Research on Time Series Modeling by Genetic Programming and Model De-noising

YONGQIANG ZHANG, LILI WU
School of Information and Electricity-Engineering
Hebei University of Engineering
Guangming South Street 199#, Handan City, Hebei Province 056038
P.R.China.

Abstract: - In order to cast off the subjective assumptions of traditional methods for modeling, this paper brings forward the Genetic Programming (GP for short) algorithm to establish a reasonable system model dynamically for time series signal. Meanwhile, the approach of wavelet threshold is adopted to de-noising for the GP models. On the basis of these theories, the simulation experimentations about two instances are carried on. The results indicate that the threshold approach of wavelet de-noising for time series signal models take on better impacts, which can improve the GP models to some extent, and enhance the forecast precision of the model.

Key-Words: - Genetic Programming, De-noising, Wavelet Threshold, Time Series, Modeling, GP Model

1 Introduction

In the areas of engineering technology, economic management, natural science, social science and so on, there are many complex systems and nonlinear phenomenon, which change over time. Such as price fluctuation, urban expansion, weather variations, population growth and so on. People often need to establish a reasonable dynamic system model (dynamic model) according to the observation data of dynamical system, which provides a basis for the system analysis, design as well as the forecast of the system future state.

However, traditional methods of modeling presume a group of dynamic equation structures in the view of intuition or experience. And then the numerical method is adopted to determine the parameters of these presumed equations. Its poor rationality, accuracy and complicated calculation limit the application of traditional methods. Genetic Programming (GP for short) algorithm can combine the constructive estimation with parameter estimation and establish dynamic system model using observation data when there is not complete structure information of model, which can also effectively solve difficulties of complex system modeling that is nonlinear, multidimensional and high-order. It provides people with a new way of thinking and a new approach of complex system modeling.

In that the model based on GP algorithm is optimizing model, which can be interfered and contaminated inevitably by various noises and make the prediction accuracy of models reduced. Therefore, the model de-noising is required. While, the filter de-noising is a most extensive method in practical application, which often filter out useful signal when de-noising. The method of filter de-noising analyses which noises among the frequency ranges should be eliminated in the view of pure frequency-domain. However, for the advantages of time-frequency domain localization and the flexible selection for wavelet radix, wavelet transformation provides an effective solution tool for this problem [1].

2 Modeling by GP Algorithm

2.1 Summary of GP Algorithm

Genetic Programming (GP) is a technique based on biological evolution, which is developed from Genetic Algorithm (GA). In Genetic Programming, individuals are computer programs that try to solve a particular problem, which uses trees' delamination structure to denote solution spaces, and each tree structure relates to a computer programming in the solution spaces. The leave nodes are initial variables of the problem, the middle nodes presents the functions which composes to initial variables, and the root nodes are functions of final outputs.

At the beginning of evolution, all of these programs are generated at random. This is the first generation. Each program is run in order to see how well it solves the given problem. Since the programs were completely random, they will be quite different from each other and some will do better on the problem than others. After testing all individuals (programs) on the given problem, evolution goes on to the next generation. Just like in biological evolution, reproduction takes place between two generations. The most suitable individuals are the ones which can solve the problem in the best way, making the fewest errors. These individuals get the best chance of reproduction. So, the more errors an individual makes, the smaller its chance of reproduction is.

When reproduction, crossover or mutation may occur. Crossover occurs between two individuals. In each individual, a crossover point is randomly chosen and the corresponding parts of the two individuals are swapped. Mutation means that part of the program is removed and a new part is randomly generated and inserted at that point.

Here, we define the depth of the tree is less than N , which is a given positive integral; T_i are all the trees, and each tree's root nodes $r_k \in F$; their leave nodes are defined as $l_j \in T$, so $S = \{T_i \mid r_k \in F, l_j \in T\}$, in which F is the function sets and T is the terminal sets. We also have to define the fitness functions $f : S \to R$, which can search for the best individuals T_i^* in the space S in order to make the formula $f(T_i^*) = f^* = \max\{f(T_i): T_r \quad T_r \in S\}$.

The approach of creating the trees randomly by growth method, whose depth can't exceed the maximum depth D , is defined as: for the root nodes, we have to select them from the function sets in order to generate non-ordinary individuals; while for the other nodes, we may select them from $T \cup F$ if the depth of nodes to be selected is less than D ; and if the depth equals D , we can only select them from T [2,4]. We also adopt the "best-reserved strategy" to make the best individuals reserved in the next generation, but these individuals won't attend the genetic operations.

