

基于经验模式分解与混沌分析的直接多步预测模型

谢景新^{1,2} 程春田² 周桂红³ 孙玉梅¹

摘要 直接多步预测不依靠单步预测的结果而直接使用测量数据, 效果理想, 但往往要求模型能够学习多种不同的目标函数. 本文建立的直接多步预测混合模型, 使用模式分解方法把原始时间序列分解成不同尺度的基本模式分量, 再经混沌分析和神经网络进行组合预测, 减小了各步预测模型之间的差别, 提高了模型对多种目标函数的学习能力, 有效提高了预测精度. 最后, 通过基准时间序列验证了本模型的优越性.

关键词 直接多步预测, 经验模式分解, 混沌分析
中图分类号 TP18

A New Direct Multi-step Ahead Prediction Model Based on EMD and Chaos Analysis

XIE Jing-Xin^{1,2} CHENG Chun-Tian² ZHOU Gui-Hong³ SUN Yu-Mei¹

Abstract The direct multi-step ahead prediction model, which employs observation values and does not depend on the result of single-step prediction, provides more accurate prediction than indirect model. But in this case, the model could be asked to learn various object functions. In this paper, a hybrid model is presented based on empirical mode decomposition (EMD) and chaos analysis. The model employs EMD to decompose the original sequences into many basic modal partitions which can significantly represent potential information of original time series. And chaos features of those data sequences can be used to design DRNN. By these means, the model can be improved to learn various objective functions. And then, more precious prediction can be obtained. Finally, a benchmark time series is tested to display the advantage of this model.

Key words Direct multi-step ahead prediction, empirical mode decomposition, chaos analysis

在时间序列分析和预测中, 直接多步预测模型能够取得与单步迭代模型相当、甚至更好的预测效果, 因而受到许多学者的重视. 随着一些新的非线性技术, 例如神经网络、小波、混沌和局域波法相继出现, 直接多步预测模型又有了新的进展. 从 Weigend 提出在神经网络中引入时间窗的概念^[1] 进行时间序列预测后, 简单的递归神经网络及其改进的网络模型分别被用于直接多步预测^[1-4]. 近年来, 有关学者提出了基于小波分解与重构的时间序列预测模

型, 用于对非线性非平稳信号的预测^[5-7]. 但是小波分析不具有自适应性的特点. 一旦基本小波被选定, 就必须用它来分析所有待分析数据. 一种被称为经验模式分解 (Empirical mode decomposition, EMD) 的方法^[8] 于 1998 年由 Huang 提出后, 经验证在很多方面的应用效果都优于其他的信号处理方法^[8-12].

本文将经验模式分解法和混沌分析相结合, 建立了基于 EMD 和递归神经网络结构模型 EMD-DRNN, 用于非线性、非平稳的时间序列预测. 经太阳黑子年平均数序列测试表明, 该模型是一个性能良好的直接多步预测模型.

1 经验模式分解与基于混沌分析的动态递归神经网络

1.1 经验模式分解法

经验模式分解算法的思路是: 使用局部最大值与最小值分别定义的包络来求均值, 当找到信号中的所有局部极值点后, 其中所有的局部最大值被一个三次样条连接成为上包络, 同理局部最小值生成下包络, 上下包络应将所有的数据都包含在它们之间. 筛选过滤过程的主要作用是去除叠加波, 使波形

收稿日期 2007-02-08 收修改稿日期 2007-07-10
Received February 8, 2007; in revised form July 10, 2007
国家自然科学基金 (50679011, 50479055) 和博士点基金 (20050141008), 河北省科技厅资助项目 (072135125, 072135124) 资助

Supported by National Natural Science Foundation of China (50679011, 50479055), Research Fund for the Doctoral Program of Higher Education of China (20050141008), and Research Fund of the Science and Technology Department of Hebei Province (072135125, 072135124)

