

基于 DEMETER 卫星观测数据的电离层离子温度时间序列预测模型

徐方舟^①, 宋现锋^{①,②}, 马灵玲^②, 唐伶俐^②

(^① 中国科学院研究生院, 北京 100049; ^② 中国科学院光电研究院, 北京 100094)

摘要:提出了一种利用 DEMETER 卫星观测量构建电离层离子温度背景场及预测模型的时间序列分析方法。首先,通过分析卫星轨道重访周期与建模格网大小的关系,确定了时间序列采样轨道数据的空间区域大小为 $1.75^{\circ} \times 1.75^{\circ}$ 、时间间隔约为 14 天。然后,基于电离层离子温度明显的季节变化特征,采用自回归移动平均模型 (ARIMA) 构建时间序列预测模型来描述电离层离子温度及其周期性变化。最后,采用 2006 年~2009 年的 DEMETER 卫星观测数据验证了该时间序列预测模型,结果表明 ARIMA 季节模型能较好地模拟电离层离子温度在时间上的变化趋势,建立较为可靠的电离层背景场。

关键词: DEMETER 卫星; 电离层背景场; 离子温度; ARIMA 模型; 时间序列分析

doi: 10.3969/j.issn.1000-3177.2012.02.005

中图分类号: TP79 文献标识码: A 文章编号: 1000-3177(2012)120-0025-06

A Time Series Prediction Model of Ionosphere Ion Temperature(Ti) Based on DEMETER Satellite Dataset

XU Fang-zhou^①, SONG Xian-feng^{①,②}, MA Ling-ling^②, TANG Ling-li^②

(^① Graduate University of Chinese Academy of Sciences, Beijing 100049;

^② Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing 100094)

Abstract: The recent researches revealed there is some relationship of anomalous signals observed by DEMETER satellite sensors with seismic abnormalities. Since ionosphere is affected heavily by solar radiation and earth magnetic field, only after removing those background signals the true abnormalities related to seismic activities can be obtained. So, it is essential to establish a background field reflecting the changes of various features in ionosphere. This paper presents an advanced time-series analysis approach to modeling the background field of ionosphere ion temperature (Ti) using DEMETER earth observation dataset. First, we analyzed the relations between the satellite repeat cycle and modeling grid size, and determined to prepare time-series Ti data with an 14-days interval and falling within a box of $1.75^{\circ} \times 1.75^{\circ}$ centered at a particular geographic location. Then, the ARIMA model (Seasonal Autoregressive Integrated Moving Average) was adopted to model the background field of ionosphere Ti by concerning on the periodic changes of Ti values in time-series dataset. Finally, the observation data of 2006—2009 was used to calibrate and validate the ARIMA model, and the results show that the proposed time-series model could well predict the changes of Ti in future and give a high-quality background field of ionosphere Ti.

Key words: DEMETER satellite; ionospheric background field; temperature-ions; ARIMA model; time-series analysis

收稿日期: 2001-04-01 修订日期: 2011-05-09

基金项目: 国家科技支撑计划项目 (2008BAC35B05)、(2008BAC35B02)。

作者简介: 徐方舟 (1986~), 男, 研究方向: 地理信息系统及数据挖掘。

E-mail: xufangzhou@live.cn

1 引言

近年来,地震工作者研究发现电磁卫星观测到的电离层异常信号与地震前兆变化有着显著的关联性,为开展地震预测及预报工作提供了一种新型技术手段^[1~3]。DEMETER (Detection of Electro-Magnetic Emission Transmitted from Earthquake Regions) 卫星是法国于 2004 年 6 月 29 日发射的一颗专用于监测地震电磁活动的微小卫星,它可实现对全球 65°S~65°N 纬度范围的电离层环境进行连续观测,发射以来为地震电磁耦合机理、地震前兆信息研究积累了宝贵的资料。朱荣等分析 DEMETER 电磁卫星数据发现云南普洱地震前电离层发生扰动^[4];安张辉等对经过汶川震区上空的卫星轨道进行分析,发现汶川地震前 DEMETER 卫星探测的电离层离子温度(Temperature-ions, Ti)发生明显变化^[5],验证了电离层中离子温度的变化与地震异常信号是有关的。