The parameters of modeling are given in Table 1 as follows:

Table 1 Parameters in GP Operation

Parameters	Span solutions
Function set (F)	+、-、/、ln、sin、tan、sqrt
Terminal set (T)	Argument x
Population size	30
The probability of crossover	0.90
The probability of mutation	0.05
Generative method	Growth
Selection method	Tournament, best reserved strategy
Terminal condition	100 generations
Great number of generation	100
Maximum of tree depth	5
Max depth after crossover	7
Max depth after mutation	7

2.2 Examples for modeling by GP

Here, two examples for GP modeling are given, whose parameters are all from Table 1. The first one is based on a group of surveyed data series which is to forecast content of gas in the air [5], shown in Table 2.

Table 2 Surveyed Data Series of Gas Content

x	1	2	3	4	5
y	0.25	0.35	0.361	1.051	1.361
x	6	7	8	9	10
y	1.402	2.305	3.0702	3.5892	3.9059

Where, x represents the time series for forecasting gas content, y is the content of gas in the air. Here, the new GP model is created based on the parameter of x. After 50 generations operation under MATLAB programming environment, we can get a better fitting individual, which is shown in (1) as follows.

$$f(x) = 0.1 \times \left(x \cdot \ln\left(\frac{x^{\ln(x+\sin x)}}{\sqrt{x - \cos(x^{\sin x})}} \right) + \sin\left(\frac{\left(x + \sin x - e^{(x+\ln x)}\right) \cdot \ln x}{\sin\left(\frac{x^{\sqrt{x}}}{\sqrt{x}} \right) + \sin x} \right) + \sin x \right) \quad (1)$$

Another example is based on a group of data, which is for forecasting surveyed data series of service lifetime of electron apparatus [6], shown in table3.

Table 3 Surveyed Data Series of Service Lifetime of Electron Apparatus

x	1	2	3	4	5
y	2.01	4.35	6.92	9.61	12.50
x	6	7	8	9	10
Y	15.49	18.74	22.31	26.06	30.03

Where, x represents the time series for forecasting data, y is the service lifetime of electron apparatus. Here, the new GP model is created based on the parameter of x. After 50 generations operation under MATLAB programming environment, we can get a better fitting individual, which is shown in (2) as follows.

$$f(x) = ((x) + \sqrt{(x)}) \times \ln(x) \quad (2)$$

3 Wavelet Threshold De-noising

The method of wavelet threshold de-noising on theory is mainly based on that: the main energy of signal which is in the domain of wavelet and belonged to the space of Besov, concentrated in a limited number of coefficients. While the noises' energy is distributed throughout the wavelet domain. Therefore, after the wavelet decomposition, the wavelet transform coefficients of signal are larger than the wavelet transform coefficients of noise, and then a suitable number λ as the threshold can be obtained. When $W_{j,k}$ is less than the threshold, $W_{j,k}$ is mainly caused by noise; when $W_{j,k}$ is larger than the threshold, $W_{j,k}$ is mainly caused by signal, thus it can achieve signal-to-noise separation.

The specific steps of wavelet threshold de-noising method are shown as follows [3]:

(1) Discrete wavelet transformation for signal with noise $f(k)$ is carried on and each scale wavelet coefficients $W_{j,k}$ can be gotten;

(2) Each scale wavelet coefficients $W_{j,k}$ should be processed by threshold and then get the estimate wavelet coefficients $\hat{W}_{j,k}$, which satisfies $\left\|\hat{W}_{j,k}-\mu_{j,k}\right\|$ much smaller. This step is the key one for wavelet threshold de-noising method. However, the selection of threshold will affect the de-noising result to a large extent. There are four threshold functions in MATLAB for using:

Sqtwolog: this adopts a fixed threshold style, which will get good visual de-noising result by the soft threshold processing;

②Minimaxi: this method adopts minimum and maximum principle for the selection of threshold, which can create extremum of least mean-square error;

③ Rigrsure: this method adopts unbiased likelihood estimate principle of Stein for the selection of threshold;

④heursure: this method selects the heuristic threshold, which combine the sqtwolog method with the rigrsure one.

