李嘉康, 赵颖, 廖洪林, 等. 2017. 基于改进 EMD 算法和 BP 神经网络的 SST 预测研究 [J]. 气候与环境研究, 22 (5): 587-600. Li Jiakang, Zhao Ying, Liao Honglin, et al. 2017. SST forecast based on BP neural network and improved EMD algorithm [J]. Climatic and Environmental Research (in Chinese), 22 (5): 587-600, doi:10.3878/j.issn.1006-9585.2017.16180.

基于改进 EMD 算法和 BP 神经网络的 SST 预测研究

李嘉康 赵颖 廖洪林 李其杰

中国人民解放军理工大学理学院,南京 211101

摘 要 海洋表面温度(Sea Surface Temperature,SST)具有非平稳、非线性的特征,直接将处理平稳数据序列的方法应用到非平稳非线性特征明显的序列上显然是不合适的,预测的误差将会很大。为了提高预测精度,更好地解决非平稳非线性序列预测的问题,本文以东北部太平洋(40°N~50°N、150°W~135°W)区域的月平均海洋表面距平温度为例,首先分别应用集合经验模态分解(EEMD)和互补集合经验模态分解(CEEMD)方法将 SST 分解为不同尺度的一系列模态分量(IMF),再运用 BP(Back Propagation)神经网络模型对每一个模态分量进行分析预测,最后将各IMF 预测结果进行重构得到 SST 的预测值。数值试验的结果表明,CEEMD 分解精度比 EEMD 分解精度高,CEEMD 提高了基于 BP 神经网络的预测精度。系列试验统计分析说明应用这种方法对 SST 的 1 年预测是有效的。

 关键词
 集合经验模态分解
 互补集合经验模态分解
 BP 神经网络
 海洋表面月平均温度
 预测

 文章编号
 1006-9585 (2017) 05-0587-14
 中图分类号
 P468.1
 文献标识码
 A

doi:10.3878/j.issn.1006-9585.2017.16180

SST Forecast Based on BP Neural Network and Improved EMD Algorithm

LI Jiakang, ZHAO Ying, LIAO Honglin, and LI Qijie

Institute of Science, PLA University of Science and Technology, Nanjing 211101

Abstract Monthly mean sea surface temperature (SST) is characterized by non-stationary and nonlinear feature. It is obviously unreasonable to apply linear data processing methods directly to non-stationary and nonlinear time series, which would produce large prediction errors. In order to improve the prediction accuracy and better address the non-stationary and nonlinear sequence prediction problem, in this paper, we present an example based on monthly mean SST anomalies (SSTA) of the Northeast Pacific (40°N–50°N, 150°W–135°W). We first use ensemble empirical mode decompose (EEMD) and complementary ensemble empirical mode decomposition (CEEMD) to decompose monthly mean SST into a series of Intrinsic Mode Function (IMF). BP (Back Propagation) neural network model is then utilized to predict each IMF. Finally, the forecast results of each IMF are reconstructed to obtain the predicted value of monthly mean SST. Results of the experiment indicate that the accuracy of CEEMD is better than that of EEMD, and CEEMD has improved the forecast accuracy based on BP neural network. Statistical analysis of the results of a series of experiments shows that this method is effective for SST prediction at the 1-year scale.

Keywords Ensemble empirical mode decomposition, Complementary Ensemble Empirical Mode Decomposition, BP neural network model, Monthly mean sea surface temperature, Prediction

收稿日期 2016-09-27; 网络预出版日期 2017-04-08

作者简介 李嘉康, 男, 1993 年出生, 硕士研究生, 主要从事资料同化方面的研究。E-mail: lijiakang1993@foxmail.com

通讯作者 赵颖, E-mail: zhaoy@mail.iap.ac.cn

资助项目 国家自然科学基金项目 91530204、41430426

Funded by National Natural Science Foundation of China (Grants 91530204 and 41430426)

1 引言

海洋表面温度(SST)对人类的生产生活具有非常重要的影响,由于水的比热容大,海洋温度微小的变化都会对全球的气候产生巨大的影响,我们所熟知的厄尔尼诺现象和拉尼娜现象都是海洋表面温度异常变化所引起的,因此,人们很早就开始了对海洋表面温度的观测,尤其是近年来对海洋表面温度的观测更加重视。对于海洋表面温度的精确观测并且进行有效的预报是十分重要的,提前对SST进行预报可以使人们做好相应的应对措施来减少对日常生产生活的影响,减少不必要的损失。但是由于海洋表面月平均温度变化随机性非常强,非线性和非平稳性特征十分明显,目前还没有一种明确可行且精度较高的方法对海洋表面温度进行有效的预测。

经验模态分解(Empirical Mode Decomposition, EMD) 是由 Huang et al. (1998) 创造性提出的一种 新型的信号处理方法,这种方法可以依据数据自身 的特征将不同频率的信号数据逐级分解,得到若干 个彼此正交的具有周期性和趋势性的信号, 能将非 线性、非平稳性强的信号分解为非线性、非平稳性 较弱的一些系列信号。但是, EMD 方法会出现模 态混叠的问题(Huang et al., 1998; Gai, 2006)。 针对这一问题, Wu and Huang (2011) 通过对原始 数据序列加入不同白噪声提出了集合经验模态分 解方法 (Ensemble Empirical Mode Decomposition, EEMD), 达到了抑制模态混叠的目的。进一步, Yeh et al. (2011) 在原始数据序列中加入两个符号 相反的白噪声,提出了一种 EEMD 的改进算法—— 互补集合经验模态分解方法(Complete Ensemble Empirical Mode Decomposition, CEEMD), 在保证 分解效果与 EEMD 相当的同时,减小了由于添加白 噪声所引起的重构误差。目前 EMD 及其改进算法 已经广泛应用于经济、大气环境、海洋、机械故障 等多个领域(万仕全等, 2005; 毕硕本等, 2012; 蔡吉花等,2014; 王德青等,2014; 张承钊和潘和 平, 2015; Fang, 2015; Liu et al., 2015; Niu et al., 2016)。这些研究和应用都反映出了 EMD 及其改进 算法可以有效降低数据的非平稳性,给后面进一步 的分析处理带来一定的帮助。