电离层环境复杂,深受太阳活动、季节变化等多种因素影响,为了有效地分离出地震相关活动引起的电离层变化,必须先建立地震前期电离层的背景场。背景场是用来表现电离层对应物理量(如:电场、磁场、离子温度等)自身变化规律的时空模型,目前建立背景场的方法尚处研究阶段,主要采用均值或者统计方法^[3~4,6]。均值法是将一段时间内数据集合的均值作为背景值。这种方法简单易行,但是无法体现物理量随季节变化的特征,也无法辨识数据自身的扰动特征;统计方法是先根据观测数据的累积概率分布图求得当前数据对应的概率值,再对该值进行处理的方法。这种方法也没有解决物理量随时间变化问题。由于不同时间段上数值之间存在差异,如果不将此种差异消除,很难建立高质量的背景场,最终地震异常提取结果也很难准确。时间序列分析方法克服了上述两种方法的缺点,时间序列模型考虑了任何事物发展都具有一定惯性(即延续性)特点,来构建背景场、预测未来变化,其过程简便、适用,短期预测精度也较高。

由于地理位置及地表环境的不同,DEMETER 电磁卫星观测数据存在着明显的空间异质性,同时,相邻区域之间的观测值又具有明显的空间渐变性。空间异质性要求对不同经纬度范围的数据做分块处理,而空间渐变性又可以将相邻经纬度范围的数据关联处理,弥补卫星观测对某一地区重访周期过长引起的数据缺乏缺陷,从而为建立连续固定时间间

隔的时间序列模型提供了条件。针对 DEMETER 电磁卫星的电离层离子温度观测数据,本文分析了利用时间序列分析方法建立电离层离子温度背景场所需要的经纬度网格大小与轨道重访周期的相关关系,提出了利用 ARIMA 季节模型对离子温度时间序列进行建模与预测的方法,深入讨论了利用 ARIMA 模型构建电离层背景场模型的可行性。

2 ARIMA 建模方法

2.1 季节 ARIMA 模型原理

由于季节性变化(包括季度、月度、周度等变化)或其他一些固有因素,某些时间序列观测数据存在着明显的周期性变化,这类时间序列称作季节性序列。ARIMA 模型(Seasonal Autoregressive Integrated Moving Average)是一种典型的时间序列模型^[7~9],利用自回归与分布滞后方法来预测即将发生的未来变化,具体原理如下。

假设已知时间序列 $y_t, t=1 \cdots n$, 其季节性序列变化周期为 s , 构建模型首先用季节差分方法消除时间序列的周期性变化,季节差分算子(Δ_s)定义为公式(1)。

$$\Delta_s = 1 - L^s \quad (1)$$

通常,平稳时间序列可以采用一次季节差分(公式(2)),但是对于非平稳季节性时间序列,有时需要进行 D 次季节差分之后才能转换为平稳的序列。

$$\Delta_s y_t = (1 - L^s) y_t = y_t - y_{t-s} \quad (2)$$

基于差分获得的平稳时间序列,建立周期为 s 的 P 阶自回归 Q 阶移动平均季节时间序列模型。

假定 u_t 是平稳的、非自相关的,则模型为:

$$A_P(L^s) \Delta_s^D y_t = B_Q(L^s) u_t \quad (3)$$

假定当 u_t 非平稳且存在 ARIMA 成分时, u_t 描述为:

$$\Phi_p(L) \Delta^d u_t = \Theta_q(L) v_t \quad (4)$$

其中, v_t 为白噪声过程, p, q 分别表示非季节自回归、移动平均算子的最大阶数, d 表示 u_t 的一阶(非季节)差分次数。由式(4)得

$$u_t = \Phi_p^{-1}(L) \Delta^{-d} \Theta_q(L) v_t \quad (5)$$

把式(5)代入式(3),于是得到季节时间序列模型的一般表达式。

$$\Phi_p(L) A_P(L^s) (\Delta^d \Delta_s^D y_t) = \Theta_q(L) B_Q(L^s) v_t \quad (6)$$

