A New Hybrid Grey Neural Network Based on Grey Verhulst Model and BP Neural Network for Time Series Forecasting

Deqiang Zhou
School of Information and Mathematics, Yangtze University, Jingzhou, China
E-mail: zdqmfk@yahoo.com.cn

Abstract— The advantages and disadvantages of BP neural network and grey Verhulst model for time series prediction are analyzed respectively, this article proposes a new time series forecasting model for the time series growth in S-type or growth being saturated. From the data fitting’s viewpoint, the new model named grey Verhulst neural network is established based on grey Verhulst model and BP neural network. Firstly, the Verhulst model is mapped to a BP neural network, the corresponding relationships between grey Verhulst model parameters and BP network weights is established. Then, the BP neural network is trained by means of BP algorithm, when the BP network convergences, the optimized weights can be extracted, and the optimized grey Verhulst neural network model can be obtained. The experiment results show that the new model is effective with the advantages of high precision, less samples required and simple calculation, which makes full use of the similarities and complementarities between grey system model and BP neural network to settle the disadvantage of applying grey model and neural network separately. It is concluded that grey Verhulst neural network is a feasible and effective modeling method for the time series increasing in the curve with S-type.

Index Terms— Time Series Prediction, BP Neural Network, Grey Verhulst Model, Grey Neural Network, Grey Verhulst Neural Network

1. Introduction

Time series prediction refers to the process by which the future values of a system is forecasted based on the information obtained from the past and current data points\(^1\). At present, there are a lot of methods for time series prediction, from traditional statistical method such as ARMA (Auto Regressive Moving Average) model to artificial intelligence based approaches, the core of these models is to establish a prediction model\(^2\). Neural Network (NN) based models are widely used as an artificial intelligence-based approach, back propagation (BP) being the most widely used technique for updating the parameters of the model. BP neural network is the most used neural network at present. It has unique approximation ability and simple structure, and it is a neural network with good performance. The BP learning process works in small iterative steps, and the network produces some output based on the current state of it's synaptic weights (initially, the output will be random). This output is compared to the known-good output, and a mean-squared error signal is calculated. The error value is then propagated backwards through the network, and small changes are made to the weights in each layer. The weight changes are calculated to reduce the error signal for the case in question. The whole process is repeated for each of the example cases, then back to the first case again, and so on. However, not only are the statistical models not as accurate as the neural network-based approaches for nonlinear problems, they may be too complex to be used in predicting future values of a time series. One major criticism about the BP model is that it demands a great deal of training data\(^3\) and its application was inhibited largely by the slow convergence rate and over-prolonged training time, primarily the results of inappropriate sample preprocessing for a large initial sample domain. On the other hand, it is well known that selecting the number of neurons in hidden layer is also an important and tough problem because it affects the overall performance of neural networks. However, there is still no definite theory to settle it out.

As the neural network, the large amount of data that can be used to provide information, but also increase the difficulties of dealing with these data\(^3\). Grey system theory is an interdisciplinary scientific area that was first introduced in early 1980s by Deng\(^4\). In the field of information research, deep or light colors represent information that is clear or ambiguous, respectively. Meanwhile, black indicates that the researchers have absolutely no knowledge of system structure, parameters and characteristics, while white represents that the information is completely clear. Colors between black and white indicate systems that are not clear, such as social, economic or weather systems. The fields covered by grey theory include systems analysis, data processing, modeling, prediction, decision making and control. The grey theory mainly works on systems analysis with poor, incomplete or
uncertain messages. Because the grey system model needs little origin data, has simple calculate process and higher forecasting accuracy, it has been widely used in the time series prediction of a lot of research fields. In these studies and the others, grey system theory-based approaches can achieve good performance characteristics when applied to real-time systems, since grey predictors adapt their parameters to new conditions as new outputs become available. Because of this reason, grey predictors are more robust with respect to noise and lack of modeling information when compared to conventional methods[12]. A grey prediction model is one of the most important parts in grey system theory[13], and that, the grey Verhulst model is a special kind of model within the grey system. The grey Verhulst model[6-10] is used to forecast the raw sequence growth in S-type or growth being saturated. However, there are many defects in traditional grey model[11-12]. First, the traditional model need to face the reasonable selection of background value[13]. On the other hand, in order to obtain forecasting model, in this model, parameters which obtained from the grey difference equation is substituted into the grey differential equation[14], thus we also can not avoid the jumping errors from the difference equation to differential equation in traditional grey modeling.

