

基于最小二乘与 ARMA 组合模型的海平面异常预报^{*}

孙 文 王庆宾

(解放军信息工程大学地理空间信息学院,郑州 450052)

摘 要 利用最小二乘模型对海平面异常信号中的周期项及线性趋势项进行拟合预报,残差部分则采用 ARMA 模型预报。对北太平洋试验区域的预报结果表明,1 年的预报精度为 0.29 cm,两年的预报精度为 0.63 cm,说明该模型能够以亚厘米级精度有效进行短期 SLA 预报。

关键词 最小二乘;ARMA;组合模型;海平面异常;预报精度

中图分类号:P228

文献标识码:A

SEA LEVEL ANOMALY FORECASTING BASED ON COMBINED MODEL OF LEAST SQUARE AND ARMA

Sun Wen and Wang Qingbin

(*Institute of Geospatial Information, Information Engineering University, Zhengzhou 450052*)

Abstract A method based on the least square (LS) model combined with the ARMA model is used to predict the sea level anomaly (SLA), where the periodic terms and linear trends are fitted using LS while the stochastic terms are predicted with ARMA model. A test was carried out in the North Pacific Ocean by using satellite altimetry data and the results show that the accuracy for 1 year and 2 years' prediction are 0.29 cm and 0.63 cm, respectively. It indicates that this combined model can be used for the short term SLA prediction in sub-centimeter accuracy.

Key words:LS, ARMA, combined model;sea level anomaly, forecasting accuracy

1 引言

海平面变化主要由三方面因素引起:陆地和冰山水系的补充、海水温度的变更及海底构造的变化,这些因素造成海平面高度呈区域性周期振荡及全球性上涨的趋势。海平面异常是由实时海平面扣除平均海平面得到,对其进行精确预报能够为海洋气候变化预测提供重要参考。现今预报算法主要有:时间序列分析方法、遗传算法及神经网络算法。文献[1]基于遗传算法利用验潮站数据对澳大利亚西部

海域的海平面进行了预报,并与神经网络算法进行了比较分析;文献[2、3]探讨了遗传算法在预报海洋表面温度和海洋风速场方面的应用;文献[4、5]基于 AR 模型对全球平均海平面和格网海平面异常进行了预报,并重点研究了该模型在预报厄尔尼诺和拉尼娜现象方面的应用效果。国内的研究热点主要集中在海面温度及风速场的预报^[6-9],对于 SLA 预报研究的相关文献与论述较少。本文利用时间序列分析中的 ARMA 模型与最小二乘模型组合的方法来进行短期 SLA 的预报。

* 收稿日期:2012-04-28

基金项目:国家自然科学基金(41104047,41174026)

作者简介:孙文,男,1987 年生,博士研究生,研究方向为物理大地测量、卫星测高。E-mail: gravitysunwen@gmail.com

2 预报方法

2.1 最小二乘模型

海平面异常信号主要由年周期项、半年周期项及线性趋势项等组成。对于这些非随机信号,可以采用最小二乘方法进行拟合,即 t 时刻海平面异常 $s(t)$ 可用如下表达式来逼近^[5]:

$$\tilde{s}(t) = A_1 \sin(\omega_1 t + \phi_1) + A_2 \sin(\omega_2 t + \phi_2) + at + b \quad (1)$$

其中, A_1 、 A_2 分别为周期项振幅, ϕ_1 、 ϕ_2 为初始相位, a 、 b 为线性趋势项的系数,若时间 t 以 Topex/Poseidon (TP) 卫星的周期 (约为 9.9 天) 为单位,则有 $\omega_1 = 2\pi/36.83$, $\omega_2 = 2\pi/18.42$ 。式 (1) 拟合结果与实际信号 $h(t)$ 的残差值 $r(t)$ 表示为:

$$r(t) = h(t) - \tilde{s}(t) \quad (2)$$

该残差将利用 ARMA 模型进行预报。

2.2 ARMA 模型

若平稳序列 $\{X_t\}$ 满足

$$\sum_{j=0}^p a_j X_{t-j} = \sum_{j=0}^q b_j \varepsilon_{t-j} \quad (3)$$

其中 $a_0 = b_0 = 1$, $\varepsilon_t \sim N(0, \sigma^2)$ 为白噪声序列,则称其为 ARMA(p, q) 序列,式 (3) 称为 ARMA(p, q) 模型^[11]。 p 为 AR 模型的阶数, q 为 MA 模型的阶数。应用 ARMA 模型的关键在于其阶数的确定,文献 [11、12] 对此作了详细分析,在此不再赘述。

应用式 (3) 给出的模型,式 (2) 中的残差 $r(t)$ 可以利用 ARMA 模型进行预报:

$$\tilde{r}(t) = \sum_{j=0}^p a_j r_{t-j} + \sum_{j=0}^q b_j \varepsilon_{t-j} \quad (4)$$