其中,下标 P, Q, p, q 分别表示季节与非季节自回归、移动平均算子的最大滞后阶数, d, D 分别表示非季节和季节性差分次数。式(6)称作 $(p, d, q) \times (P, D, Q)_s$ 阶季节时间序列模型或乘积季节模型。

2.2 电离层离子温度背景场的 ARIMA 模型构建

2.2.1 DEMETER 离子温度观测数据时间序列预处理

ARIMA 模型需要对地表某点以固定时间间隔进行电离层离子温度连续采样的时间序列数据,才可以预测未来变化。DEMETER 电磁卫星每日绕地球飞行约 15 圈,对于地面某点而言,重访轨道很难做到精确重叠。为了形成有效的时间序列数据,我们将地球表面以一固定经纬度间隔划分经纬度格网(研究地点的经纬度坐标为格网中心点),选出经过每个格网覆盖区域的卫星轨道,并收集其相关离子温度观测数据形成网格中心点的时间序列数据。

上述生成的时间序列间隔与经纬度网格间隔密切相关。经纬度网格间隔太小,时间序列间隔将很大,即轨道的重访周期很长,建立时间预测模型将失去预测意义;经纬度网格间隔增大,时间序列间隔将变小,但是时间序列数据质量却降低了,这是因为时间序列内部各数据采集位置之间的空间偏离增大,受空间异质性因素影响,造成了时间序列数据不能够准确反映某点或某区域的电离层温度场的实际变化。可以利用数据波动率来描述这种变化的程度,数据波动率计算方法如下:将原始时间序列序列后一个值与前一个值做对比,求出这些比值的自然对数 $r_i = \ln x_i - \ln x_{i-1}$,得到新序列 r_i ,求出 r_i 的标准差 s , s 即为数据的波动率。

在兼顾重访周期和数据波动率两项指标后,可以确定出网格大小,在一个重访周期内,会有一组卫星数据落在网格内,本文对其取均值来进行处理。

(1) DEMETER 卫星轨道选择原则

在建立某个网格的时间序列时,首先从 DEMETER 电磁卫星观测数据库中通过空间查询选出经过该网格的 DEMETER 卫星离子温度数据,按时间进行数据排序建立起原始的时间序列数据。然后参考 DEMETER 卫星数据获取当日的 K_p 指数以及其昼夜特征,从时间序列中剔除部分不合格的数据。

在地震研究中, K_p 指数用来标识磁静和磁扰,即 $0 \leq K_p \leq 2$ 为磁静, $2 < K_p \leq 9$ 为磁扰。相关研究表明,采用电磁卫星数据建立电离层背景场需要选择磁静时段的 DEMETER 数据,以排除地磁干扰。此外,白天电离层受太阳影响较大,应选择夜间 DEMETER 观测数据以避免太阳辐射干扰。

(2) 经纬度格网采样间隔确定

地震严重影响范围自几十公里至几百公里不等,空间展布变化很大。我们采用 2006 年至 2008 年的 DEMETER 卫星电离层离子温度数据,分析了北纬圈 38° 的 DEMETER 卫星重访周期,网格大小位于 $1^\circ \sim 2^\circ$ 之间,间隔 0.25° 。沿纬线 38° ,以 20° 经度为间隔,分析了 18 个样点处的轨道重访周期和数据波动率。统计分析结果如下:

网格大小越大,重访周期越短,越有利于离子温度背景场的预测。表 1 表明 $1.5^\circ \sim 2^\circ$ 间隔的格网内的轨道重访周期小于 20 日,比较适合短期预测。此外,ARIMA 模型要求时间序列的数据波动率低于 0.3,表 2 显示小于 2° 网格的数据波动率都适中。综合上述因素, 1.5° 和 1.75° 格网大小比较适合建立电离层离子温度 ARIMA 预测模型。这里采用了 1.75° 格网,卫星轨道的重访周期约为 14.28 天约为两周。