针对非线性预测,较为常用的方法有曲线拟合、

均生函数模型、灰色系统模型、神经网络等。其中BP (Back Propagation)神经网络(李友坤,2012)在处理非线性问题上有一定的优势,它的原理简单、操作性强,因此在自动化、经济、环境等众多领域(朱信忠,2005;刘艳荣,2006;王德明等,2012;彭基伟等,2013)中得到了极为广泛的应用。

针对 SST 具有非平稳性强这一特点,本文拟利用 EEMD、CEEMD 和 BP 神经网络方法研究如何提高 SST 预测精度。

2 试验资料简介

近年来的观测发现,东北部太平洋区域的冬季海面温度比往年高出不少,尤其是 2014 年 2 月,这一区域的月平均海表温度距平已经超过 2.5 °C。到了 2014 年的春夏季,高温区域已经扩展到了沿海海域,对沿岸地区的天气和渔民的生活都产生了影响,甚至影响到了美国华盛顿的气温,给当地居民的日常生活和生产劳动都造成了干扰(Bond et al., 2015)。本文选取东北部太平洋(40°N~50°N,150°W~135°W)区域 1982 年 1 月至 2015 年 12 月的 SST 资料,资料长度为 408 个月,实际分析计算时使用的是月平均海温距平序列 SSTA(Sea Surface Temperature Anomaly)(图 1)。从图像直观来看,整体数据显得非常杂乱无章,非线性和随机性强。

3 SSTA 的分解

本文将 EEMD 分解算法、CEEMD 分解算法分别与 BP 神经网络算法结合起来,建立一种新的预测模型。首先对 SSTA 数据进行 EEMD 和 CEEMD 分解,得到若干个本征模函数 IMF_i (Intrinsic Mode Function);再通过 BP 神经网络对每个 IMF_i进行预测;最后将各个 IMF_i重构得到 SSTA 的预测值。

3.1 SSTA 的 EEMD 分解

基于 EEMD 分解算法将图 1 的 SSTA 进行分解, 共分解出 7 个 IMF 分量和 1 个剩余分量 RES (Residue)(图 2)。

从图 2 可以看出,前三个模态分量 IMF₁、IMF₂、IMF₃ 仍然表现出很强的非线性和非平稳性, IMF₄ 至 IMF₇和最后的趋势项 RES 已经具备一定的周期性和较为规律的波动性,非平稳和非线性性质比前 3 个模态有所下降。趋势项 RES 反映了 SSTA 的整

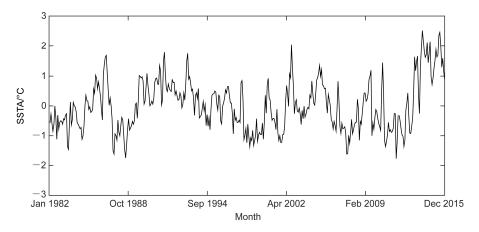


图 1 1982~2015 年东北太平洋(40°N~50°N,150°W~135°W)月平均 SSTA 分布

Fig. 1 Monthly mean SSTA (Sea Surface Temperature Anomaly) over the Northeast Pacific (40°N-50°N, 150°W -135°W) during 1982-2015

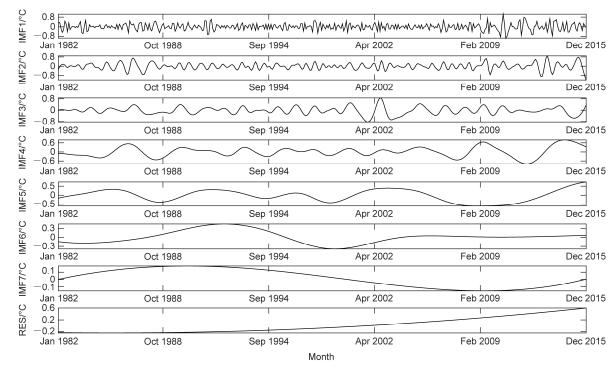


图 2 1982~2015 年东北太平洋月平均 SSTA EEMD 分解的 IMF 分量和趋势项 RES

Fig. 2 IMF (Intrinsic Mode Function) components and the trend item RES (residue) of monthly mean SSTA over the Northeast Pacific from EEMD (Ensemble Empirical Mode Decomposition) decomposition during 1982–2015

体变化趋势从 1982 年以来是逐渐上升的。由于各个模态的非平稳性逐渐降低, EEMD 分解将会降低 非平稳性对预测的影响。

分解的绝对误差 ERR 设为

$$a(t) = \left| S(t) - \left[\sum_{i=1}^{7} I_i(t) + R(t) \right] \right|, \tag{1}$$

其中,a(t) 表示绝对误差 ERR,S(t) 表示原始 SSTA 观测数据, $I_i(t)$ 表示模态分量 IMF 的第 i 个模态, R(t) 表示趋势项 RES。

图 3 为 EEMD 分解的绝对误差,图中显示,分解后的 408 个月的数据误差基本都在 0.008 °C 以下,有 5 个月的数据误差超过 0.01 °C,分别是 1989年 7 月、1993年 9 月、1998年 6 月、1999年 5 月和 2010年 1 月,除 1989年外,另外 4 个误差较大的数据月均出现在厄尔尼诺现象时间内,其中最大误差位于 2010年 1 月,实际值为—0.1195 °C,EEMD分解结果为—0.1314 °C,分解绝对误差为 0.0119 °C;最小误差位于 1987年 2 月,为 1.69×10⁻⁵ °C,EEMD

分解的总体平均绝对误差为 0.0032 °C, 数量级 10⁻³。

3.2 SSTA 的 CEEMD 分解

基于 CEEMD 分解算法将 SSTA 进行分解,共分解出7个 IMF 分量和1个剩余分量 RES(图4)。

对比 EEMD 和 CEEMD 的分解结果可以看到,虽然 CEEMD 分解出的模态与 EEMD 分解出的相应模态 不尽相同,但是这两种分解方法分解出的 8 个模态的非线性和非平稳性都是逐渐降低的,而且最后的