In the actual forecast, the same problem can have many different forecasting methods. Given many approximation methods, it has become difficult for an engineer to select the most appropriate method for the problems under study[15]. On the other hand, in the actual application, because of time series has irregular, chaos and nonlinear characteristics, it is difficult to establish the ideal prediction model system[2]. So it would be very meaningful if different methods can be combined, which consolidates the advantage of the different models and obtains better forecasting accuracy. With the application of the grey system theory and the BP neural network in the field of forecasting, the time series forecasting technology has obtained the great development. The grey forecast model has the very strong fusion strength and penetrability to the general model, so it can be combined with other models to improve the forecasting precision. Therefore, analysis and research of the grey model with other model's combination are carried, the combination forecasting models are established[16-20].

Based on the approach of “mixture of experts”, some researchers have integrated the GM(1,1) and the NN to enhance the prediction according to their complementary merits. Generally, the ways to combine the GM(1,1) and the NN can be categorized into four classes[16,21]: simple combination, serial combination, strengthening grey system with neural network, building neural network with the aid of grey system. The work in [12]and [22] established grey neural network combing GM(1,1) model and neural network. Grey neural network is an innovative intelligent computing approach, which makes full use of the similarities and complementarities between grey system model and BP neural network to settle the disadvantage of applying grey model and neural network separately[23]. However, it is imperfect when the time series increase in the curve with S-type, or the increment of time series is in the saturation stage[23]. One major reason is that the structure of this kind of grey neural network model is established based on GM(1,1) model, therefore, it has the defect of GM(1,1) model. The essence of the GM (1,1) is a model fitting sequence such as $Ae^\alpha$ exponential function, which request primitive data has the approximate exponential incremental characteristic. In this case, the time series forecasting error of Grey neural network model will become larger and the result is unaccepted in the real world. To solve this problem and expand the application range of the grey neural network model, this paper proposes a new hybrid grey neural network combing the grey Verhulst model and BP neural network. The new grey neural network based on grey Verhulst model and BP neural network consolidates the advantage of the different models, and is perfect for the time series increasing in the curve with S-type.

The remainder of this paper is organized as follows. In Section 2, the grey Verhulst model is introduced. In Section 3, the disadvantages of grey Verhulst model for time series prediction is analyzed, this article establishes a new hybrid grey neural network based on grey Verhulst model and BP neural network for time series Forecasting. The experiment results and discussions are presented in Section 4. Finally, conclusion is drawn in Section 5.

II. Grey Verhulst Model

The grey Verhulst model is a special kind of model within the grey system. It has been extensively used in numerous applications to explain the phenomenon of population increasing, living creature breeding and its individual growth. Grey Verhulst model is also a time series forecasting model, the constructing process is described below[23]:

1) Denote the nonnegative original sequence by

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$$  \hspace{1cm} (1)

2) The accumulated generating operation (AGO)\footnote{AGO stands for Accumulated Generating Operation.} formation of $X^{(0)}$ is defined as:

$$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n))$$ \hspace{1cm} (2)

$$x^{(1)}(k) = \sum_{j=1}^{k} x^{(0)}(j)$$

3) The grey Verhulst model can be established by constructing a first order differential equation for $X^{(1)}$
\[
\frac{dx^{(1)}}{dt} + ax^{(1)} = b(x^{(1)})^2
\]  
(3)

The solution of Eq.(3) can be obtained by using the least square method, that is,

\[
\hat{x}^{(1)}(k + 1) = \frac{\hat{a}x^{(1)}(1)}{bx^{(0)}(1) + (\hat{a} - \hat{b}x^{(0)}(1))e^{\hat{a}t}}
\]  
(4)

where

\[
\hat{a}, \hat{b} = (B^T B)^{-1}B^TY.
\]

\[
B = \begin{bmatrix}
-z_2^{(1)}(z_2^{(1)})^2 \\
-z_3^{(1)}(z_3^{(1)})^2 \\
\vdots \\
-z_n^{(1)}(z_n^{(1)})^2
\end{bmatrix},
\]

\[
z_k^{(1)} = 0.5(z^{(1)}(k) + x^{(1)}(k - 1)).
\]

Applying the inverse accumulated generating operation (IAGO)\(^{(6)}\), the grey Verhulst model fitted and predicted values are obtained by

\[
\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k) \quad (k = 2, 3, \ldots)
\]

\[
\hat{x}^{(0)}(1) = \hat{x}^{(0)}(1).
\]  
(5)

Eq.(4) and Eq.(5) are constructed by the original data sequence, we call this model as \(X^{(0)}\)-Verhulst model. In actual application, if the original data sequence increases in the curve with S-type, we can also use it as the “\(X^{(1)}\)" to construct Verhulst equation directly. The Verhulst model is relatively applicable to describe the monotonous variety process increases in the curve with S-type or the increment of nonnegative original sequence is in the saturation stage.