表 1 经纬度格网大小与 DEMETER 电磁卫星重访周期的关系

样点位置($^\circ$)		重访周期(天)				
Lon	Lat	1°	1.25°	1.5°	1.75°	2.0°
-170	38	33.45	25.58	19.20	14.72	13.44
-150	38	35.38	28.15	21.22	15.06	12.94
-130	38	38.92	26.31	21.48	15.06	12.25
-110	38	32.17	25.90	17.57	14.00	11.85
-90	38	46.73	24.48	17.54	12.98	10.95
-70	38	59.75	32.10	21.83	15.55	11.47
-50	38	67.00	32.39	21.28	12.73	10.25
-30	38	55.45	35.58	19.49	13.24	10.32
-10	38	95.94	41.56	23.67	13.44	10.23
10	38	64.50	33.40	17.00	11.45	9.58
30	38	44.62	31.28	20.70	14.82	11.20
50	38	86.17	36.22	18.69	13.31	11.11
70	38	33.70	23.76	19.76	14.66	11.42
90	38	36.78	23.56	19.66	15.39	12.56
110	38	38.10	30.44	18.83	14.82	12.74
130	38	37.53	22.62	18.10	15.03	13.48
150	38	32.56	23.59	20.51	15.01	12.74
170	38	55.86	31.40	19.95	15.72	14.32
均值		49.70	29.35	19.80	14.28	11.83

表 2 不同格网大小生成时间序列数据的数据波动率

样点位置(°)		数据波动率				
Lon	Lat	1°	1.25°	1.5°	1.75°	2.0°
-170	38	0.143	0.243	0.126	0.256	0.324
-150	38	0.126	0.226	0.148	0.150	0.319
-130	38	0.181	0.281	0.236	0.262	0.462
-110	38	0.116	0.216	0.225	0.246	0.386
-90	38	0.120	0.220	0.239	0.257	0.461
-70	38	0.208	0.308	0.478	0.237	0.222
-50	38	0.163	0.136	0.250	0.238	0.231
-30	38	0.122	0.222	0.163	0.224	0.326
-10	38	0.121	0.221	0.213	0.391	0.248
10	38	0.124	0.224	0.153	0.292	0.336
30	38	0.126	0.226	0.242	0.248	0.125
50	38	0.186	0.286	0.230	0.256	0.339
70	38	0.161	0.261	0.139	0.243	0.478
90	38	0.122	0.222	0.184	0.156	0.250
110	38	0.131	0.231	0.169	0.250	0.363
130	38	0.126	0.226	0.141	0.262	0.213
150	38	0.148	0.248	0.142	0.246	0.353
170	38	0.136	0.137	0.140	0.257	0.442
均值		0.142	0.229	0.234	0.248	0.326

2.2.2 DEMETER 离子温度 ARIMA 预测模型构建

以 DEMETER 卫星电离层离子温度观测数据的时间序列为基础,构建 ARIMA 预测模型分析未来离子温度变化^[10~12],具体过程如下:

(1)时间序列平稳化

平稳性是 ARIMA 模型时间序列分析的基础,在序列平稳性的基础上才可以对未来观测值进行预测。ARIMA 模型的时间序列具有平稳性特点,针对不平稳的时间序列数据,需要采用公式(2)和(3)进行一阶差分或季节差分从而使序列平稳化。

(2)模型周期识别

序列 x_t 的间隔为 1 的样本的自相关系数定义为:

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^T (x_t - \bar{x})(x_{t-l} - \bar{x})}{\sum_{t=1}^T (x_t - \bar{x})^2} \quad (7)$$

通过对差分时间序列数据的自相关分析,构建自相关图 (Auto-Correlation-Figure, ACF),并依据其自相关特征,确定模型的季节周期 s 。

(3)参数估计与检验

将时间序列数据分为两部分,第一部分数据用来估计模型参数,即 (p, d, q, P, D, Q) ;第二部分数据用来检验模型预测效果。针对效果差的模型,则回到第二阶段重新提取季节周期。

3 结果与分析

3.1 时间序列平稳性分析

(1)时间序列数据准备

为了研究电离层离子温度物理量的自身变化规律,确认其相应的模型周期以及参数,我们选择的数据区域(东经 113°,北纬 31°)在近 3 年内没有发生 5 级以上地震。在该地区 $1.75^\circ \times 1.75^\circ$ 区域范围内,从 DEMETER 电磁卫星数据库中^[13],提取 2006 年~2008 年的离子温度数据,按照 2.2 节所述处理方法进行处理,生成 Ti 物理量(离子温度)的 3 年时间序列数据 x (如图 1(a)所示)。从图中可以看出,该序列呈现逐年下降的趋势(图 1(a)实线所示),每年均有一个高峰值(图 1(a)虚线所示),表现出较为明显的季节性变化。