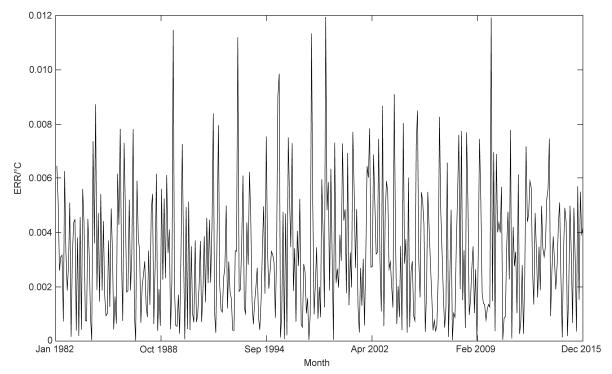


图 3 1982~2015 年东北太平洋月平均 SSTA EEMD 分解绝对误差 ERR

Fig. 3 Absolute error (ERR) of monthly mean SSTA over the Northeast Pacific from EEMD decomposition during 1982-2015

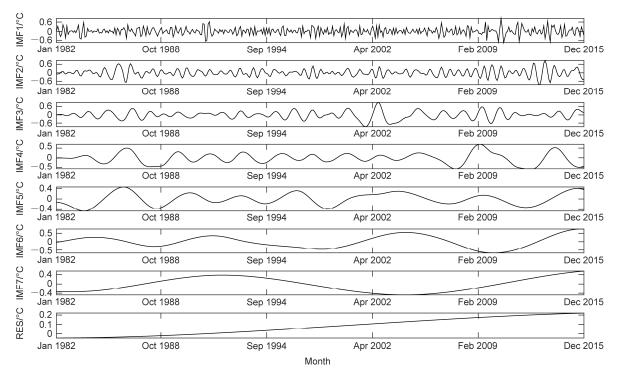


图 4 1982~2015 年东北太平洋月平均 SSTA CEEMD 分解的 IMF 分量和趋势项 RES

Fig. 4 IMF components and the trend item RES of monthly mean SSTA over the Northeast Pacific from CEEMD (Complete Ensemble Empirical Mode Decomposition) decomposition during 1982–2015

趋势项 RES 均为上升趋势,两种分解方法均印证了数据序列整体趋势逐渐升高的特征。

CEEMD 分解绝对误差见图 5,图中显示,分解后的 408 个月数据误差均小于 5×10^{-16} °C,精度非常高。其中误差最大值位于 2015 年 7 月,为 4.4409×10^{-16} °C;最小误差为 0;CEEMD 分解的总体平均绝对误差为 6.2008×10^{-17} °C,数量级在 10^{-17} 。

通过对以上两种分解方法的结果和误差的对比可以看出,改进算法 CEEMD 的误差远远小于 EEMD 分解误差,这是由于 CEEMD 比 EEMD 多加入了符号相反的白噪声,使得在分解效果与 EEMD 相当的情况下,减小了由白噪声引起的重构误差,与其他文献结论一致(郑近德等,2013)。

4 BP 神经网络对 SSTA 模拟预测

为了研究 EEMD和 CEEMD两种方法对预测结果的影响,分析 BP 神经网络预测效果和预测能力,进行如下若干试验:设计 2015年 SSTA 预测试验,分析基于 EEMD和 CEEMD 分解出的不同模态数据的预测效果;基于 2015年的试验结果,设计系列试验分析预测模型的时效性、稳定性、以及对不同

起报月份的敏感性。

4.1 2015 年 SSTA 的 1 年模拟预测试验

2015 年海温异常升高,出现了厄尔尼诺现象(邵勰和周兵,2016),因此,选择这一年进行模拟预测可以验证预测模型对厄尔尼诺现象预报的准确性。试验设计如下:用各模态 1982~2014 年的资料作为分析资料对 BP 网络进行训练,用训练好的网络预测 2015 年 12 个月的 SSTA,并与真实的 2015 年 12 个月的观测值进行对比和分析。

由于 IMF₁至 IMF₃模态的非线性还比较强,我们选择 3 层 BP 网络结构,各月独立分析预测。对 IMF₄及之后的各个模态,由于其非线性和非平稳性相对于前三个模态已经有一定的下降,因此我们直接采用输入层节点和输出层结点个数均为 12 的 BP 网络结构进行训练和预测。

EEMD 各模态分量预测结果如图 6, 预测值与实际值的绝对误差见表 1。通过图 6 和表 1 可以看出,EEMD-BP 预测的第一模态最大误差位于 1 月份,为 0.2216 °C;最小误差为 0.0015 °C,位于 8 月份,第二模态预测效果与第一模态大致相当,前三个模态的预测平均绝对误差均在 0.10 °C 至 0.15 °C 之间。第四和第五模态的平均绝对误差分别为 0.0671 °C 和 0.0092 °C,预测精度和 EEMD 分解

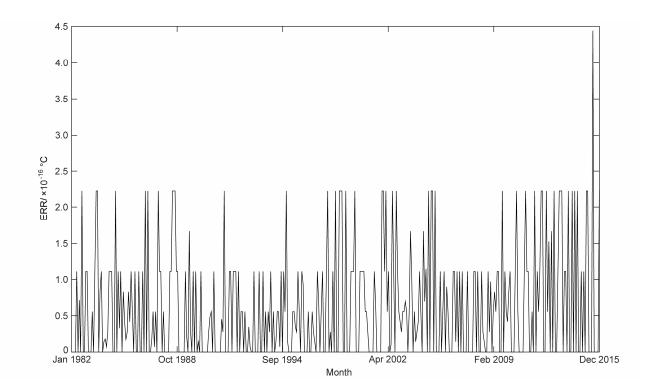


图 5 1982~2015 年东北太平洋月平均 SSTA CEEMD 分解绝对误差 ERR

Fig. 5 Absolute error (ERR) of monthly mean SSTA over the Northeast Pacific from CEEMD decomposition during 1982-2015