1. 河北农业大学机电工程学院 保定 071001 2. 大连理工大学水利信息研究所 大连 116024 3. 河北农业大学信息与技术学院 保定 071001

1. College of Mechanical and Electronic Engineering, Hebei Agricultural University, Baoding 071001 2. Institute of Hydro Informatics, Dalian University of Technology, Dalian 116024 3. College of Information Science and Technology, Hebei Agricultural University, Baoding 071001
DOI: 10.3724/SP.J.1004.2008.00684

更加对称. 这个过程可以被重复多次, 直到 h_{1k} 是一个基本模式分量, 即

$$h_{1k} = h_{1(k-1)} - m_{1k} \quad (1)$$

则从原始数据中处理得到的第一个基本模式分量 C_1 为

$$C_1 = h_{1k} \quad (2)$$

并确定一个迭代停止准则即标准差 S_d , S_d 是通过两个连续处理结果计算得出的, 如下式所示

$$S_d = \sum_{t=0}^T \frac{|h_{1(k-1)}(t) - h_{1k}(t)|^2}{h_{1k}^2(t)} \quad (3)$$

S_d 的典型值可以在 0.2 到 0.3 之间, 从原始数据中分离出 C_1 后, 得到

$$r_1 = X(t) - C_1 \quad (4)$$

由于剩余部分 r_1 仍然包含较长周期分量的信息, 所以 r_1 仍被当作新的数据按上述方法进行处理. 当剩余分量小于预定值或为单调函数时停止. 最终将信号分解成 n 个基本模式分量 C_1, \dots, C_n 和一个剩余分量 r_n , 该剩余分量是一个平均趋势或常数, 即

$$X(t) = \sum_{i=1}^n C_i + r_n \quad (5)$$

1.2 基于混沌分析的动态递归神经网络

由混沌时间序列分析可知: 系统中任一分量的演化都是由与之相互作用的其他分量决定的. 因此, 每个分量的演化过程中都隐含着系统的全部信息. 某动力系统中, 若唯一可观察到的是单变量一维时间序列 $X(t)$, 假设 D 是吸引子的分

维数, 把一维时间序列嵌入到 m 维空间中, 则 $X(t) = \{x(t), x(t - \tau), \dots, x(t - (m - 1)\tau)\}$, 其中, $X(t)$ 表示 t 时刻系统的动力学状态, τ 为延迟时间, m 是嵌入空间矩阵的维数, 由此可以建立相空间到嵌入空间的映射. 则在相空间中必然存在函数 $F(\cdot)$ 使得延时后的状态 $\hat{x}(t + \tau)$ 和当前状态 $X(t)$ 之间满足

$$\hat{x}(t + \tau) = F(X(t)) \quad (6)$$

式中 $F(\cdot)$ 是待寻找的预测函数. 嵌入维数 m 可以通过不同方法获得.

用神经网络模拟可构建上述系统模型. 利用混沌相空间重构理论, 获得有关神经网络结构构建的先验知识, 以便合理构筑网络模型. 当进行混沌时间序列多步预测时, 若时间序列为一个有限离散序列时, 可使用内插法对序列进行扩充, 而后再进行混沌参数的分析. 设 $\eta = 1, 2, \dots$, 在 \mathbf{R}^n 空间中存在光滑函数, 满足

$$\hat{x}(t' + \eta\tau) = F'(X(t')) \quad (7)$$

式中 $\eta\tau = p$ 为预测步数, 证明可见文献 [13]. 最佳延迟和嵌入维数可使用互信息量法和 G-P 算法得到. 直接多步预测的最大特点是: 预测模型的输入数据均为测量值, 据此, 得到的 p 步预测模型可由下式表示

$$\hat{x}(t + p) = f(x(t), x(t - \tau), \dots, x(t - (M - 1)\tau)) \quad (8)$$

式中的 M 为神经网络输入节点数, $\{x(i) | i \leq t + p\}$ 为测量值.

