulties in selecting variables and over-fitting models.

LASSO regression aims to identify the variables and corresponding regression coefficients that lead to a model that minimizes the prediction error.

When the number of features p of the data set is greater than the total number of samples n, the solution of the problem is not unique, and the solution with different properties needs to be selected with the help of regular terms. If we expect the solution x to be sparse, we'd better consider adding the l_1 norm to the regular term, i.e.

$$\min_{x \in \mathbb{R}^n} \mu||x||_1 + \frac{1}{2}||Ax - b||_2^2. \tag{11}$$

Due to the existence of regular terms, the objective function of this problem is strongly convex, and the properties of the solution are improved. In the ridge regression problem, the function of the regular term $||x||_2^2$ is actually to punish x with a larger norm. When $\mu = 0$, LASSO regression degenerates into ordinary least squares regression. As the value of μ increases, more and more coefficients are compressed to zero, which helps in feature selection and reduces model complexity [10].

3.2. Alternating Direction Method of Multipliers

Because of the wide application of the LASSO problem, there have been many methods to solve the LASSO problem, such as the gradient descent method, sub-gradient algorithm, approximate point algorithm, etc. We choose the ADMM algorithm in this paper and consider the dual problem at the same time.

There are many complex optimization problems in statistics, machine learning, and scientific computing that may be complex, possibly non-convex and non-smooth. The ADMM naturally provides a wide range of applicable, easy-to-understand, and implement, reliable solutions.

Consider a convex problem

$$\begin{aligned} &\min_{x_1, x_2} f_1(x_1) + f_2(x_2) \\ &s.t.\ A_1x_1 + A_2x_2 = b, \end{aligned} \tag{12}$$

Where f_1, f_2 Are Properly Closed Convex, But Not Necessarily Smooth, $\mathbf{x_1} \in \mathbb{R}^N, \mathbf{x_2} \in \mathbb{R}^N$, $A_1 \in \mathbb{R}^{p\times n}$, $A_2 \in \mathbb{R}^{p\times m}$, $\mathbf{b} \in \mathbb{R}^p$. The characteristic of this problem is that the objective function can be divided into two pieces that are separate from each other, but the variables are held together by linear constraints. Actually, some common unconstrained and constrained optimization problems can be expressed in this form. The LASSO problem is solved by the ADMM.

The unconstrained composite optimization LASSO problem is transformed into a standard form that can be solved by the ADMM

$$\min_{x,z} f(x) + h(z) \\ s.t. x = z, \tag{13}$$

Where $f(\mathbf{x}) = \frac{1}{2}||A\mathbf{x} - \mathbf{b}||_2^2, h(\mathbf{z}) = \mu||\mathbf{z}||_1$. The iterative form of ADMM is

$$\begin{aligned} x^{k+1} &= \underset{x}{argmin}\{\tfrac{1}{2}||Ax-b||_2^2 + \tfrac{\rho}{2}||x - z^k + \tfrac{1}{\rho}y^k||_2^2\} \\ &= (A^TA+\rho I)^{-1}(A^Tb + \rho z^k - y^k) \\ z^{k+1} &= \underset{z}{argmin}\{\mu||z||_1 + \tfrac{\rho}{2}||x^{k+1} - z + \tfrac{1}{\rho}y^k||_2^2\} \\ &= prox_{(\frac{\mu}{\rho})||\cdot||_1}(x^{k+1} + \tfrac{1}{\rho}y^k) \\ y^{k+1} &= y^k + \tau\rho\,(x^{k+1} - z^{k+1}), \end{aligned} \tag{14}$$

Where τ is the step size and usually take the value in $(0, \frac{1+\sqrt{5}}{2})$ [11].