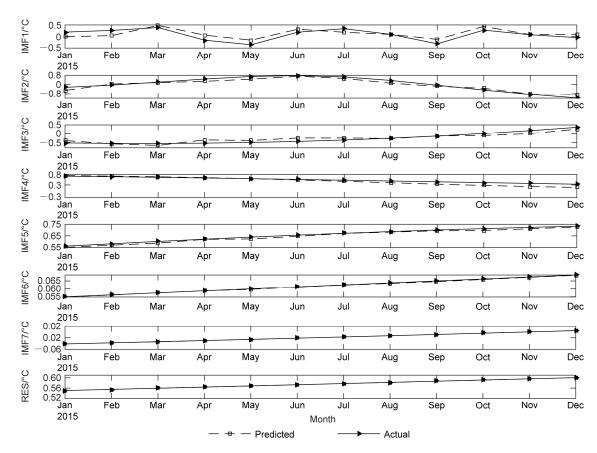


图 6 SSTA EEMD-BP 2015 年各分量预测结果

Fig. 6 SSTA EEMD-BP forecast results of each individual component in 2015

的精度基本保持一致,最后 3 个模态的预测误差数量级为 10⁻⁴。可见随着数据序列非线性和非平稳性的程度降低,预测结果的误差也越来越小。

表 1 SSTA EEMD-BP 各模态预测结果误差分析 Table 1 SSTA Analysis of prediction error of each EEMD-BP mode °C

	最大绝对误差	最小绝对误差	平均绝对误差	均方根误差
IMF_1	0.2216	0.0015	0.1424	0.1597
$IMF_2 \\$	0.2686	0.0300	0.1491	0.1716
$IMF_{3} \\$	0.1922	0.0049	0.1070	0.1224
IMF_4	0.1546	1.7092×10^{-4}	0.0671	0.0861
IMF_5	0.0158	0.0010	0.0092	0.0104
IMF_6	3.9970×10^{-4}	1.9145×10^{-5}	1.7218×10^{-4}	0.0002
IMF_7	5.4704×10^{-4}	1.2487×10^{-5}	1.6107×10^{-4}	0.0002
RES	5.3484×10^{-4}	3.4006×10^{-6}	2.1767×10^{-4}	0.0002

按照同样的方法对 CEEMD 分解出的 8 个模态 进行分析预测,预测结果和误差分析如图 7 和表 2 所示。通过图 7 和表 2 可以看到,CEEMD-BP 预测的第一模态最大误差位于 5 月份,为 0.1843 °C;最小误差为 0.0073 °C;其中 1 月、4 月、8 月、11 月

的预测精度较高。第二模态预测效果与第一模态大致相当,除 5 月、9 月、10 月、11 月这 4 个月的误差较大外,其余月份预测效果均较为理想。前 3 个模态预测效果和真实数据的趋势基本相同。第四模态除了 12 月份误差稍大外,预测效果都比较好。后面几个模态的预测值和实际观测值基本一致。

表 2 CEEMD-BP 各模态预测结果误差分析

Table 2 Analysis of prediction errors of each SSTA CEEMD-BP mode °C

	21			· ·
	最大绝对误差	最小绝对误差	平均绝对误差	均方根误差
IMF_1	0.1843	0.0073	0.0767	0.0999
IMF_2	0.1568	0.0303	0.0753	0.1020
IMF_3	0.1538	0.0154	0.0671	0.0976
IMF_4	0.1051	0.0104	0.0377	0.0453
IMF_{5}	0.0051	8.8828×10^{-5}	0.0026	0.0030
IMF_6	0.0113	6.6653×10^{-5}	0.0049	0.0062
IMF ₇	0.0018	2.7272×10^{-5}	9.6158×10^{-4}	0.0011
RES	2.6836×10^{-5}	2.5860×10^{-6}	1.3428×10^{-5}	1.5851×10^{-5}

将各个模态分量重构后得到 2015 年月平均 SSTA 的预测值(图 8),误差分析见表 3。通过图

5期

No. 5

表可以看到, EEMD-BP 预测方法在 1 月、8 月、11 月误差较大,超过0.3°C,其余月份预测误差较为 理想。CEEMD-BP 预测方法除了 8 月预测误差超过 0.2 °C、10 月份预测误差比 EEMD-BP 方法略大之 外, 其每个月的预测误差均小于 EEMD-BP 方法的

预测误差。从整体上看,CEEMD-BP 预测结果的预 测值和真实值在 0.001 的显著性水平下,相关系数 达到了0.92,比较准确地预测出了2015年的SSTA。 通过分析可以看出, EEMD 和 CEEMD 分解的误差 会影响最终的预测结果,分解误差较小的 CEEMD

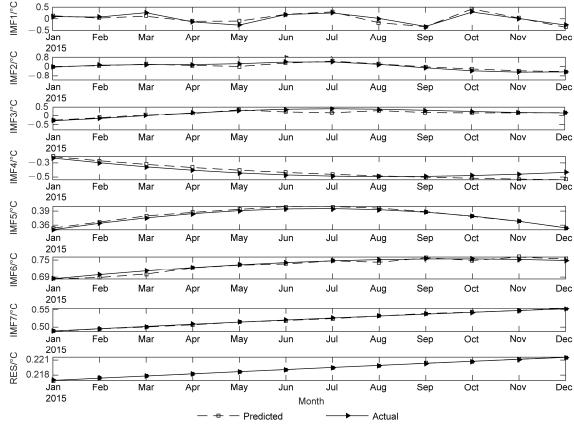


图 7 SSTA CEEMD-BP 2015 年各分量预测结果

Fig. 7 SSTA CEEMD-BP forecast results of each individual component in 2015

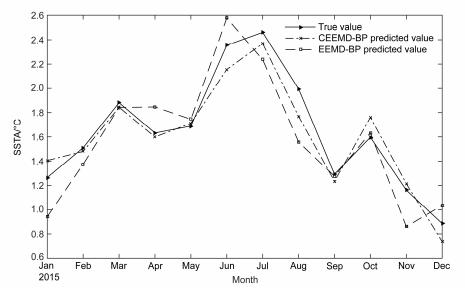


图 8 各个模态分量重构得到的 2015 年月平均 SSTA 预测结果

Fig. 8 Monthly SSTA predicted by the IMF components in 2015

算法在最终的预测结果中的误差也较小,证明了 CEEMD方法在数据分解上比EEMD方法更具有优势。同时还可看到两种预测方法最终的预测误差主 要来自于前三个模态,后 5 个模态的误差对最终的 预测结果影响不大。