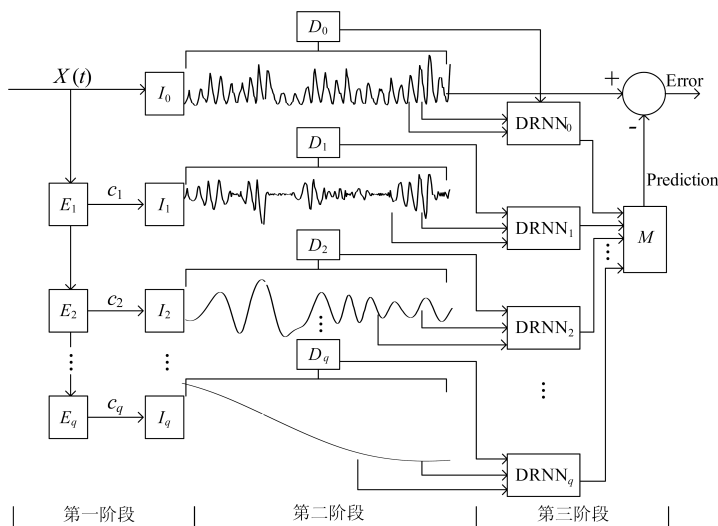


图 1 EMD-DRNN 混合模型结构

Fig. 1 Sketch of EMD-DRNN

2 EMD_DRNN 多步预测模型

2.1 模型结构

本文将经验模式分解法和神经网络结合起来, 建立一个新的神经网络集成结构 (如上页图 1 所示), 经验模式分解算法能充分体现原始信号中包含的信息, 简化了系统间特征信息的干涉或耦合, 并通过一个分离的神经网络预测每一个尺度的模式分量, 大大简化了网络的学习任务. 在第一阶段中, E 为模式分解单元, I 为插值单元; 第二阶段中的 D 为 DRNN 结构设计单元; 在第三阶段中, DRNN 是动态递归神经网络, M 是一个没有隐层的线性神经网络. p 为原始数据经 EMD 法分解后得到的 IMF 分量的个数.

2.2 算法

给定一时间序列 $\{x(t)|t = 1, 2, \dots, n\}$, n 是时间序列的样本点数. 第一步, 原始时间序列通过 EMD 分解成不同的基本模式分量; 第二步, 通过一个分离的神经网络, 对每个模式分量进行预测; 第三步, 通过另一个线性神经网络, 将前一阶段中的每个分量的预测结果作为输入, 对原始序列的下一个样本点进行预测.

1) EMD 分解用于时间序列预测

在第一阶段中, EMD 方法如上节所述, 时间序列 $\{x(t)|t = 1, 2, \dots, n\}$, 被归一化到 $[0, 1]$ 之间, 然后分解该序列, 成为有限个基本模式分量 $\{imf_i(t)|t = 1, 2, \dots, n; i = 1, 2, \dots, q\}$, 方法如第 1.1 节所述, 注意在对每个基本模式分量预测之前, 仍需要对每个分量进行归一化处理, 以提高预测精度.

2) 基于相空间重构理论的动态递归神经网络预测

在序列进入神经网络之前, 使用内插扩展序列, 得到序列 $X'(t') = \{x'(t')|t' = 1, 2, \dots, n'\}$ 及 $\{imf'_i(t')|t' = 1, 2, \dots, n'; i = 1, 2, \dots, q\}$. 若对其进行序列内插, 则称 λ 为内插倍数, 其含义为: 当 $\lambda = 1$ 表示序列中插入了一倍数量的值, 使原始时间序列 $X(t)$ 变为 $X'(t')$, 上式表示为

$$\hat{x}(t' + (\lambda + 1)p) = f(x'(t'), x'(t' - \tau), \dots, x'(t' - (M - 1)\tau)) \quad (9)$$

关于时间延迟, 由 $S(m, N, r, \tau)$ 为基础构造统计量来确定最佳时间延迟 τ' , 在 S 和 τ 关系图上第一个极小值对应最佳延迟 τ' [14]. 嵌入维数的选择如文献 [15] 所述, 使用 $\ln(C_m(r))$ 与 $\ln(r)$ 的关系图求得. 在选取相空间重构中的最佳延迟为神经网络中

的 τ 后, 网络的输入层节点数 $M = m$. 故用于直接多步预测的递归神经网络结构如图 2 所示.