(2)时间序列数据差分处理

对原始数据进行一阶差分后,去除了逐年下降的趋势,获得新的时间序列(图 1(b)所示),平稳性得到很大改进(图 1(b)实线所示)。对原始数据进行季节差分,去除季节变化量,原始数据同样也去除了逐年下降的趋势(图 1(c)实线所示)。

上述两种方法的差分数据结果的平稳性是否符合要求还要通过序列的自相关图来验证。

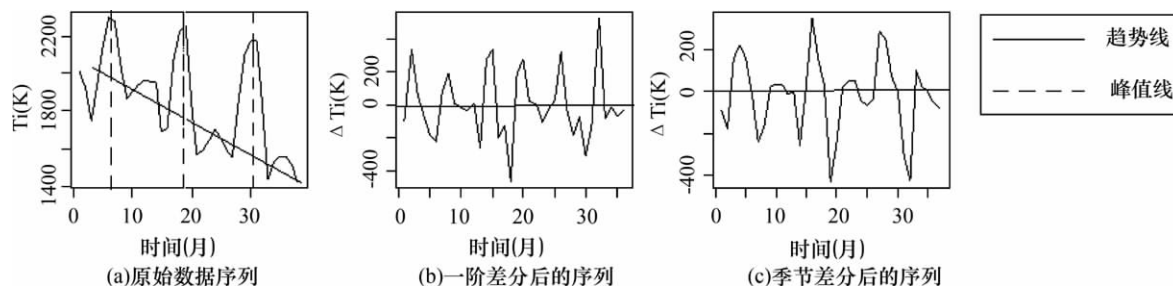


图 1 DEMETER 卫星电离层离子温度时间序列数据平稳性分析

3.2 时间序列自相关分析

对上述差分处理后的时间序列数据 x 进行自相关分析(图 2),横轴的自相关滞后期和纵轴的自相关系数关系表明:序列 x 为非平稳序列,在 1 倍和 12 倍滞后期上存在明显的自相关性。

对 x 序列进行一阶差分,得 Δx 。从 Δx 的自相关图(图 2(b))看出,差分 Δx 序列的平稳性有所改

进,但其在 12 倍滞后期存在着显著的自相关关系。

对 x 进行一次季节性差分,得 $\Delta_s x$ 。从 $\Delta_s x$ 的相关图(图 2(c))看出, $\Delta_s x$ 仍然是非平稳的。

对 x 进行混合差分,即进行一阶差分和一阶季节性差分,得 $\Delta\Delta_s x$ 。从 $\Delta\Delta_s x$ 的相关图(图 2(d))看出, $\Delta\Delta_s x$ 近似为一个平稳过程,因此,确定模型的季节周期 s 为 12 个月(约 50 周)。