表 3 2015 年 SSTA 预测结果误差分析

Table 3 Error analysis of SSTA predictions in 2015 °C

Tubic C Ellor un	analysis of SS 111 predictions in 2016			
	EEMD-BP 绝对误差	CEEMD-BP 绝对误差		
2015年1月	0.3232	0.1406		
2015年2月	0.1363	0.0302		
2015年3月	0.0391	0.0374		
2015年4月	0.2144	0.0333		
2015年5月	0.0557	0.0234		
2015年6月	0.2226	0.1990		
2015年7月	0.2238	0.0916		
2015年8月	0.3376	0.2264		
2015年9月	0.0167	0.0591		
2015年10月	0.0369	0.1641		
2015年11月	0.3014	0.0534		
2015年12月	0.1504	0.1484		
最小绝对误差	0.0167	0.0234		
最大绝对误差	0.3376	0.2264		
平均绝对误差	0.2465	0.1006		
均方根误差	0.2982	0.1219		

4.2 SSTA 的 1 年模拟预测统计试验

通过对 2015 年 1 年 SSTA 预测的试验可以发现,使用 CEEMD 方法比 EEMD 方法误差更小,预测结果更加精确,因此,后续试验均基于 CEEMD 分解进行预测试验。

为了进一步分析 CEEMD-BP 预测 SST 的有效性,下面我们分别对 2006~2015 年的各年 SSTA 进行预测,统计分析 10 个 1 年的预测结果。试验中均按照上述 2015 年的预测试验方法,使用统一的神经网络结构和参数,试验结果如图 9、表 4 所示。通过 10 年的试验预测结果来看,预报的平均绝对误差在 0.3 ℃ 左右,准确率较高。对照 4.1 节中CEEMD-BP 预测 2015 年的 1 年试验结果,本次 10年预测的平均精度有所下降,但是从整体上来看预测值和真实值的相关系数为 0.83,且通过了 99.9%的置信度检验。而且 10 年试验的统计结果说明,使用统一的神经网络结构和参数的 CEEMD-BP 预测模型预测一年 SSTA 是有效的。

4.3 SSTA 的有效预报时间长度模拟预测试验

以上试验说明使用 CEEMD-BP 预测模型预测一年 SSTA 是可行的;为了进一步分析本预测模型的可预报性,下面分别进行试验尝试预测 2 年和预

测 3 年的 SSTA。在预测两年的试验中: 用 1982~2012 年的数据作为训练样本,连续预测 2013~2014年 2 年的 SSTA;在预测 3 年的试验中: 用 1982~2012年的数据作为训练样本,分别进行连续预测和逐年预测 2013~2015年 3 年的 SSTA。

表 4 2006~2015 年 SSTA 预测结果误差分析

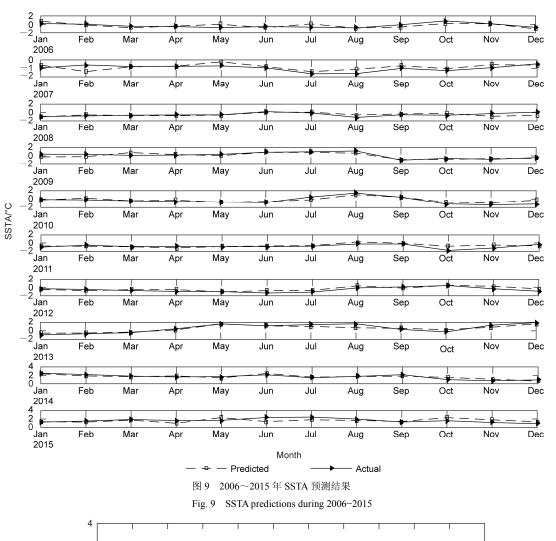
Table 4 Error analysis of SSTA predictions during 2006-

2015				°C
模拟时间	最大绝对误差	最小绝对误差	平均绝对误差	均方根误差
2006年	0.7893	0.0057	0.3815	0.4590
2007年	0.7860	0.0146	0.2969	0.3634
2008年	0.7576	0.0296	0.2923	0.3730
2009年	0.6657	0.0120	0.3067	0.3820
2010年	0.9079	0.0139	0.3376	0.4282
2011年	1.0027	0.0357	0.3245	0.4247
2012年	0.5941	0.0444	0.3411	0.3885
2013年	0.9318	0.0031	0.3242	0.4065
2014年	0.5352	0.0368	0.2614	0.2924
2015年	0.7571	0.0826	0.3706	0.5422

图 10 和表 5、图 11 和表 6 分别是连续预测 2 年、3年的结果和误差分析,试验结果显示,随着 预报时间的延长,误差越来越大,第二年预测值和 实际值的相关性明显降低,预测效果不理想。图 12 和表 7 是 3 年逐年预测的结果和误差分析,从预测 3 年的效果来看,逐年预报比连续预报效果好,但 两种预测方法对第三年的预测值走势均出现了问 题,相关性不到 0.1。另一方面,由于 3 个试验在 进行 BP 神经网络训练时的结构不同: 2013~2014 年的连续 2 年预测时 BP 训练输入层和输出层节点 个数为 24, 预测时连续输出 24 个月; 2013~2015 年连续 3 年预测时 BP 训练输入层和输出层节点个 数为 36, 预测时连续输出 36 个月; 2013~2015 年 逐年预测时 BP 训练输入层和输出层节点个数为 12, 预报时每次输出 12 个月, 并将输出的 12 个月 的预测值加入下一个 12 月预测的训练结构; 不同 的 BP 网络结构导致上述 3 个试验中第一年预测、 第二年预测及第三年预测结果不同,说明 BP 网络

表 5 2013~2014 年 SSTA 连续预测结果误差分析 Table 5 Error analysis of SSTA continuous forecast results during 2013-2014

