

基于时间序列ARMA模型的建筑物沉降预测研究

田 渊, 杨 可, 张晓星, 寇少磊, 刘 基, 杨 伟

中国地质调查局西安矿产资源调查中心, 陕西 西安
Email: 670584979@qq.com

收稿日期: 2021年5月17日; 录用日期: 2021年6月24日; 发布日期: 2021年7月1日

摘 要

在我国城市化进程中, 高层建筑物越来越多。为了保证建筑物的正常使用寿命和建筑物的安全性, 很有必要对建筑物进行长期和系统的监测。本文采用时间序列分析法, 利用建筑物观测的15期沉降数据建立ARMA模型, 对未来3期的沉降值进行预测, 通过与实际值进行对比, 从检验效果上看, 此模型预测精度高, 效果好。

关键词

沉降数据, 时间序列分析, ARMA模型

Study on Building Settlement Prediction Based on Time Series ARMA Model

Yuan Tian, Ke Yang, Xiaoxing Zhang, Shaolei Kou, Ji Liu, Wei Yang

Xi'an Mineral Resources Survey Center of CGS, Xi'an Shaanxi
Email: 670584979@qq.com

Received: May 17th, 2021; accepted: Jun. 24th, 2021; published: Jul. 1st, 2021

Abstract

In the process of urbanization in China, there are more and more high-rise buildings. In order to ensure the normal service life of buildings and the safety of buildings, it is necessary to carry out long-term and systematic monitoring of buildings. In this paper, the time series analysis method is used to establish ARMA model based on the settlement data of 15 periods of building observation to predict the settlement value of the next 3 periods. Through the comparison with the actual value,

from the perspective of test effect, this model has high prediction accuracy and good effect.

Keywords

Settlement Data, Time Series Analysis, ARMA Model

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着国民经济的发展, 高层建筑物越来越多, 而建筑物在施工和运营期间, 受到多种主观原因和客观因素的影响, 会产生变形, 变形如果超出了规定的限度, 就会影响建筑物的正常使用, 严重时还会危及建筑物的安全, 给社会和人民生活带来巨大的损失。为了保证建筑物的安全, 避免事故, 因此对建筑物进行变形监测显得尤为重要。

目前对建筑物沉降预测的研究方法有灰色模型法、小波神经网络法、卡尔曼滤波分析法、回归分析法和时间序列分析法。灰色模型法, 是从统计规律的角度进行研究, 各因素之间的关系不明确, 预测精度不高[1]。小波神经网络法在进行预测时存在容易陷入局部最优和训练速度慢等缺点, 需要不断修正预测值与实际值的误差, 而且预测速度慢[2]。卡尔曼滤波分析法应用的前提是首先知道系统的初始情况, 而实际中系统初始情况很难确定, 如果估计初始值与实际值相差很大, 则会导致模型预测的误差变大[3]。回归分析法最大的特点就是在偶然中发现必然, 而建筑物沉降因素很难确定, 因此应用回归分析法预测精度不高[4]。时间序列分析法是一种动态数据的分析法, 能够在数据量很少的情况下对未来进行预测, 和对短期预测比较适合[5]。本文采用时间序列分析法, 对不等时间间隔的建筑物沉降观测数据进行处理, 研究其随时间变化的规律, 利用 Eviews 软件建立 ARMA 模型, 采用逆函数预报法预测后 3 期沉降数据值, 通过与实际观测值进行对比分析, 说明了时间序列分析在建筑物沉降预测的可行性。

2. 时间序列分析法

2.1. 时间序列模型

时间序列分析的基本思想是对于平稳、正态、零均值的时间序列 $\{x_t\}$, 若 x_t 的取值不仅与其前 n 步的各个取值 $x_{t-1}, x_{t-2}, \dots, x_{t-n}$ 有关, 而且还与前 m 步的各个干扰 $a_{t-1}, a_{t-2}, \dots, a_{t-m}$ 有关 ($n, m = 1, 2, \dots$), 按照多元线性回归的思想, 可得到一般的 ARMA 模型[6]:

$$x_t = \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \dots + \varphi_n x_{t-n} - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_m a_{t-m} + a_t \quad (1)$$

公式(1)中, $\varphi_i (i = 1, 2, \dots, n)$ 称为自回归系数, $\theta_j (j = 1, 2, \dots, m)$ 称为滑动平均系数, $\{a_t\}$ 称为白噪声序列。该式称为 x_t 的自回归滑动平均模型, 简记为 ARMA (n, m) 模型。

当 $\theta_j = 0$ 时, 称为 n 阶自回归模型, 简记为 AR (n)。模型变为

$$x_t = \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \dots + \varphi_n x_{t-n} + a_t \quad (2)$$