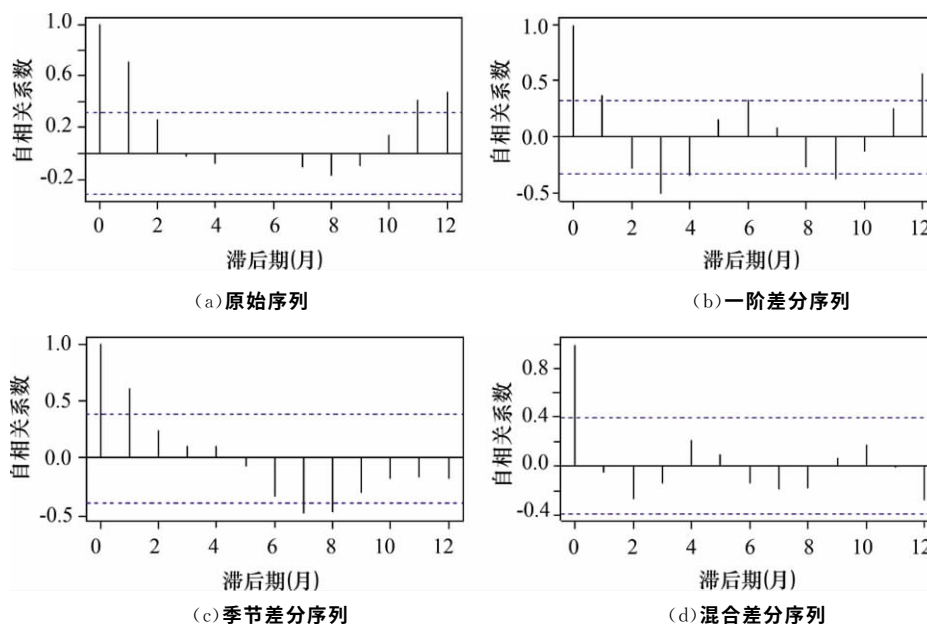


图 2 时间序列数据的自相关分析

3.3 模型参数估计

根据 3.2 的分析,可确定季节和非季节性差分次数 d, d 均为 1,季节与非季节自回归、移动平均算子的最大滞后阶数 P, Q, p, q 一般不会超过 1。因此,采用 $(1, 1, 1) \times (1, 1, 1)_{12}$ 季节时间序列模型进行建模。采用 2006 年 1 月~2008 年 12 月期间观测数据,估计序列 x 的季节时间序列模型参数如表 3 所示,得到结果如下:

$(1+0.5924B)(1-0.5129B^{12})x_t = (1+0.0239B)(1+0.1385B^{12})a_t$ ($\{x_t\}$ 为时间序列, $\{a_t\}$ 为 0 均值独立同分布随机变量序列)

各参数如下:

表 3 ARIMA 预测模型参数

	arl	mal	sral	smal
Coefficients:	-0.2584	0.0239	-0.5129	0.1385
s. e.	0.1725	0.1808	0.2041	0.2566

注:arl 为自回归参数, mal 为滑动平均参数, sral 为季节自回归参数, smal 为季节滑动平均参数。

3.4 模型验证

采用上述模型预测 2009 年前 16 周的离子温度 T_i 的变化趋势,同时对比 DEMETER 卫星实际观测数据(表 4)。结果显示模型预测值的动态变化趋势与实际观测情况基本一致,平均误差 2%,最大误差小于 4.5%,说明采用 $(1, 1, 1) \times (1, 1, 1)_{12}$ 季节时间序列模型是可行的。

表 4 预测值与实际值对比

T_i	2 周	4 周	6 周	8 周	10 周	12 周	14 周	16 周
预测值	1603.8	1544.8	1596.4	1541.3	1519.5	1487.7	1451.5	1360.3
实际值	1571	1554	1531.8	1544.3	1455.1	1463.9	1420.3	1365.5
控制上限	1431.4	1337.1	1341	1245.9	1188.8	1125.2	1059.8	941.3
控制下限	1776.3	1752.6	1851.8	1836.8	1850.1	1850.2	1843.3	1779.3
误差	2.1%	0.6%	4.2%	0.19%	4.4%	1.6%	2.1%	0.38%

注意:控制限为 95%置信区间

2009 年前 16 周的实际离子温度数据虽然与预测值不完全一样,但是各周预测值都落入了预测值的 95% 置信区间范围内(表 4)。图 3 左图实线为物理量的实测值,虚线为预测值,右图是对预测部分的