3.3. The dual LASSO problem

Now consider the duality of the LASSO problem

$$\min\ b^Ty + \frac{1}{2}||y||^2 \\ s.t.\ ||A^Ty||_\infty \le \mu. \tag{15}$$

By turning $||A^T\mathbf{y}||_\infty \le \mu$ into a characteristic function and putting it in the objective function, and introducing constraints, we get the dual problem

$$\min\ b^Ty + \frac{1}{2}||y||^2 + I_{||\boldsymbol{z}||_\infty \le \boldsymbol{\mu}(\boldsymbol{z})} \\ s.t. A^Ty + z = 0. \tag{16}$$

If we denote $f(\mathbf{y}) = b^T\mathbf{y} + \frac{1}{2}||\mathbf{y}||^2$and $h(\mathbf{z}) = I_{||\mathbf{z}||_\infty \le \mu(\mathbf{z})}$, by introducing multipliers λ to the constraints, the augmented Lagrange function of the dual problem is that

$$L_\rho(y,z,x) = b^Ty + \frac{1}{2}||y||^2 + I_{||z||_\infty \le \mu(z)} - \lambda^T(A^Ty + z) + \frac{\rho}{2}||A^Ty + z||^2. \tag{17}$$

By deduction, it can be shown that λ is exactly the independent variable of the original problem, and we will rewrite it as $\mathbf{x}$. Fixed $\mathbf{y}$ and $\mathbf{x}$, do a Euclidean projection to the infinite norm sphere to achieve an update of $\mathbf{z}$. Fixed $\mathbf{z}$ and $\mathbf{x}$, solving linear equations

$$(I + \rho AA^T)y = A(x^k - \rho z^{k+1}) - b \tag{18}$$

To achieve an update of $\mathbf{y}$. Therefore, the iterative form of ADMM is

$$\begin{gathered} x^{k+1} = \mathcal{P}_{||z||_\infty \le \mu}(\frac{x^k}{\rho} - A^Ty^k) \\ y^{k+1} = (I + \rho AA^T)^{-1}(A(x^k - \rho z^{k+1}) - b) \\ x^{k+1} = x^k - \tau\rho(A^Ty^{k+1} + z^{k+1}). \end{gathered} \tag{19}$$

3.4. Advantages

On the basis of the original least squares method, we add the regular term, the solution is sparser, and the feature selection of the original problem is realized. In order to solve this LASSO problem, we adopted the ADMM, which is suitable for solving distributed convex optimization problems, especially statistical learning problems. ADMM decomposes the large-scale global problem into several local small-scale problems that are easy to solve through the decomposition and coordination process and obtains the solution of the large global problem by coordinating the solution of the sub-problems.

Both the original and dual ADMM problems need to solve the linear equations, but the particularity of the LASSO problem is noted, that is, the computational amount required to solve the linear equations with $\mathbf{y}$ update is $\mathcal{O}(\mathrm{m}^3)$, which can be further reduced by combining the cache decomposition technique, and the computational amount is smaller than that of the ADMM solution for the original problem [12]. The ADMM here does not use a continuous strategy to adjust μ, which highlights its superiority over other algorithms. The number of iteration steps needed to solve the original problem is less, but the time required for each iteration to solve the dual problem is shorter. In general, ADMM is more efficient in solving the dual problem.

4. The Hiking Optimization Algorithm

4.1. Algorithm background

The Hiking Optimization Algorithm (HOA) is an improvement on Tobler's hiking function (THF) [13]. Tobler's Hiking Function is determined only by the speed and the slope of the terrain, i.e.

$$W_{i,t} = 6e^{-3,5|S_{i,t}+0.05|}, \tag{20}$$

Where $W_{i,t}$ is the speed at t, and $S_{i,t}$ is the slope of terrain.

$$S_{i,t} = \frac{dh}{dx} = tan\Theta_{i,t}, \tag{21}$$

Where dh is the altitude difference, dx is the distance, $\Theta_{i,t}$ is the inclination angle, which fluctuates between $[0, 50^{\circ}]$.

4.2. Algorithm construction

In the HOA algorithm, travel speed is determined by THF speed, leader position, member position, and scanning factor.

$$W_{i,t} = W_{i,t-1} + \gamma_{i,t}(\beta_{best} - \alpha_{i,t}\beta_{i,t}), \tag{22}$$

Where $\gamma_{i,t}$ is a uniformly distributed number within the range [0,1], β_{best} is the position of the lead hiker, $\alpha_{i,t}$ is the sweep factor (SF) of the hiker i and lies in [1, 3], $\beta_{i,t}$ is the position of the hiker i.