III. Grey Neural Network based on Grey Verhulst Model and BP Neural Network

3.1 Grey Verhulst Model and BP Neural Network Mapping Relationships

We will build grey Verhulst model directly. Let’s consider the following non-linear function,

\[
y(t) = \frac{a\hat{x}^{(0)}(1)}{bx^{(0)}(1) + (a - bx^{(0)}(1))e^{-a(t-1)}}.
\]  
(6)

where \(a, b, \hat{x}^{(0)}(1)\) are undetermined parameters.

Assume that the AGO data sequence \(X^{(1)}\) of the original sequence \(X^{(0)}\) increases in the curve with S-type or the increment of the original sequence is in the saturation stage, compared with grey Verhulst model, the function value of (6) at \(t = 2, 3, \ldots, n\) can be taken as the predicted value in the corresponding time \(t\) of the grey Verhulst model, then from the data fitting’s viewpoint, the sequence \(X^{(1)}\) can be fitted by (6). Therefore, we can fit the AGO data sequence \(X^{(1)}\) by (6), the key question is how to estimate the parameters \(a, b, \hat{x}^{(0)}(1)\). In order to obtain a better solution, this initial value does not take the first data of the original sequence data, but as a parameter to be determined by data fitting.

Because (6) is nonlinear, we now resolve the above problem by means of BP network which can approximate nonlinear function with any precision.

First, (6) is transformed as follows:

\[
y(t) = \frac{b}{a + (\hat{x}^{(0)}(1) - b)e^{-a(t-1)}}
\]  
(7)

\[
\frac{1}{y(t)} = \frac{1}{a} + \frac{1}{\hat{x}^{(0)}(1) - b}e^{-a(t-1)}
\]  
(8)

Since \(e^{-a(t-1)} > 0\), therefore,

\[
\frac{1}{y(t)} = \left(\frac{1}{\hat{x}^{(0)}(1)} - \frac{b}{a} \frac{e^{-a(t-1)}}{1 + e^{-a(t-1)}} + \frac{b}{a} \frac{1}{1 + e^{-a(t-1)}}\right)(1 + e^{-a(t-1)})
\]

\[
= \left(\frac{1}{\hat{x}^{(0)}(1)} - \frac{b}{a} \left(1 - \frac{1}{1 + e^{-a(t-1)}}\right) + \frac{b}{a} \frac{1}{1 + e^{-a(t-1)}}\right)(1 + e^{-a(t-1)})
\]

\[
= \left(\frac{1}{\hat{x}^{(0)}(1)} - \frac{b}{a} - \frac{\hat{x}^{(0)}(1)}{\hat{x}^{(0)}(1) - \frac{1}{1 + e^{-a(t-1)}}} + \frac{2b}{a} \frac{1}{1 + e^{-a(t-1)}}\right)(1 + e^{-a(t-1)})
\]

The above equation is mapped to the BP neural network, the network structure is described as Fig.1:
The corresponding relationships between the network weight and the grey Verhulst model parameters are established as follows:

\[
\begin{align*}
\omega_1 &= a, \\
\omega_{21} &= -\frac{1}{\hat{x}^{(0)}(1)}, \\
\omega_{22} &= \frac{2b}{a}, \\
\omega_{31} &= \omega_{32} = 1 + e^{-a(1-t)}
\end{align*}
\]  

(9)

The threshold value of the LB level is

\[
\theta = (1 + e^{-a(1-t)}) \left( \frac{b}{a} - \frac{1}{\hat{x}^{(0)}(1)} \right).
\]

And the LB level neuron's transfer function is taken as Sigmoid type function: \(f(x) = \frac{1}{1 + e^{-x}}\), there is a high-gain area which ensures that the network eventually reach a steady\(^{[22]}\). Other neuron's transfer function is taken as a linear function: \(f(x) = x\).

Through the above method, the grey Verhulst model is mapped to a BP neural network. Figure 1 shows this neural network structure. The corresponding relation between the grey Verhulst model parameters and BP network weights is established. At the same time, a new grey neural network based on grey Verhulst model and BP Neural network is built, for simple, we call it grey Verhulst neural network.