放大图,从放大图中可以看出预测曲线与实测曲线很接近,说明模型对未来几周的变化情况能够进行较好的跟踪和预测(图 3)。

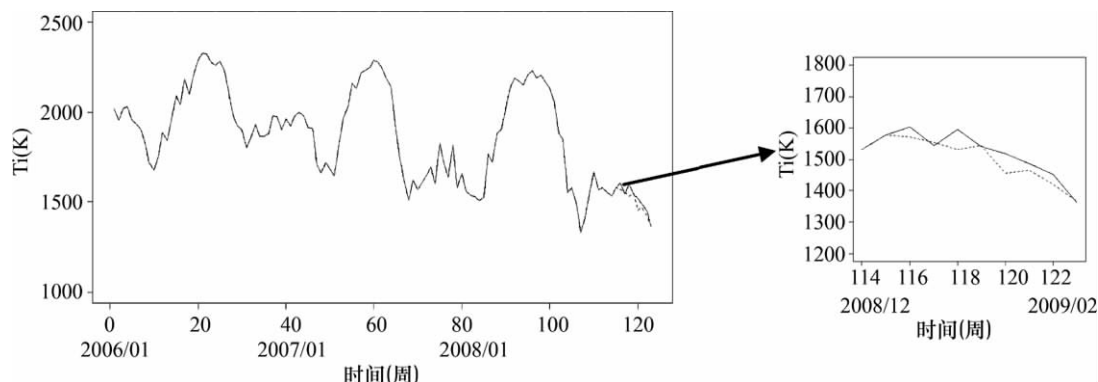


图 3 离子温度 T_i 的预测值与实测值对比

4 结束语

本文通过分析 DEMETER 电磁卫星轨道重访周期和经纬度网格大小的关系,确定了以 $1.75^\circ \times 1.75^\circ$ 网格划分经纬度网格,收集落于网格内的 DEMETER 卫星轨道数据及其电离层离子温度 T_i 观测数据,其重访周期为 14 天、数据波动率小于 3% 比较平稳,生成了以周为单位的离子温度时间序列数据。通过时间序列数据的平稳性分析与自相关分析,确认了 ARIMA 预测模型的周期与参数,构建了电离层离子温度 T_i 的预测模型。采用 2006 年~2008 年的离子温度时间序列建立序列,并用 2009

年前 16 周数据进行验证,结果表明预测值和实际值的周最大相对误差小于 4.5%,说明该模型能够反映电离层电子温度的时间变化周期性特征,能够用于地震前兆异常探测。

目前构建出了单个网格点的在非震条件下基于时间序列的离子温度 ARIMA 预测模型,下一步的工作将结合震例,用文中的方法构建地震发生前的时间序列模型,并预测地震发生时(后)的物理量的值,将此预测值与卫星观测值对比,比直接从卫星观测值序列中寻找异常更有效,从而有可能提取出更加有效的地震前兆异常信息。

参考文献

- [1] Hayakawa M. Atmospheric and ionospheric electromagnetic phenomena associated with earthquakes[C]//Tokyo: TERRAPUB,1999:997.
- [2] Hayakawa M, Molchanov O A. Seismo-Electromagnetics: Lithosphere-atmosphere-ionospheric coupling[C]//Tokyo: TERRAPUB,2002:477.
- [3] 何宇飞. 地震电磁前兆观测[J]. 国际地震动态,2009,32(6):28-36.
- [4] 朱荣,杨冬梅,荆风. DEMETER 卫星观测到的云南普洱地震前的电离层扰动[J]. 地震学报,2008,30(1):76-81.
- [5] 安张辉,范莹莹,刘君. 汶川 M_S8.0 地震前 DEMETER 卫星探测的离子温度变化分析[J]. 地震学报,2011(6):755-759.
- [6] 李志刚,程宗颐,冯初刚,等. 电离层预报模型研究[J]. 地球物理学报,2007,50(2):327-337.
- [7] Brockwell, P. J., Davis, R. A. . Introduction to Time Series and Forecasting[M]. Springer, New York. , Sections 3. 3 and 8. 3.
- [8] Harvey, A. C. Time Series Models[M]. 2nd Edition, Harvester Wheatsheaf.
- [9] Durbin, J. , Koopman, S. J. Time series analysis by state space methods[D]. Oxford University Press.
- [10] Jones, R. H. Maximum likelihood fitting of ARMA models to time series with missing observations[J]. Technometrics 1980(20):389-395.
- [11] 张树京,齐立新. 时间序列分析简明教程[M]. 清华大学出版社,2003:137-138.
- [12] 潘国荣. 基于时间序列分析的动态变形预测模型研究[J]. 武汉大学学报:信息科学版,2005,30(6):484-485.
- [13] 董承玮,宋现锋,芮小平,徐方舟,汪超亮. DEMETER 卫星数据格式解析与数据入库效率研究[J]. 地理与地理信息科学,2010,26(4):42-46.