The updated position $\beta_{i,t+1}$ is

$$\beta_{i,t+1} = \beta_{i,t} + W_{i,t}. \tag{23}$$

The initialization of $\beta_{i,t}$ is determined by the upper bound ϕ_j^2 and lower bound ϕ_j^1 of the solutions

$$\beta_{i,t} = \phi_j^1 + \delta_j(\phi_j^2 - \phi_j^1), \tag{24}$$

Where δ_j is a uniform distribution number within the range [0, 1], ϕ_j^1 and ϕ_j^2 are the lower and upper bounds of the decision variables about the optimization problem.

Algorithm1: HOA

Input: $\mathbf{UB}, \mathbf{LB}, \mathbf{T}, \mathbf{I}, \mathbf{d}$
Γ ← Preallocate fitness vector $R^{I\times 1}$
Γ_{best} ← Preallocate best fitness vector $R^{T+1\times 1}$
β ← Initialize the hikers' position randomly $R^{I\times d}$
for i ← 1 to I **do**
  Γ_i ← Evaluate a hiker's fitness
end for
$\Gamma_{best,1}$ ← initial position of the leader
for t ← 1 to T **do**
  $\Gamma_{best,t}$ ← Determine the leader
  $\beta_{best,t}$ ← Extract position of $\Gamma_{best,t}$
  for i ← 1 to I **do**
    $\beta_{i,t}$ ← Extract initial position of hiker i
    $\theta_{i,t}$ ← Determine terrain angle of elevation
    $S_{i,t}$ ← Compute the slope
    $W_{i,t}$ ← Determine the actual speed of hiker i
    $W_{i,t-1}$ ← Compute the initial hiking speed
    $\beta_{i,t+1}$ ← Update the hiker's position
    $\beta_{i,t}$ ← Bound $\beta_{i,t+1}$ within LB and UB
     if $\Gamma_n \leq \Gamma(n)$ **then**
     $\gamma(n,:) \leftarrow \Phi$
     $\Gamma(n) \leftarrow \Gamma_n$
    end if
  end for
  Γ_{sol}(i+1) ← min (Γ)
end for
Γ_{sol}^{*} ← argmin (Γ_{sol})
return
Output: Bestpos ←The best solution
    Bestscore ←The best optimal value

5. Numerical example

In this part, we cite three examples of specific data to show that the CFGM-L1 is better than CFGM at some times, and the obtained data prediction results are closer to the raw data. The data that we use are introduced from EI Statistical Review of World Energy. We have the first set of raw data at the case of the Peru's different renewable consumption from 1965 to 1986, $X_1^{(0)}$ = (0.0300, 0.0400, 0.0500, 0.0600, 0.0700, 0.0800, 0.0900, 0.1000, 0.1100,0.1300). And we use the first six data points to calculate the corresponding parameters, while the last four data points are used for prediction purposes. The first six data points are also included during the prediction. By running GM (1,1), CFGM and CFGM-L1 codes separately, we can get the corresponding data values in the Table 1. We also get a trend comparison map of the predicted data and the raw data of the three methods as Figure 1. Then there are the trend maps of MAPE of CFGM and CFGM-L1 as Figure 2 and Figure 3, where MAPE is the evaluation function to get the optimal prediction value. Finally, to show the trend change of the prediction data between CFGM and CFGM-L1more clearly, we draw the trend comparison of the predicted results and the original data in Figure 4.

Table 1. The corresponding data values of $X_1^{(0)}$.