**3.2 Character Analysis for Grey Verhulst Neural Network**

The above grey Verhulst neural network is a hybrid model which consolidates the character both of the grey Verhulst model and neural network. However, the model is different from the two single models. In this model, the traditional grey Verhulst model parameters estimation problem is transformed into an optimization problem for the weights of neural network. As we all know, for the neural network design, the key is focused on the number of hidden layers and the number of nodes in each layer of the network. An appropriate number of hidden layers can minimize the system error of the network. The number of hidden layer units is directly related to the requirements of the problem and the number of input/output units. Therefore, the selection of an optimal number of hidden layer units is crucial. However, it is difficult to design a neural network when the problem is particularly complex because there is few rigorous design criteria\(^{[24]}\). Grey Verhulst neural network in this paper is also a neural network, this new neural network is built based on grey Verhulst model, and some useful advice is put forward for the choice of some key parameters in the BP model. Therefore, figure 1 shows this neural network structure is simple and clear. Compared to the traditional BP neural network the model is much more convenient and direct. The neural network is trained by means of data sets \((1-t, \frac{1}{\hat{x}^{(0)}(t)}), t = 1,2,\cdots, n\), when the BP network convergence, the optimized network model weights can be obtained\(^{[13]}\), and the optimization modeling for the grey Verhulst model based on BP algorithm can be realized. Then, applying the inverse accumulated generating operation, the grey Verhulst model fitted and predicted values are obtained by

\[
\begin{align*}
\hat{x}^{(0)}(k) &= \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k), (k = 2, 3, \cdots n) \\
\hat{x}^{(0)}(1) &= -\frac{1}{\omega_{21}},
\end{align*}
\]

where \(\hat{x}^{(1)}(k) = \frac{1}{y(k)}\).

As we can see, compared with the grey Verhulst model, the computation is simpler and the modeling process is more intuitive. On the other hand, it does not require the transform from discrete form to continuous form, therefore it can avoid the jumping errors from the difference equation to differential equation, and overcome the defects of the grey Verhulst model.

**IV. Little Sample Experiment Validation and Discussions**

Grey model due to whose distinguishing features: modeling by less data(suiting the data as few as 4)\(^{[23]}\), thus underlay grey modeling and grey forecasting. The attraction of neural networks is that they are best suited to solving the problems that are the most difficult to solve by traditional computational methods. As mentioned above, the grey Verhulst neural network is an hybrid model which consolidates the advantages both of the grey Verhulst model and neural network. Now, in this section, we validate the performance of the grey Verhulst neural network by a little sample example.

**4.1 Experiment Data and Design**

We use the following example with little data and
little sample to carry on the experiment. The raw series demonstrate as follows\(^{(25)}\):

\[
x^{(0)} = (1, 2.0890, 3.6663, 3.0117)
\]

The original sequence \(x^{(0)}\) approximate the single-peak type, and the AGO series of \(x^{(0)}\) presents the saturation characteristic, therefore, it is suitable for grey Verhulst modeling\(^{(26)}\). According to the proposed method in this paper, the BP neural network with three hidden layer is established. The network structure is described as Fig.1. The first level neuron's transfer function is taken Log-Sigmoid type function, other neuron's transfer function is taken as a linear function.

The maximum training epoch is \(M = 300\), the allowance permissible error is \(E = 0.0000005\). The LM (Levenberg-Marquardt) algorithm in this paper is used as the training algorithm, the learning rate is dynamically determined by LM algorithm. To use the Matlab6.5 programming, when the network achieves the accuracy requirement, we can obtain the optimized grey Verhulst BF neural network model.

The results obtained by the model in this paper is compared with traditional GM(1,1) model and grey Verhulst model. To demonstrate the accuracy of above models, the actual value and the predicted value can be compared. Mean absolute percentage error (MAPE)\(^{(26)}\) approach has been recommended to validation. Mean absolute percentage error is defined as

\[
\text{MAPE} = \frac{1}{l} \sum_{i=1}^{l} \frac{|y_i - \hat{y}_i|}{y_i}
\]

where \(y_i\) is the actual value,\( \hat{y}_i \) is the predicted value, \(\frac{|y_i - \hat{y}_i|}{y_i}\) is absolute percentage error of \(y_i\).

### 4.2 Experiment Results and Error Analysis

Compared with the traditional grey Verhulst model and GM(1,1), the Experiment results and MAPE are shown in Tab.1.