(3) The wavelet should be reconstructed by $\hat{W}_{j,k}$ and then we can get estimate signal $\hat{f}(k)$ of signal $f(k)$, which is de-noising signal.

4 Simulation of Model De-noising

Hereon, we adopt two of the wavelet de-noising function "wden" and "wdencmp" that are in wavelet toolbox of MATLAB for model (1) and (2) de-noising, which has been established above. Because of selecting different parameters of wavelet de-noising function can get different de-noising effect, hereon, we mainly adopt wavelet sym (N) for de-noising to the testing example in this paper, the simulation results have been shown as follows:

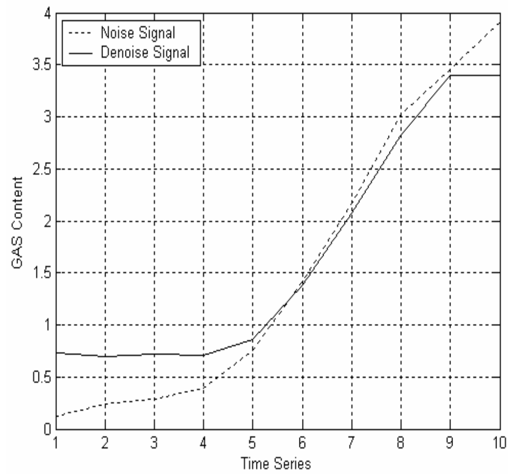

Fig. 1 The de-noising signal of model 1 with "wden"

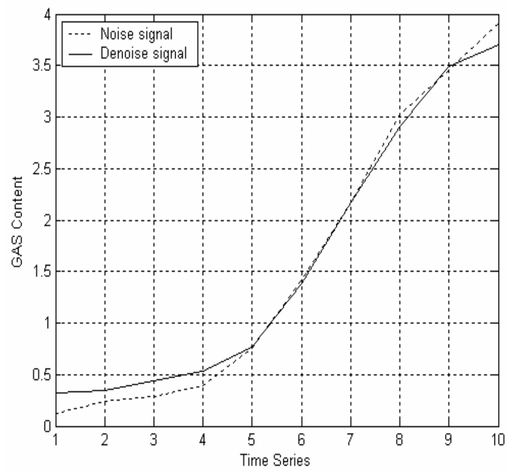

Fig. 2 The de-noising signal of model 1 with "wdencmp"

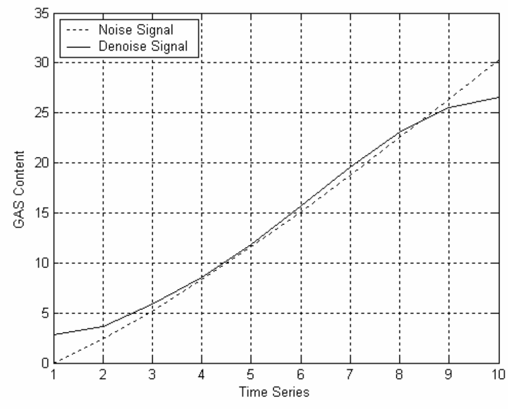

Fig. 3 The de-noising signal of model 2 with "wden"

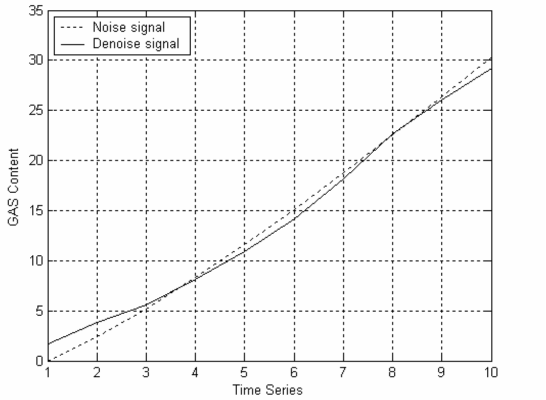

Fig. 4 The de-noising signal of model 2 with "wdencmp"

In order to compare the de-noising results between the initial model signals and the signals which have been de-noised, the concept of model noise [7, 8] should be presented.