Raw Data	0.0300	0.0400	0.0500	0.0600	0.0700	0.0800	0.0900	0.1000	0.1100	0.1300
GM (1,1)	0.0300	0.0399	0.0748	0.1055	0.1323	0.1559	0.1766	0.1947	0.2106	0.2245
CFGM	0.0300	0.0390	0.0505	0.0606	0.0699	0.0785	0.0867	0.0945	0.1020	0..1093
CFGM-L1	0.0300	0.0399	0.0501	0.0600	0.0699	0.0800	0.0905	0.1016	0.1132	0.1255

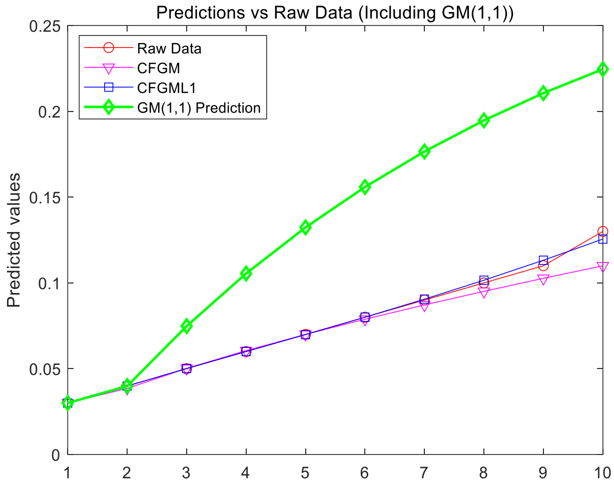

Figure 1. Trend comparison map of the predicted data by three methods of $X_1^{(0)}$.

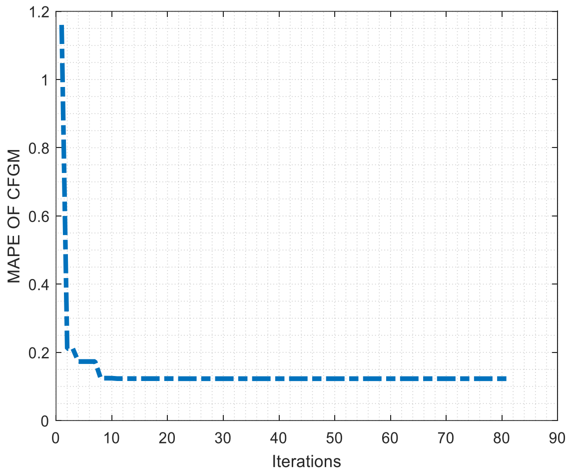

Figure 2. The MAPE trend of CFGM by $X_1^{(0)}$.

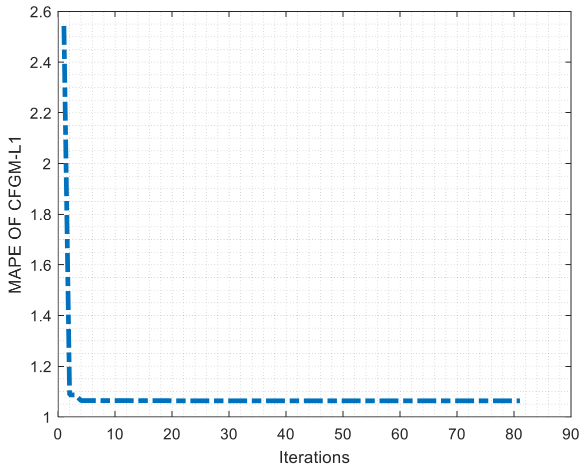

Figure 3. The MAPE trend of CFGM-L1 by $X_1^{(0)}$

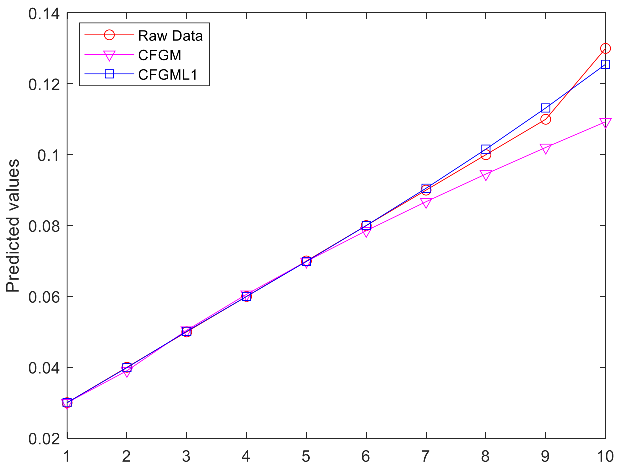

Figure 4. Predicted data trend comparison of $X_1^{(0)}$.