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Actual</th>
<th>Grey Verhulst</th>
<th>This paper</th>
<th>GM(L1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.0001</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2.0890</td>
<td>2.4231</td>
<td>2.0858</td>
<td>2.5265</td>
</tr>
<tr>
<td>3</td>
<td>3.6663</td>
<td>3.1203</td>
<td>3.7219</td>
<td>2.9041</td>
</tr>
<tr>
<td>4</td>
<td>3.0117</td>
<td>3.2680</td>
<td>2.8895</td>
<td>3.3381</td>
</tr>
<tr>
<td>MAPE(%)</td>
<td>9.85</td>
<td>1.43</td>
<td>13.14</td>
<td></td>
</tr>
</tbody>
</table>

In Tab.1, the model in this paper presented higher overall fitted accuracy with a MAPE of 1.43%. Compared with the traditional Grey Verhulst model, the precision enhances 85.48%. Compared with GM(1,1) model, the precision enhances 89.12%. As opposed to the traditional two grey model, it can be seen that the improvement of accuracy of the model in this paper is very significant. Analyzing the errors of competitive models by means of comparing the MAPE, we conclude that the method in this paper is a better fitting model. Furthermore, according to the criteria purpose by Lewis\(^{(26)}\), the model is better than the other forecasting models because of lower MAPE, indicating that the performance for fitting little sample of the model in this paper is good. It can also be noticed in Tab.1, the initial value \(x^{(0)}(1)\) of the model in this paper is 1.0001, it does not take the first data of the original sequence data, but as a parameter to be determined by data fitting, therefore, relative to grey Verhulst model and GM(1,1)model simply take the first data as the initial value, it is more reasonable.

Fig.2 shows comparison of the actual and different models fitted values respectively. It can be noticed in Fig.2 that the original sequence \(x^{(0)}\) approximate the single-peak type, as mentioned above, relative to GM(1,1) model, it is more suitable for grey Verhulst modeling, which can be proven by the experiment results listed in Tab.1. It can also be noticed in Fig. 2 that the model in this paper fitting curve is very closer, what’s more the curve also approximate the single-peak type, which supports the view that the new grey Neural network makes full use of the complementarities between grey model and neural network\(^{(12,22)}\).

![Fig. 2: Actual and fitted value of different models](image-url)
The grey theory is suitable to little sample, incomplete information and the BP algorithm is capable of processing non-linear adaptable information, and the grey Verhulst neural network is a combination of those advantages.

As we all know, the traditional BP neural network model demands a great deal of training data and its application was inhibited largely by the slow convergence rate. Fig. 3 shows the training times of our method only use 125 steps to reach the minimum error, furthermore, the traditional BP network not only need a large amount of data to train, but also the theoretical guidance of network settings is lack, moreover the training time is much greater than this article method. Many experiments show that the number of network convergence the method is generally is about 130.

In summary, it can be concluded that grey Verhulst neural network model shows excellent learning ability with fewer training data, and combines the advantages of Grey Verhulst model and BP neural network. Experiment results show the method combines the characteristics of the small sample and poor information of the grey Verhulst model, and displayed the characteristics of strong nonlinear approximation and fault-tolerant capability of the neural network. The new grey neural network based on grey Verhulst model and BP neural network consolidates the advantage of the different models, and is perfect for the time series increasing in the curve with S-type.


does not work.

V. Conclusion

The practice proves that the traditional the improved Verhulst modeling method exist some shortage. In this paper, a new grey Neural network based on grey Verhulst model and BP Neural network is presented for time series prediction. The grey Verhulst neural network can overcome the defects of the grey Verhulst. It is suitable for building by less data and little sample.

This model expand the application range of the grey neural network model. The experiment results show that the grey Verhulst neural network for time series prediction is applicable, especially, when the series increases according to the curve with S-type, not only higher accuracy can be obtained, but also the superiority and the features of grey system model can be reserved.

In summary, a grey Verhulst neural network is a new hybrid Grey Neural Network which consolidates the respective advantages of the neural network and grey Verhulst. It is feasible and effective for the time series increasing in the curve with S-type.

However, as a BP neural network, grey Verhulst Neural network is easy to converge toward some local minimum. From this angle, further research work should be studied.

References


**Author’s Profiles**

**Deqiang Zhou:** Associate Professor of school of information and mathematics in Yangtze University interested in machine learning and wavelet analysis.