综合式 (1) 和式 (4) 可得 t 时刻海平面异常的预报值为:

$$\tilde{h}(t) = \tilde{s}(t) + \tilde{r}(t) \quad (5)$$

2.3 精度评定

采用差值绝对误差最大值 (AME, Absolute Maximum Error)、均方根 (RMS, Root Mean Square error) 和平均绝对误差 (MAE, Mean Absolute Error) 评定预报结果的精度:

$$Max = \max |\tilde{h}(t_i) - h(t_i)| \quad (6)$$

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n [\tilde{h}(t_i) - h(t_i)]^2} \quad (7)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\tilde{h}(t_i) - h(t_i)| \quad (8)$$

式中, n 为待预报的 TP 周期个数。

3 算例分析

3.1 数据

算例中的数据采用法国 AVISO 中心发布的 DT SLA 数据,其中包括 TP、Jason-1 以及 Jason-2 (1992

年 10 月—2011 年 10 月) 共 992 个周期卫星测高数据。试验区域选择北太平洋,如图 1 所示,包括 $0^\circ \sim 40^\circ \text{N}$ 、 $100^\circ \sim 160^\circ \text{E}$ 之间的区域。数据的详细编辑准则参见文献 [10]。

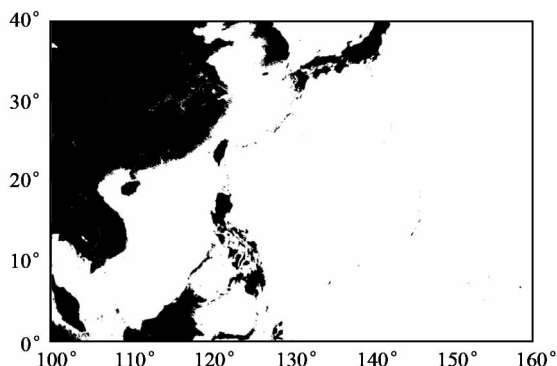


图 1 研究区域位置示意图

Fig. 1 Position of the studied region

对上述各周期 SLA 数据取空间平均,得到该区域 19 年间 SLA 时间序列如图 2。

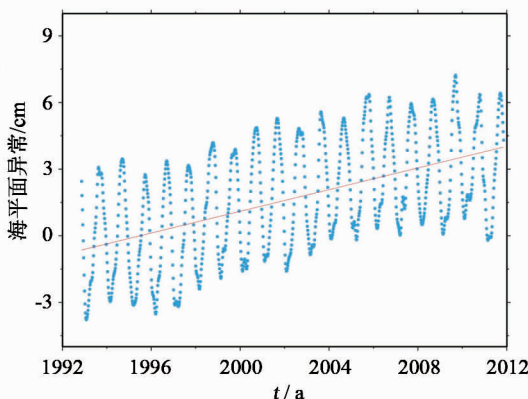


图 2 海平面异常时间序列

Fig. 2 Time series of sea level anomaly

图 2 中的时间序列即可作为预报算法的输入数据,对后续时刻的 SLA 进行预报。

3.2 结果与分析

为了能够对预报结果进行精度评定,将图 2 中的数据分为两个部分,一部分数据作为预报算法的输入数据,另一部分数据用来检核预报结果的精度。根据文献 [13] 的研究成果,预报海平面异常的输入数据时间范围应不小于 162 个 TP 周期,约为 4.4 年。