In Figure 1, we found that the CFGM and CFGM-L1 predicted values are similar to the raw data, however, the results of the traditional GM (1,1) are particularly different from the raw data. As we all know, the applicability of traditional GM models is limited when the data fluctuations are relatively small or particularly large, probably because the raw data are very small, so GM (1,1) is not accurate. In the first two MAPE trend maps Figure 2 and 3, the final MAPE value of CFGM and CFGM-L1 models reached the minimum. At this time, we think the data reached the best prediction result. From the changing trend of MAPE CFGM-L1, the corresponding MAPE value shows a ladder decrease, but the CFGM decreases vertically by the first raw data $X_1{}^{(0)}$. As a result, the predicted results of CFGM-L1 are significantly closer to the raw data than CFGM.

From Figure 4, for the prediction of the first six data, the results of CFGM-L1 almost completely fit the raw data. After the sixth data point, there is a slight difference between CFGM-L1 and CFGM predictions and the raw data. However, we can see that the trend of CFGM-L1 prediction data and the raw data is approximately straight upward, whereas the trend in CFGM prediction results shows a slow upward trend.

Then we have the second set of raw data at the case of the Peru's different renewable consumption from 1998 to 2013. So, we can give $X_2{}^{(0)}$= (0.1500, 0.1600, 0.1700, 0.1900, 0.2000, 0.2100, 0.2200, 0.2300, 0.2400, 0.2500). To clarify the advantages of CFGM-L1 in more details, we give the third set of raw data $X_3{}^{(\mathbf{0})}$= (7.246, 7.251, 7.258, 7.262, 7.266, 7.269, 7.272, 7.272, 7.273, 7.277). Run CFGM and the CFGM-L1 code separately, we get the following predicted data trend comparison map Figure 5, Figure 6, Figure 7, Figure 8 and the corresponding data values Table 2, Table 3.

Table 2. The corresponding data values of $X_2{}^{(0)}$.

Raw Data	0.1500	0.1600	0.1700	0.1900	0.2000	0.2100	0.2200	0.2300	0.2400	0.2500
GM (1, 1)	0.1500	0.1579	0.3076	0.4496	0.5843	0.7120	0.8331	0.9480	1.0569	1.1602
CFGM	0.1500	0.1594	0.1740	0.1867	0.1986	0.2100	0.2212	0.2323	0.2435	0.2547
CFGM-L1	0.1500	0.1578	0.1746	0.1881	0.1997	0.2100	0.2195	0.2283	0.2366	0.2445

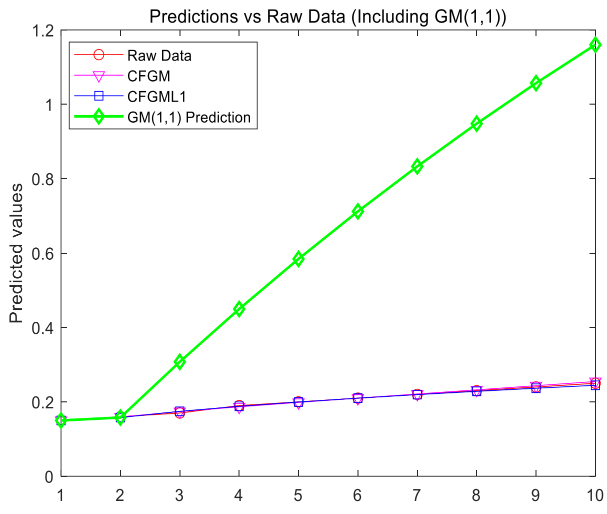

Figure 5. Trend comparison map of the predicted data by three methods of $X_2{}^{(0)}$.