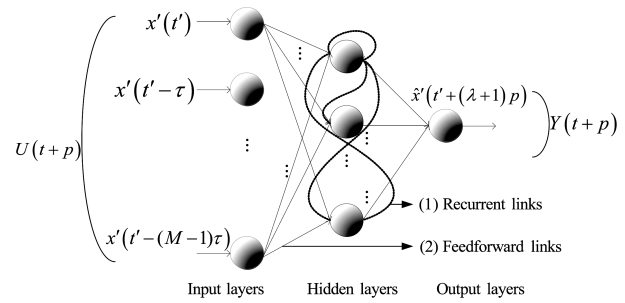


图 2 经内插处理过的数据输入动态递归神经网络

Fig. 2 Data after interpolation for dynamic recurrent neural network architecture

模型可写为:

$$\begin{cases} S(t' + (\lambda + 1)p) = \Theta(W^{[2 \rightarrow 2]}s(t') + W^{[1 \rightarrow 2]}x'(t') + b^{[2]}) \\ \hat{x}(t' + (\lambda + 1)p) = W^{[2 \rightarrow 3]}s(t' + (1 + \lambda)p) + b^{[3]} \end{cases} \quad (10)$$

其中 $\Theta(\cdot)$ 是一个 $N^{[1]}$ 维的向量集合, 包括所有隐层神经元的激活函数, 那么隐层神经元的状态集合可由下式表示

$$S(t' + (\lambda + 1)p) = [s_1^{[2]}(t' + (\lambda + 1)p), s_2^{[2]}(t' + (\lambda + 1)p), \dots, s_{N^2}^{[2]}(t' + (\lambda + 1)p)]^T \quad (11)$$

递归网络的学习算法选用正向传播算法即实时递归算法 [16], 详见文献 [4].

3) 线性组合

在第三步, 所有不同分量的预测结果, $\hat{x}_k(t + p), k = 1, 2, \dots, q + 1$, 通过一个单层的感知器, 线性地组合到一起预测原始时间序列的下一个样本点 $\hat{x}(t + p)$, 如下式

$$\hat{x}(t + p) = \sum_{k=1}^{q+1} w_j^3 \hat{x}_k(t + p) \quad (12)$$

输入层数目等于在第二步中的多层神经网络的个数, 即基本模式分量的个数 q . 该网络没有隐层, 有一个线性输出单元, 用以确定在最终的预测结果中, 每个分量对应的权值 w_j^3 , 其中 $j = 1, 2, \dots, q$. 因为输出单元是线性的, 对于权值向量, 为了计算最小范数解, 它的训练可以通过 Moore-Penrose 伪逆 [5] 来完成.

3 实验结果与讨论

3.1 太阳黑子年平均数序列实验结果

对于 Wolfer Sunspots 数据集, 选择从 1700 年 ~ 1920 年的太阳黑子数据作为训练集, 1921 年 ~ 1954 年和 1955 年 ~ 1979 年期间的数据作为测试集进行测试 (分别用 Set1 和 Set2 表示). 采用标准均方误差衡量预测结果.

1) 将原始序列数据经标准化处理后, 通过 EMD 分解后, 产生 6 个基本模式分量 (如图 3 所示).