	最大绝对	最小绝对	平均绝对	均方根误	与实际值
模拟时间	误差/℃	误差/℃	误差/℃	差/°C	相关系数
2013年	0.7716	0.0384	0.4356	0.5023	0.8672
2014年	1.2791	0.0162	0.4924	0.6174	0.5290



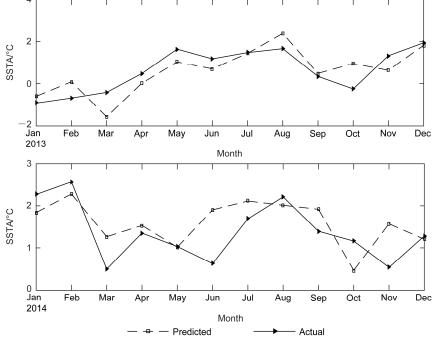


图 10 2013~2014年 SSTA 连续预测结果

Fig. 10 SSTA continuous forecast results during 2013-2014

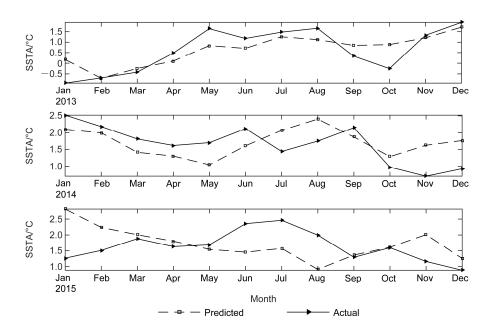


图 11 2013~2015 年 SSTA 连续预测结果

Fig. 11 SSTA continuous forecast results during 2013–2015

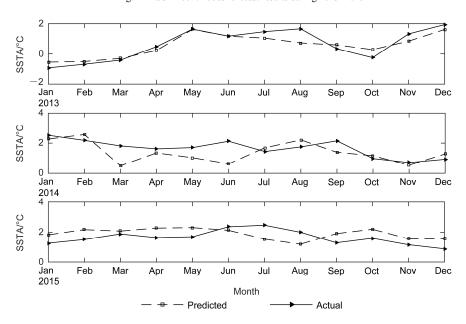


图 12 2013~2015 年 SSTA 逐年预测结果

Fig. 12 SSTA annual forecast results during 2013-2015

表 6 2013~2015 年 SSTA 连续预测结果误差分析 Table 6 Error analysis of SSTA continuous forecast results during 2013-2015

	最大绝对	最小绝对	平均绝对	均方根误	与实际值
模拟时间	误差/℃	误差/°C	误差/℃	差/°C	相关系数
2013年	1.1389	0.0160	0.4740	0.5927	0.8041
2014年	0.9205	0.1840	0.5100	0.5562	0.3128
2015年	1.5468	0.0154	0.5743	0.7431	-0.2585

表 7 2013~2015 年 SSTA 逐年预测结果误差分析

Table 7 Error analysis of SSTA annual forecast results during 2013–2015

	最大绝对	最小绝对	平均绝对	均方根误	
模拟时间	误差/℃	误差/°C	误差/°C	差/°C	相关系数
2013年	0.9318	0.0031	0.3242	0.4065	0.9330
2014年	1.4984	0.1697	0.5474	0.6924	0.4525
2015年	0.9178	0.1921	0.5753	0.6098	0.0365

本身的机制是预测过程中的误差也是最终预测误差的主要来源之一。综上所述,本模型在预测1年时效果比较准确,可以较为准确地预测出正确的趋势走向,预测第二年时效果不理想,预测时间延长到3年则基本不具备可预报性。

4.4 不同月份起报的敏感性预测试验

鉴于前人研究工作中指出的春季可预报性障碍问题(Webster and Yang, 1992; Chen et al., 1995),我们尝试从不同的月份开始预报,分析不同月份起报对预测的敏感性。本试验的设计为:针对 2010年1月至2014年12月,以每年不同月份为起报时间进行12个月的预测,即不同月份起报各5次试验。预测结果为图13~17、表8。

从上面的误差分析可以看到,不同的起报月份 得到的预测平均绝对误差均在 0.33 °C 左右,且预 测误差和 4.2 节"SSTA 的 1 年模拟预测统计试验" 中的预测误差处于同一水平,因此,用本预测模型 对此区域的海洋表面温度进行预测在不同的起报 月份下预测效果基本相同。

5 结论

针对 SST 很强的非线性和非平稳性,本文利用了 EEMD 和 CEEMD 分解算法的 BP 神经网络预测模型进行了尝试,得到如下结论:

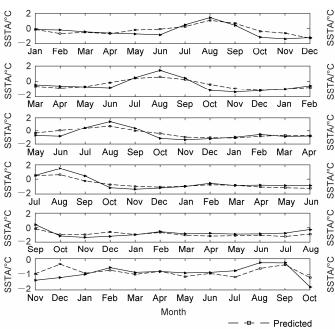


表 8 2010~2014 年不同起报月份 SSTA 预测结果误差分析 Table 8 Error analysis of SSTA forecast results with different forecast start months during 2010-2014 °C

起报月份	最大绝对误差	最小绝对误差	平均绝对误差	均方根误差
1月	0.7828	0.0301	0.2621	0.3295
2 月	0.7113	0.0220	0.3014	0.3735
3 月	0.7664	0.0379	0.3117	0.3733
4月	0.7750	0.0473	0.2850	0.3548
5月	0.6966	0.0518	0.3261	0.4039
6月	0.7571	0.0574	0.3143	0.3838
7月	0.7690	0.0282	0.3503	0.4231
8月	0.7194	0.0146	0.3262	0.4070
9月	0.7429	0.0348	0.3183	0.3892
10 月	0,7792	0.0303	0.3412	0.4178
11 月	0.8031	0.0227	0.3680	0.4558
12 月	0.6970	0.0472	0.3514	0.4541

- (1)对原始数据序列用两种基于 EMD 的改进算法进行处理,得到了若干个模态分量,有效降低了原始序列的非线性和非平稳性。并且使用CEEMD 分解比 EEMD 分解精度高、误差小,这是由于CEEMD 在保证分解效果与 EEMD 相当的情况下,减小了由白噪声引起的重构误差。对比 EEMD-BP 和 CEEMD-BP 两种方法得到的预测结果,后者比前者的预测精度要高,这和 CEEMD 比 EEMD 分解精度高、误差小的结果相一致,两种方法的最终预测误差均主要来自于 BP 神经网络的机制。
 - (2) 从各个模态预测误差和 SSTA 的预测结果