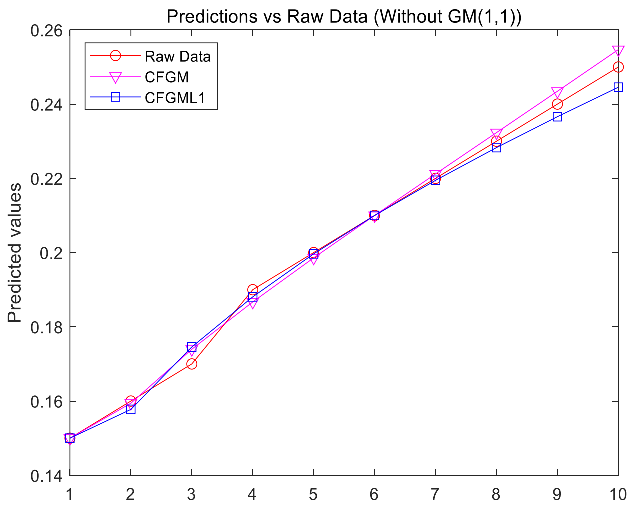

Figure 6. Predicted data trend comparison of $X_2{}^{(0)}$.

Table 3. The corresponding data values of $X_3{}^{(0)}$.

Raw Data	7.246	7.251	7.258	7.262	7.266	7.269	7.272	7.272	7.273	7.277
GM (1, 1)	7.246	7.2521	14.501	21.748	28.991	36.231	43.469	50.704	57.935	65.164
CFGM	7.246	7.2502	7.2573	7.2623	7.2661	7.2693	7.2720	7.2742	7.2763	7.2779
CFGM-L1	7.246	7.251	7.25772	7.26242	7.26599	7.26887	7.27125	7.27328	7.27503	7.27658

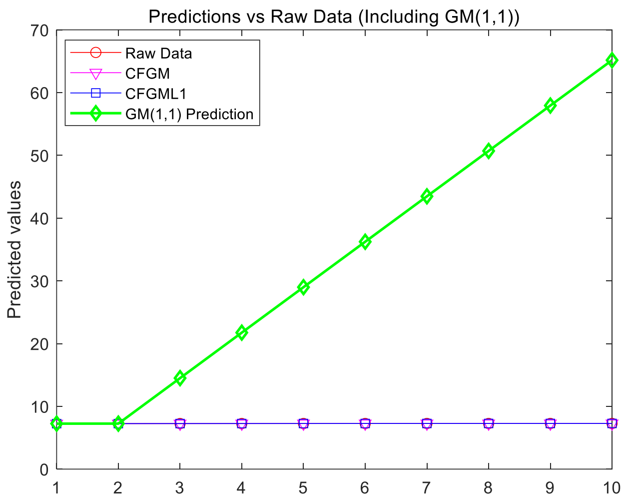

Figure 7. Trend comparison map of the predicted data by three methods of $X_3{}^{(0)}$.

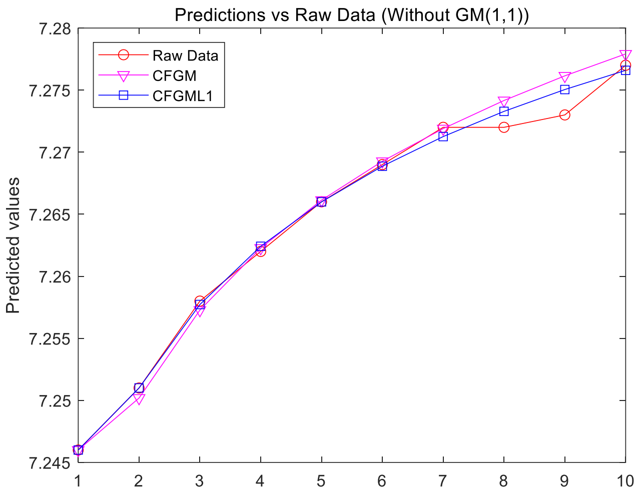

Figure 8. Predicted data trend comparison of $X_3{}^{(0)}$.