2) 根据原始数据和 6 个分解后数据设计 DRNN 的拓扑结构: 由功率谱分析可知, 功率谱的幂函数形式表明: 1) 若对应于周期序列, 谱图具有单峰或几个峰; 2) 对应于湍流或混沌序列, 谱图无明显的峰值或峰连成一片. 分析图 3 所示各序列及其谱分析

图, 可以看出, imf_5 和 imf_6 序列已经具有较平稳的特征, 所以, 本模型使用单输入 DRNN 足以表达其非线性模型, 故只需计算原始数据及其他模式分量的最佳延迟时间与嵌入维数. 由于测量数据是有限的离散数据, 可根据计算需要对其进行不同数量的插值处理, 计算出 $S - \tau$ 关系, 得到原始数据及四个分量的最佳延迟分别为: 6, 6, 4, 7, 9. 将时间延迟代入 G.P 算法计算得到: 序列及模式分量 1 ~ 4 的 $\ln C - \ln r$ 关系图 (图 4), 由此得到嵌入维与关联维数的对应关系 (如表 1 所示), 各序列的嵌入维数为: 10, 10, 8, 14, 12. 因此, 得到 DRNN 输入模式即输入节点个数和时间延迟分别等于序列的嵌入维数和最佳延迟. 隐层节点数的确定采用试凑法, 分别得到前 5 个网络结构为: 10 - 12 - 1、10 - 12 - 1、8 - 6 - 1、14 - 12 - 1、12 - 12 - 1, 另外, 对于 imf_5 和 imf_6 的预测网络为单输入网络: 1 - 3 - 1、1 - 3 - 1.