3.2.1 LS 拟合与信号分离

以 1~165 周期作为预报模型的输入数据,利用式 (1) 对其进行拟合,从信号中分离出的周期项及趋势项如图 3 所示。

将周期项及趋势项从原始信号中分离,得到最小二乘残差项如图 4。

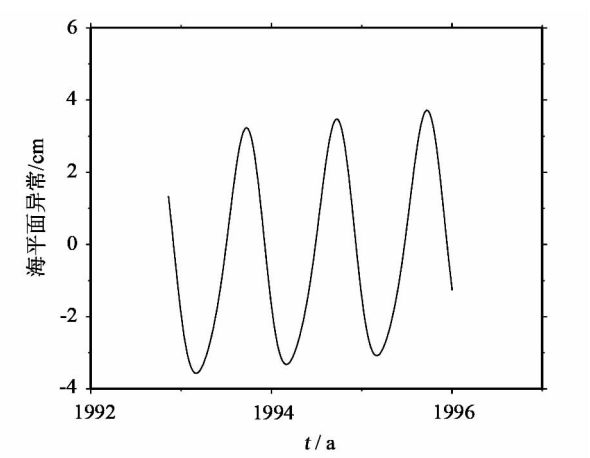


图3 原始信号分离出的周期项及趋势项

Fig. 3 Periodic and trend terms separated from the original signal

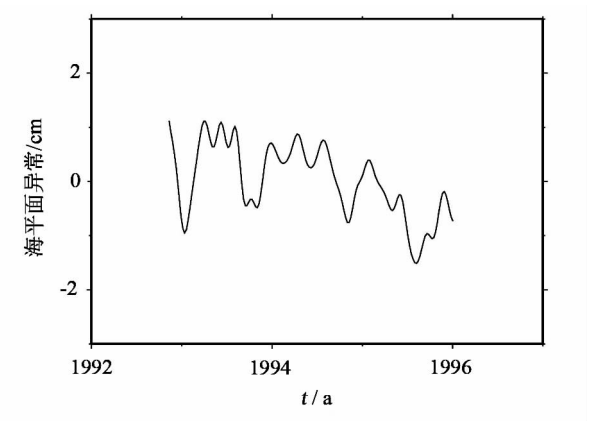


图4 最小二乘残差

Fig. 4 LS residuals

由图4可以看出,分离出的残差项在0 cm附近波动,分布具有随机性。

3.2.2 ARMA 模型残差预报

对图4中的残差用ARMA模型进行拟合,确定最优模型的阶数,最后利用ARMA模型与最小二乘模型组合对SLA进行预报。

利用该组合模型,分别对图1所示区域1、3、6、12及24个月的SLA进行预报,并与图2中的实测数据进行比较,其检核统计结果如表1。

表1 预报结果精度统计(单位:cm)

Tab.1 Accuracy of forecasted results(unit: cm)	1 个月	3 个月	6 个月	12 个月	24 个月
RMS	0.017	0.25	0.32	0.29	0.63
MAE	0.017	0.16	0.22	0.22	0.49
Max	0.02	0.55	0.63	0.63	1.49

图5是24个月预报结果的自相关系数。由图5及表1结果可以看出,利用LS+ARMA组合模型预报SLA时,预报精度随着预报时间长度的增加而

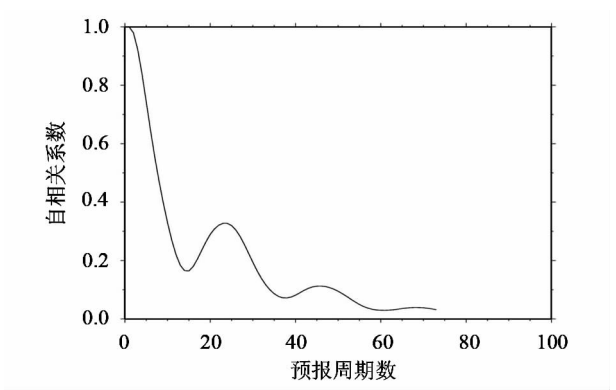


图5 24个月预报结果的自相关系数

Fig. 5 Autocorrelation coefficient of 24 months predicted results

降低,预报24个月的SLA精度约为0.63 cm,最大差值达到1.49 cm。而较短时间的预报结果良好,预报1年以内SLA的精度不超过0.35 cm,最大差值不超过0.65 cm;尤其是1个月的预报结果,其精度达到了0.02 cm,说明该方法可以用来进行短期SLA预报。

另外,由表1可以看出,使用不同的精度评价标准,所得到的评定结果也不同。例如,6个月预报结果的RMS值大于12个月,但MAE与Max值均与其相同。这说明有必要选择适当的精度评价机制和方式。根据表1的结果,建议在进行短期预报时(小于6个月),使用Max值来评价其精度,中长期预报(大于6个月)使用RMS值来进行精度评定,从而能更好地反映算法的预报精度。

4 结论

基于LS+ARMA组合模型,利用TP测高卫星1~165周期数据分别对北太平洋部分区域不同时间长度的SLA进行了预报。当预报时间在1年以内时,该方法的预报结果精度不超过0.35 cm,最大差值不超过0.65 cm,尤其当预报时间为1个月时,精度更是达到了0.02 cm;但当预报的时间长度较长时,预报精度明显下降,例如预报两年的SLA精度为0.63 cm,最大差值达到1.49 cm,此时利用该模型可以对海平面变化作亚厘米级精度的预报,但是精密预报仍然需要更加准确的预报算法。

参 考 文 献

1 Mohammad A G., et al. Sea water level forecasting using genetic programming and comparing the performance with artificial neural networks[J]. Computers & Geosciences, 2010, 36: 620 - 627.