If we have observed the continuous testing time series as $t_1, t_2, \cdots, t_{j-1}$ while we want to predict the next time t_j. Supposed that the true data series which can be computed by the new model, we can calculate the appraisal value of next time $\hat{t}_j$. Then operate the programming from beginning until the next forecasting time is obtained. We should repeat this operation of predicting among the range of j. The model noise indicates the degree of imported noise by the GP new model itself, which can be defined by (3) as follows:

$$Model - Noise = \sum_i \left| \frac{\hat{t}_i - \hat{t}_{i-1}}{\hat{t}_i} \right| \qquad (3)$$

Evidently, the smaller the model noise is, the better the new model has been created.

The comparisons of model noises for these two model signals above are given in the following tables:

Table 4 The Comparison Result Between The Initial Signals and The De-noising One of Model 1 with Wden Function

Parameter	minimaxi	rigrsure	heursure	sqtwolog
Model noise of Initial signal	2.7353	2.7353	2.7353	2.7353
Model noise of the De-noising signal	1.0310	1.4079	1.6620	0.3790

Table 5 The Comparison Result Between The Initial Signals and The De-noising One of Model 1 with Wdencmp Function

Parameter	Sym3		Sym6	
	3	5	3	5
Model noise of Initial signal	2.7353	2.7353	2.7353	2.7353
Model noise of the De-noising signal	2.0356	1.9912	2.1794	2.0469

Table 6 The Comparison Result Between The Initial Signals and The De-noising One of Model 2 with Wden Function

Parameter	minimaxi	rigrsure	heursure	sqtwolog
Model noise of Initial signal	3.0721	3.0721	3.0721	3.0721
Model noise of the De-noising signal	2.3022	2.8483	2.8483	1.9301

Table 7 The Comparison Result Between The Initial Signals and The De-noising One of Model 2 with Wdencmp Function

Parameter	Sym3		Sym4	
	3	5	3	5
Model noise of Initial signal	3.0721	3.0721	3.0721	3.0721
Model noise of the De-noising signal	2.4273	2.3251	2.4018	2.2568
Parameter	Sym5		Sym6	
	3	5	3	5
Model noise of Initial signal	3.0721	3.0721	3.0721	3.0721
Model noise of the De-noising signal	2.5048	2.4385	2.7709	2.5819

From the results of all these comparisons above, we can draw the conclusions that it is feasible of model de-noising based on Genetic Programming algorithm by the wavelet approach.

5 Conclusion

This paper makes use of GP algorithm to create the rational evolution models for the testing time series dynamically, which can cast off different kinds of the subjective assumptions to some extent. Meanwhile, the new GP models can be de-noised by the method of wavelet threshold, which also get the better effect and higher applicable values in practice. However, the method of wavelet threshold for de-noising is not suitable for all of the time series models. Which kinds of time series models can fit for de-noising in the way of wavelet threshold, and which ones can not, as well as how to distinguish to deal with, are all the problems to study further. In addition, how to select the wavelet threshold more scientifically and rationally should be researched more deeply and hardly.

Acknowledgement:
Project 60573088 supported by National Natural Science Foundation of China (NSFC); Project 603407 supported by National Natural Science Foundation of Hebei Province.

References:
[1] Haopan Du, Shuang Cong. Research on wavelet de-noising based on MATLAB [J]. *Computer Transaction on Simulation*, 2003, 20(07): 119.
[2] Lin Dan, Li Minqiang, Kou Jisong. A theorem about convergence of GP, *Transaction of Xiamen university (The natural science edition)*, 2000(1), 2-4.
[3] Hua Cui. Wavelet analysis and its applications on the signal processing [D]. *Xi'an University of electron-science and technology*, 2005.
[4] Shaoyan Wu, Huowang Chen. Approach on simulative evolution by automatic programming design [J]. *Journal of computer*, 1997, 12 (02).
[5] Caifang Wu, Yong Zeng, Yong Qin. The Application of the analytic methods of neural networks to gas-forecasting field [J]. *Advances in earth science*. 2004(10), Vol.19, NO.5, 860-865.
[6] Qin Yijun, E Jiaqiang. Forecast of service life about electron parts of an apparatus used in sanitation ceramic based on grey system theory [J]. *China ceramic industry*. 2004(02), Vol.11, NO.1, 29-30.
[7] Michael R.LYU, Liu Xicheng, Zhong Wanyi. *Handbook of Software Reliability Engineering [M]*. Press of electron industry, 1997:11-14,79-99.
[8] Michael R.Lyu. *Handbook of Software Reliability Engineering[M]*. McGraw-Hill publishing, 1995, ISBN 0-07-039400-8.