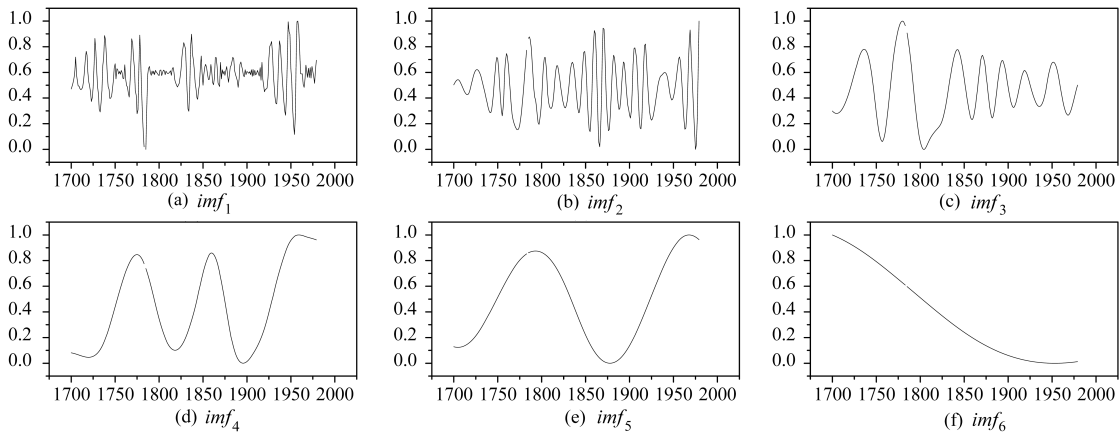


图 3 太阳黑子时间序列经过 EMD 分解后的 6 个基本模式分量

Fig. 3 Six basic modes of yearly sunspots time series decomposed by EMD

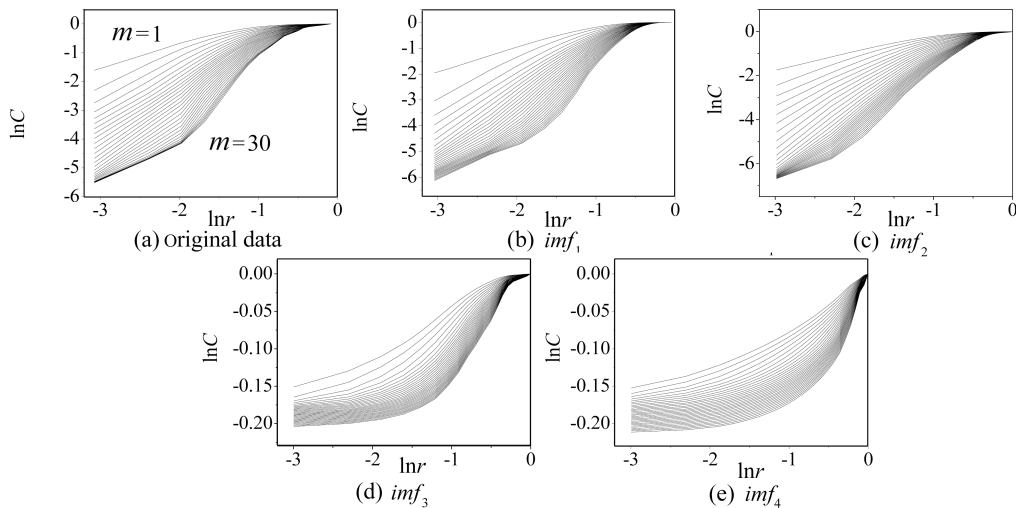


图 4 太阳黑子原始时间序列及模式分量 1 ~ 4 的 $\ln C - \ln r$ 关系图

Fig. 4 Correlation of $\ln C$ vs. $\ln r$ yearly sunspots time series and its four basic modes

表 1 嵌入维数与吸引子的关联维数的关系

Table 1 Correlaton between embedding dimension m and correlation dimension D

	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$	$m=10$
D (original)	0.8334	1.2200	1.4379	1.6245	1.8131	1.9982	2.1802	2.3559	2.5067	2.6158
D (imf_1)	0.0023	0.0036	0.0049	0.0062	0.0075	0.0088	0.0101	0.0115	0.0128	0.0142
D (imf_2)	0.9447	1.5539	1.7705	1.9063	2.0265	2.1307	2.2100	2.2489	2.2527	2.2397
D (imf_3)	0.0420	0.0427	0.0370	0.0314	0.0278	0.0254	0.0236	0.0220	0.0206	0.0195
D (imf_4)	0.0308	0.0309	0.0287	0.0258	0.0236	0.0218	0.0205	0.0193	0.0185	0.0177

	$m=11$	$m=12$	$m=13$	$m=14$	$m=15$	$m=16$	$m=17$	$m=18$	$m=19$	$m=20$
D (original)	2.6889	2.7314	2.7474	2.7368	2.7121	2.6852	2.6578	2.6295	2.5999	2.5676
D (imf_1)	0.0155	0.0169	0.0183	0.0198	0.0212	0.0226	0.0241	0.0255	0.0270	0.0285
D (imf_2)	2.2176	2.1917	2.1618	2.1291	2.0985	2.0608	2.0146	1.9614	1.9041	1.8453
D (imf_3)	0.0184	0.0174	0.0165	0.0157	0.0151	0.0146	0.0141	0.0138	0.0134	0.0132
D (imf_4)	0.0170	0.0164	0.0159	0.0154	0.0150	0.0146	0.0143	0.0140	0.0137	0.0134