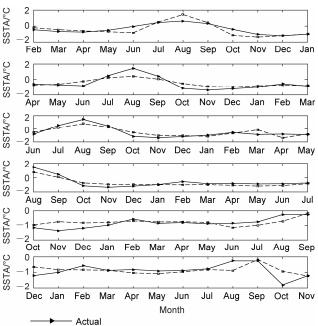


图 13 2010 年不同起报月份 SSTA 预测结果

Fig. 13 SSTA forecast results with different forecast start months in 2010

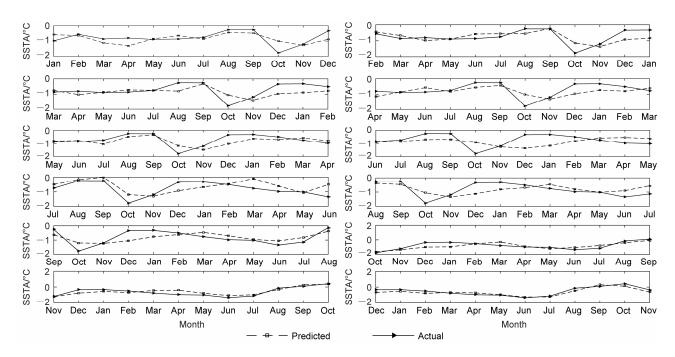


图 14 同图 13, 但为 2011 年

Fig. 14 Same as Fig. 13, but in 2011

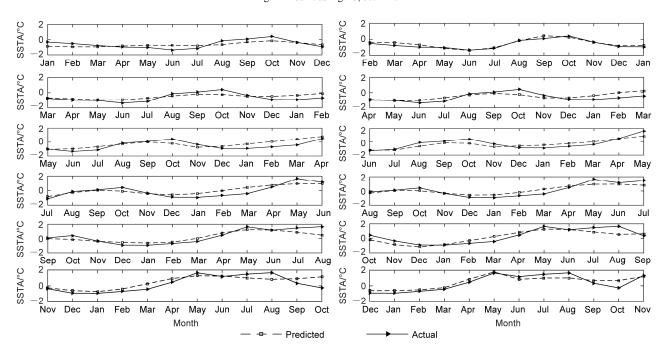


图 15 同图 13, 但为 2012 年

Fig. 15 Same as Fig. 13, but in 2012

误差来看,SSTA 的预测误差主要来自于前三个模态分量的预测,这是由于前三个模态与后面几个模态相比之下依然具有较强的非线性和非平稳性,出现奇异值和波动程度大的数值概率更大,随机性更强。随着非线性的逐渐降低,预测误差逐渐减小。

(3) CEEMD-BP 预测模型在不同起报月份下得到的预测结果均比较理想,误差处于同一水平,因此用该方法在预报文中试验区域的 SSTA 时不存在春季预报性障碍,但不排除其他区域海表温度和预测方法在预报过程中的春季可预报性障碍问题。

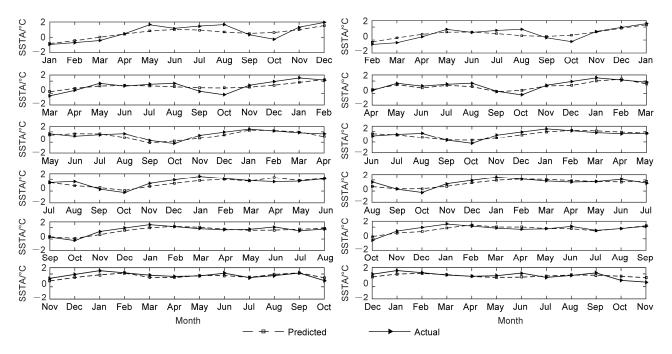


图 16 同图 13, 但为 2013 年

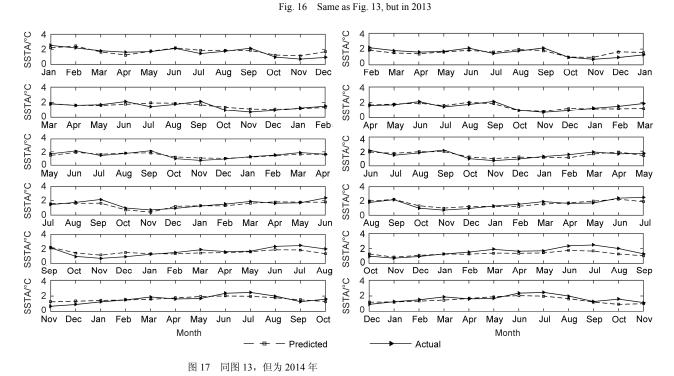


Fig. 17 2014 SSTA forecast results with different forecast start months (unit: °C)

本文基于改进的 EMD 分解算法和 BP 神经网络对 SST 的预测进行了初步研究,试验说明利用 CEEMD-BP 方法预测一年的效果比较准确,特别是超强厄尔尼诺事件出现的 2015 年,这可为今后预测厄尔尼诺现象提供参考;但预测两年及以上时效果不理想,提高预报时长是我们后续研究的重点。

参考文献 (References)

毕硕本, 陈譞, 覃志年, 等. 2012. 基于 EMD 和集合预报技术的气候预测方法 [J]. 热带气象学报, 28 (2): 283–288. Bi Shouben, Chen Xuan, Qin Zhinian, et al. 2012. Study of climate prediction method based on EMD and ensemble prediction technique [J]. Journal of Tropical