2 Alvarez A., et al. Forecasting the SST space-time variability of

- the Alboran sea with genetic algorithms[J]. *Geophysics Research Letter*, 2000, 27: 2 709 - 2 712.
- 3 Rashmi S, et al. A new technique for forecasting surface wind field from scatterometer observations; A case study for the Arabian sea [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(3): 613 - 620.
 - 4 Niedzielski T and Kosek W. Multivariate stochastic prediction of the global mean sea level anomalies based on TOPEX/Poseidon satellite altimetry [J]. *Artificial Satellite*, 2005, 40: 185 - 198.
 - 5 Niedzielski T and Kosek W. Forecasting sea level anomalies from TOPEX/Poseidon and Jason-1 satellite altimetry [J]. *Journal of Geodesy*, 2009, 83: 469 - 476.
 - 6 李祚泳, 等. 人工神经网络在台风预报中的应用初探 [J]. *自然灾害学报*, 1995, 4(2): 86 - 90. (Li Zuoyong, et al. A primary study on the application of artificial neural network to the forecasting of Typhoon [J]. *Journal of Natural Disasters*, 1995, 4(2): 86 - 90)
 - 7 冯利华. ANN 在海洋预报中的应用 [J]. *海洋预报*, 2000, 17(2): 49 - 55. (Feng Lihua. Application of the neural network in marine forecast [J]. *Marine Forecasts*, 2000, 17(2): 49 - 55)
 - 8 顾锦荣, 等. 基于遗传算法-支持向量机模型在热带气旋强度预报中的应用 [J]. *海洋预报*, 2011, 28(3): 8 - 14. (Gu Jinrong, et al. Application of genetic algorithm-support vector machine model in tropical cyclone intensity forecast [J]. *Marine Forecasts*, 2011, 28(3): 8 - 14)
 - 9 陈旋, 等. 基于权重调整的 BP 神经网络在 Nino 区海温预报中的应用 [J]. *海洋预报*, 2011, 28(5): 61 - 68. (Chen Xuan, et al. Application of BP neural network based on weight adjustment to the temperature forecast [J]. *Marine Forecasts*, 2011, 28(5): 61 - 68)
 - 10 AVISO. DT CorSSH and DT SLA product handbook [S]. CLS-DOS-NT-05.097341, 2010.
 - 11 Broersen P M T. Automatic autocorrelation and spectral analysis [M]. London: Springer-Verlag, 2006.
 - 12 Broersen P M T. Automatic spectral analysis with time series models [J]. *IEEE Transactions on Instrumentation and Measurement*, 2002, 51(2), 211 - 216.
 - 13 Weintrit A. Advances in marine navigation and safety of sea transportation [M]. Gdynia Maritime University, 2007.
-
- (上接第 90 页)
- 707 - 710. (Luo Changlin, et al. A damped least square robust estimation method for spatial three-dimensional rectangular coordinate transformation [J]. *Geomatics and Information Science of Wuhan University*, 2007, 32(8): 707 - 710)
 - 7 Awange L J and Grafarend E W. Linearized least squares and nonlinear Gauss-Jacobi combinatorial algorithm applied to the 7-parameter datum transformation C7 (3) problem [J]. *Zeitschrift fur Vermessungswesen*, 2002, 127: 109 - 116.
 - 8 Awange L J and Grafarend E W. Closed form solution of the over determined nonlinear 7 parameter datum transformation [J]. *AVN*, 2003, 4: 130 - 148.
 - 9 于彩霞, 黄文骞, 樊沛. Bursa 的 3 参数模型与 7 参数模型的适用性研究 [J]. *测绘科学*, 2008, 33(2): 96 - 97. (Yu Caixia, Huang Wenqian and Fan Pei. The adaptability study of Bursa's 3-parameter and 7-parameter models [J]. *Science of Surveying and Mapping*, 2008, 33(2): 96 - 97)
 - 10 韩雪培, 廖帮固. 海岸带数据集成的空间坐标转换方法研究 [J]. *武汉大学学报 (信息科学版)*, 2004, 29(10): 933 - 936. (Han Xuepei and Liao Bangun. Spatial coordinate system transformation for coastal data integration [J]. *Geomatics and Information Science of Wuhan University*, 2004, 29(10): 933 - 936)
 - 11 王解先. 七参数转换中参数之间的相关性 [J]. *大地测量与地球动力学*, 2007, (2): 43 - 46. (Wang Jiexian. Correlations among parameters in seven-parameter transformation model [J]. *Journal of Geodesy and Geodynamics*, 2007, (2): 43 - 46)
 - 12 谢鸣宇, 姚宜斌. 三维空间与二维空间七参数转换参数求解新方法 [J]. *大地测量与地球动力学*, 2008, (2): 104 - 109. (Xie Mingyu and Yao Yibin. A new method for solution of seven-parameter transformation between 3D and 2D spaces [J]. *Journal of Geodesy and Geodynamics*, 2008, (2): 104 - 109)
 - 13 高飞, 童恒庆. 基于改进粒子群优化的非线性最小二乘估计 [J]. *系统工程与电子技术*, 2006, 28(5): 775 - 778. (Gao Fei and Tong Hengqing. Nonlinear least squares estimation based on improved particle swarm optimization [J]. *Systems Engineering and Electronics*, 2006, 28(5): 775 - 778)