3.2 结果讨论

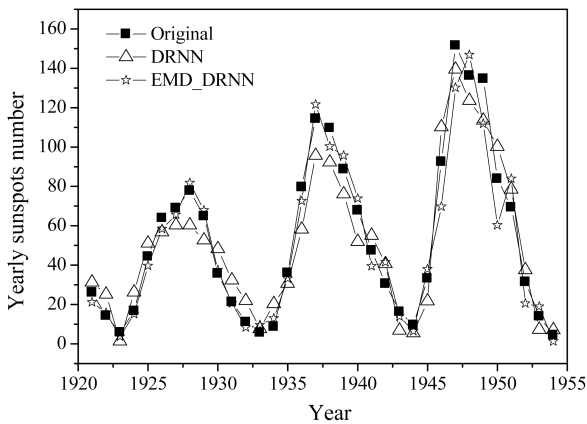


图 5 Set1 和 Set2 的预测结果

Fig.5 Prediction results for Set1 and Set2

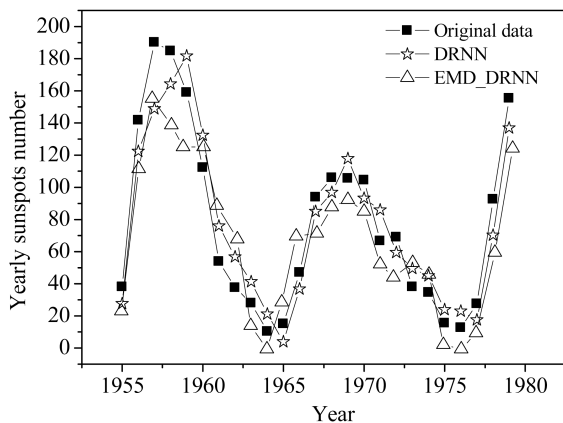


图 6 Set2 最佳预测结果

Fig.6 Best prediction result of Set2

表 2 Set1 和 Set2 的预测结果

Table 2 Predictions for Set1 and Set2

Model	NMSE (Set1)	NMSE (Set2)
Neural nets ^[1]	0.0860	0.3500
TDNN ^[6]	0.0930	0.2460
Wavelet-ANN of Soltani ^[6]	0.0760	0.2300
Wavelet-ANN of Geva ^[5]	0.0570	0.1317
DRNN	0.0839	0.1842
EMD_DRNN	0.0543	0.1313

为了验证本模型的有效性, 本文分别对 Set1 和 Set2 进行了 40 次试验并取平均值, 出于简明性和准确性的考虑, 仅仅用已有文献的相关数据和本模型作比较, 而没有模拟这些模型的结果, 此外还模拟了动态递归神经网络作比较, 其中包括针对太阳黑子年平均数的两个数据集进行直接多步预测的模型: Weigend 的 BP 前馈神经网络模型^[1], Geva 的基于小波分析与神经网络的多尺度综合模型^[5] 和 Soltani 的基于小波分解与神经网络的模型^[6]. 图 5 和图 6 是对 Set1 和 Set2 的最佳预测的预测值, 表 2 为模型对 Set1 和 Set2 的预测误差, 可以看出对 Set1 和 Set2 预测时本模型表现的最好 (用粗体标出), 此外, 本模型对 Set2 的预测误差要比 Set1 的预测有所增加, 但还是有效的. 从预测时间角度看, 由于使用了经验模式分解对原始序列进行处理, 会增加预测的时间复杂度, 但总体上讲, 本模型的运算时间与基于小波和神经网络的预测模型十分相近, 鉴于其较高的预测精度, 这样的耗时还是可以接受的. 显然, 通过模式分解方法, 把太阳黑子时间序列

分解成 6 种不同尺度的基本模式分量, 充分体现了原始信号中所包含的信息, 简化了系统间特征信息的干涉或耦合, 然后, 根据各序列的混沌特性, 设计合理的动态递归神经网络结构, 提高了多步预测的可靠性。

4 结论

本文建立的混合模型, 采用经验模式分解方法与混沌分析相结合, 有效地提高了模型学习多种目标函数的能力, 并设计合理的动态递归神经网络结构, 提高了直接多步预测模型的预测精度。经太阳黑子年平均数序列的测试, 验证了该模型的优越性。