From Figure 5 and Figure 7, we can see that the predicted values of GM (1, 1) have no common place with the raw data beside the first and second point at all, but the predicted values of CFGM and CFGM-L1 are fit with the raw data. From Figure 6, we can know that the prediction trend of CFGM-L1 is similar with the raw data, and the difference between CFGM-L1 and the raw data is smaller than CFGM. In Figure 8, the predictions of CFGM-L1 are almost common with the $X_3{}^{(0)}$ at the first six points, and the CFGM-L1 is much closer to the raw data $X_3{}^{(0)}$ at the following four points.

6. Conclusion

In this paper, we introduced the definitions of Conformable Fractional Difference (CFD) and Conformable Fractional Accumulation (CFA) and utilized them to construct the Conformable Fractional Grey Model (CFGM). By transforming the traditional least squares solution into a Least Absolute Shrinkage and Selection Operator (LASSO) problem with an added l_1 regular term, we achieved more stable and accurate model parameters. The dual form of the ADMM (Alternating Direction Method of Multipliers) algorithm was employed to efficiently solve the LASSO problem. Additionally, for the fractional order parameter α, we proposed the HOA algorithm for optimization, significantly improving computational efficiency. Finally, through simulation prediction using three sets of power production data, our improved algorithm demonstrated higher accuracy compared to the traditional CFGM on small sample data.

References

[1] Liu S, Tao L, Tang W. Memorabilia of the Establishment and Development of Grey System Theory [J]. Journal of Grey System, 2024, 36 (1).

[2] Li Q X, Liu S F. The foundation of the grey matrix and the grey input–output analysis [J]. Applied Mathematical Modelling, 2008, 32 (3): 267 - 291.

[3] Xie N. A summary of grey forecasting models [J]. Grey Systems: Theory and Application, 2022, 12 (4): 703 - 722.

[4] Xiong P, Xiao L, Liu Y, et al. Forecasting carbon emissions using a multi-variable GM (1, N) model based on linear time-varying parameters [J]. Journal of Intelligent & Fuzzy Systems, 2021, 41 (6): 6137 - 6148.

[5] Liu C, Lao T, Wu W Z, et al. An optimized nonlinear grey Bernoulli prediction model and its application in natural gas production [J]. Expert Systems with Applications, 2022, 194: 116448.

[6] Ma X, Wu W, Zeng B, et al. The conformable fractional grey system model [J]. ISA transactions, 2020, 96: 255 - 271.

[7] Lijin ZHANG, Xiayan WEI, Jiaqi LU, Junhao PAN, LASSO Regression: From interpretation to prediction, Advances in Psychological Science, Volume 28, Issue 10, 2020, Pages1777 - 1788, ISSN1671 - 3710, https://doi.org/10.3724/SP.J.1042.2020. 01777.

[8] Elgabli A, Elghariani A, Al-Abbasi A O, et al. Two-Stage LASSO ADMM Signal Detection Algorithm for Large Scale MIMO [J]. 2018.DOI: 10.1109/acssc.2017.83356 41.

[9] Oladejo, Sunday O., Stephen O. Ekwe, and Seyedali Mirjalili. "The Hiking Optimization Algorithm: A novel human-based metaheuristic approach." *Knowledge-Based Systems* 296 (2024): 111880.

[10] Ranstam, J., & Cook, J. A. (2018). LASSO regression. British Journal of Surgery, 105 (10), 1348. https://doi.org/10.1002/bjs.10895.

[11] H. Liu, J. Hu, Y. Li, Z. Wen, M. Optimization: Modeling, Algorithms and Theory (in Chinese), Beijing, Higher Education Press, 2020, 88 - 89.

[12] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein and others, Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, Foundations and Trends® in Machine Learning, 3 (1): 1 - 122, 2011.

[13] Goodchild, Michael F. "Beyond Tobler’s hiking function." *Geographical Analysis* 52.4 (2020): 558 - 569.