- Meteorology (in Chinese), 28 (2): 283-288.
- Bond N A, Cronin M F, Freeland H, et al. 2015. Causes and impacts of the 2014 warm anomaly in the NE Pacific [J]. Geophys. Res. Lett., 42 (9): 3414–3420, doi: 10.1002/2015GL063306.
- 蔡吉花, 张世军, 杨丽. 2014. EMD-SVM 在南京市月平均气温预测中的应用 [J]. 数学的实践与认识, 44 (22): 103–111. Cai Jihua, Zhang Shijun, Yang Li. 2014. Nanjing monthly average temperature prediction base on empirical mode decomposition and support vector machine [J]. Mathematics in Practice and Theory (in Chinese), 44 (22): 103–111.
- Chen D K, Zebiak S E, Busalacchi A J, et al. 1995. An improved procedure for El Niño forecasting: Implications for predictability [J]. Science, 269 (5231): 1699–1702, doi: 10.1126/science.269.5231.1699.
- Fang Y. 2015. A study on the correlations between investor sentiment and stock index and macro economy based on EEMD method [J]. Journal of Financial Risk Management, 4 (3): 206–215, doi: 10.4236/jfrm.2015. 43016
- Gai G H. 2006. The processing of rotor startup signals based on empirical mode decomposition [J]. Mechanical Systems & Signal Processing, 20 (1): 222–235, doi: 10.1016/j.ymssp.2004.07.001.
- Huang N E, Shen Z, Long S R, et al. 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences, 454 (1971): 903–995, doi: 10.1098/rspa.1998.0193.
- 李友坤. 2012. BP 神经网络的研究分析及改进应用 [D]. 安徽理工大学硕士学位论文, 18-55. Li Youkun. 2012. Analysis and improvement applications of BP neural network [D]. M. S. thesis (in Chinese), Anhui University of Science and Technology, 18-55.
- Liu J, Duan J D, Lu H L, et al. 2015. Fault location method based on EEMD and traveling-wave speed characteristics for HVDC transmission lines[J]. Journal of Power and Energy Engineering, 3 (4): 106–113, doi: 10.4236/jpee.2015.34016.
- 刘艳荣. 2006. 基于 BP 网络的彩票预测系统的研究 [D]. 天津大学硕士学位论文, 20–39. Liu Yanrong. 2006. Study on the predictive system of the welfare lottery based on BP network [D]. M. S. thesis (in Chinese), Tianjin University 20–39.
- Niu M F, Wang Y F, Sun S L, et al. 2016. A novel hybrid decompositionand-ensemble model based on CEEMD and GWO for short-term PM2.5 concentration forecasting [J]. Atmos. Environ., 134: 168–180, doi: 10.1016/j.atmosenv.2016.03.056.
- 彭基伟, 吕文华, 行鸿彦, 等. 2013. 基于改进 GA-BP 神经网络的湿度 传感器的温度补偿 [J]. 仪器仪表学报, 34 (1): 153–160. Peng Jiwei, Lv Wenhua, Xing Hongyan, et al. 2013. Temperature compensation for humidity sensor based on improved GA-BP neural network [J]. Chinese

- Journal of Scientific Instrument (in Chinese), 34 (1): 153–160.
- 邵勰,周兵. 2016. 2015/2016 年超强厄尔尼诺事件气候监测及诊断分析 [J]. 气象,42 (5): 540–547. Shao Xie, Zhou Bing. 2016. Monitoring and diagnosis of the 2015/2016 super El Niño event [J]. Meteorological Monthly (in Chinese), 42 (5): 540–547, doi: 10.7519/j.issn.1000-0526. 2016.05.003.
- 万仕全, 封国林, 周国华, 等. 2005. 基于 EMD 方法的观测数据信息提取与预测研究 [J]. 气象学报, 63 (4): 516-525. Wan Shiquan, Feng Guolin, Zhou Guohua, et al. 2005. Extracting useful information from the observations for the prediction based on EMD method [J]. Acta Meteorologica Sinica (in Chinese), 63 (4): 516-525.
- 王德明, 王莉, 张广明. 2012. 基于遗传 BP 神经网络的短期风速预测模型 [J]. 浙江大学学报 (工学版), 46 (5): 837–841, 904. Wang Deming, Wang Li, Zhang Guangming. 2012. Short-term wind speed forecast model for wind farms based on genetic BP neural network [J]. Journal of Zhejiang University (Engineering Science) (in Chinese), 46 (5): 837–841, 904, doi: 10.3785/j.issn.1008-973X.2012.05.010.
- 王德青, 王斐斐, 朱万闯. 2014. 基于 EMD 技术的非平稳非线性时间序列预测 [J]. 系统工程, 32 (5): 138–143. Wang Deqing, Wang Feifei, Zhu Wanchuang. 2014. Prediction for non-stationary non-linear time series based on empirical mode decomposition [J]. Systems Engineering (in Chinese), 32 (5): 138–143.
- Webster P J, Yang S. 1992. Monsoon and Enso: Selectively interactive systems [J]. Quart. J. Roy. Meteor Soc., 118 (507): 877–926, doi: 10.1002/qj.49711850705.
- Wu Z H, Huang N E. 2011. Ensemble empirical mode decomposition: A noise-assisted data analysis method [J]. Advances in Adaptive Data Analysis, 1 (1): 1–41, doi: 10.1142/S1793536909000047.
- Yeh J R, Shieh J S, Huang N E. 2011. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method [J]. Advances in Adaptive Data Analysis, 2 (2): 135–156, doi: 10.1142/ S1793536910000422.
- 张承钊, 潘和平. 2015. 基于前向滚动 EMD 技术的预测模型 [J]. 技术 经济, 34 (5): 70–77. Zhang Chengzhao, Pan Heping. 2015. Prediction model based on forward rolling EMD [J]. Technology Economics (in Chinese), 34 (5): 70–77.
- 郑近德, 程军圣, 杨宇. 2013. 改进的 EEMD 算法及其应用研究 [J]. 振动与冲击, 32 (21): 21–26, 46. Zheng Jinde, Cheng Junsheng, Yang Yu. 2013. Modified EEMD algorithm and its applications [J]. Journal of Vibration and Shock (in Chinese), 32 (21): 21–26, 46.
- 朱信忠. 2005. 车牌自动识别技术的研究与实现 [D]. 国防科学技术大学硕士学位论文, 49–74. Zhu Xinzhong. 2005. The research and implementation of license plate automatic recognition technology [D]. M. S. thesis (in Chinese), National University of Defense Technology, 49–74.