References

- Weigend A S, Huberman B A, Rumelhart D E. Predicting the future: a connectionist approach. *International Journal of Neural System*, 1990, **1**(3): 193–209
- Suykens J A K, Vandewalle J. Learning a simple recurrent neural state space model to behave like Chua's double scroll. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 1995, **42**(8): 499–502
- Lin T, Horne B G, Tino P, Giles C L. Learning long-term dependencies in NARX recurrent neural networks. *IEEE Transactions on Neural Networks*, 1996, **7**(6): 1329–1338
- Parlos A G, Rais O T, Atiya A F. Multi-step-ahead prediction using dynamic recurrent neural networks. *Neural Network*, 2000, **13**(7): 765–786
- Geva A B. Scalenet-multiscale neural-network architecture for time series prediction. *IEEE Transactions on Neural Networks*, 1998, **9**(6): 1471–1482
- Soltani S. On the use of the wavelet decomposition for time series prediction. *Neurocomputing*, 2002, **48**(1): 267–277
- He Guo-Guang, Ma Shou-Feng, Li Yu. A study on forecasting for time series based on wavelet analysis. *Acta Automatica Sinica*, 2002, **28**(6): 1012–1014
(贺国光, 马寿峰, 李宇. 基于小波分解与重构的时间序列预测法. *自动化学报*, 2002, **28**(6): 1012–1014)
- Huang N E, Shen Z, Long S, Wu M, Shih H, Zheng Q. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series. In: *Proceedings of the Royal Society of London Series A*. London, UK: 1998. 903–995
- Loutridis S J. Instantaneous energy density as a feature for gear fault detection. *Mechanical Systems and Signal Processing*, 2006, **20**(5): 1239–1253
- Liang H L, Bressler S L, Desimone R, Fries P. Empirical mode decomposition: a method for analyzing neural data. *Neurocomputing*, 2005, **65-66**: 801–807
- Delechelle A, Nunes J C, Lemoine J. Empirical mode decomposition synthesis of fractional processes in 1D- and 2D-space. *Image and Vision Computing*, 2005, **23**(9): 799–806
- Coughlin K T, Tung K K. 11-year solar cycle in the stratosphere extracted by the empirical mode decomposition method. *Advances in Space Research*, 2004, **34**(2): 323–329

- Han M, Xi J H, Xu S G, Yin F L. Prediction of chaotic time series based on the recurrent network. *IEEE Transactions on Signal Processing*, 2004, **52**(12): 3409–3416
- Zhang Xu-Tao, He Guo-Guang, Yu Lu. A combined algorithm for real-time on-line rapid identification of the chaos in traffic flow. *Systems Engineering*, 2005, **23**(9): 42–45
(张旭涛, 贺国光, 宇卢. 一种在线实时快速地判定交通流混沌的组合算法. *系统工程*, 2005, **23**(9): 42–45)
- Gu Sheng-Shi, Wang Zhi-Qian, Cheng Ji-Tai. The fractal research and predicating on the time series of sunspotrelativenumber. *Applied Mathematics and Mechanics*, 1999, **20**(1): 81–86
(顾圣士, 王志谦, 程极泰. 太阳黑子数时间序列的分形研究及预测. *应用数学和力学*, 1999, **20**(1): 81–86)
- Williams R J, Zipser D. A learning algorithm for continually running fully recurrent neural networks. *Neural Computation*, 1989, **1**(2): 270–280



谢景新 河北农业大学副教授, 博士。主要研究方向为计算智能。本文通信作者。
E-mail: xjx@hebau.edu.cn
(**XIE Jing-Xin** Ph.D., associate professor at Hebei Agricultural University. Her main research interest is computational intelligence. Corresponding author of this paper.)



程春田 大连理工大学教授。主要研究方向为知识管理及其应用。
E-mail: ctcheng@dlut.edu.cn
(**CHENG Chun-Tian** Professor at Dalian University of Technology. His main research interest is knowledge management system and its application.)



周桂红 河北农业大学副教授, 吉林大学博士生。主要研究方向为人工智能。
E-mail: 9902888@163.com
(**ZHOU Gui-Hong** Associate professor at the Hebei Agricultural University, Ph.D. candidate at Jilin University. Her main research interest is artificial intelligence.)



孙玉梅 河北农业大学副教授, 主要研究方向为人工智能。
E-mail: jddqsym@126.com
(**SUN Yu-Mei** Associate professor at the Hebei Agricultural University. Her main research interest is artificial intelligence.)