当 $\varphi_i = 0$ 时, 称为 m 阶滑动平均模型, 简记为 MA (m)。模型变为

$$x_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_m a_{t-m} \quad (3)$$

2.2. 时间序列模型的建模步骤

1) 数据预处理。采用 ADF 检验法对时间序列进行平稳性检验，对于非平稳的时间序列进行差分处理后再进行检验。

2) 模型识别。根据模型自相关系数和偏相关系数的变化特征进行模型识别[7]，见表 1。

Table 1. Model categories

表 1. 模型类别

模型	AR (n)	MA (m)	ARMA (n, m)
自相关系数	拖尾	截尾	拖尾
偏相关系数	截尾	拖尾	拖尾

3) 模型优化。模型优化的目的是选择相对最优的模型。模型优化采取最小信息量准则(AIC)方法，优先考虑的模型是 AIC 值最小的[8]。

4) 模型参数估计。模型参数估计是指求出模型中的自回归系数和滑动平均系数，本文利用极大似然估计法进行参数估计。

5) 模型适用性检验。通过检验模型的残差序列是否为白噪声序列来判断建立的模型是否适用，本文利用 LB 统计量的方法进行检验。

对于给定的置信概率 $1-\alpha$ ，可查 χ^2 分布表中自由度为 m 的 χ^2 值 $\chi_\alpha^2(m)$ ，将 LB 统计量值与 $\chi_\alpha^2(m)$ 进行比较。

当 LB 检验统计量 $\leq \chi_\alpha^2(m)$ ，接受原假设，认为残差序列为白噪声序列，拟合模型显著有效，可用于预测。

当 LB 检验统计量 $> \chi_\alpha^2(m)$ ，拒绝原假设，说明残差序列中还残留着相关信息，不适合用于预测的时间序列数据，拟合模型不显著[9]。

6) 模型预测。本文采取逆函数预报法对 ARMA 模型进行预报，所谓逆函数预报法就是用现在和过去的观察数据通过逆函数对未来值进行预报。

$$\hat{x}_k(l) = \sum_{i=1}^{\infty} \lambda_i^{(l)} x_{k+l-i} \quad (4)$$

公式(4)称为 ARMA (n, m)的逆函数直接预报公式[10]。它避免了利用 x_{k+l-i} 的预报值 $\hat{x}_k(l-i)$ 对 x_{k+l} 进行预报产生的误差。由于公式(4)中右边是无穷项求和，因此在实际预测中，只取有限项和进行预测。

其中系数 $\{\lambda_i^{(l)}\}$ 可由公式(5)求得

$$\begin{cases} \lambda_i^{(l)} = \lambda_i (l=1) \\ \lambda_i^{(l)} = \lambda_{i+l-1} + \sum_{j=1}^{l-1} \lambda_j \lambda_i^{(l-j)} (l>1) \end{cases} \quad (5)$$

3. 工程实例

现以某住宅楼 2019 年 4 月 9 日至 2020 年 8 月 30 日的沉降数据为样本进行分析。根据建筑区的地质情况，结合建筑物的类别，选定沉降观测的精度为二级，在楼上布置 8 个沉降观测点，总共观测了 18 次。

3.1. 数据预处理

本文以 1 号点的沉降数据为样本进行分析，根据前 15 期的观测数据对后 3 期进行预测，沉降数据见表 2。

Table 2. Settlement observation data**表 2.** 沉降观测数据

观测时间	沉降量(mm)	观测时间	沉降量(mm)
2019年4月9日	0	2019年8月5日	1.23
2019年4月20日	-0.21	2019年8月24日	2.23
2019年5月3日	1.04	2019年9月14日	1.73
2019年5月14日	0.70	2019年9月30日	2.13
2019年5月26日	0.97	2019年10月30日	2.47
2019年6月8日	1.07	2019年12月18日	2.26
2019年6月21日	0.58	2020年3月9日	2.44
2019年7月3日	0.98	2020年5月29日	1.32
2019年7月14日	0.69	2020年8月30日	1.02

由于时间序列分析法要求的是等间隔、平稳、正态、零均值的数据[11]。原始数据不利于分析，先将原始数据处理为等时间间隔的观测序列。以 14 天为间隔，对原始数据进行处理，其处理方法如下[12]：

1) 将每两次之间测得的沉降量看成是随时间均匀变化。

2) 根据每次间隔的天数，将每次的沉降量除以天数，得到每次测量的沉降速率。

3) 根据整体观测天数的间隔情况，以 14 天作为间隔，用每次测量的日沉降速率乘以 14，此时得到的数据即为等时间间隔的沉降数据，见表 3，将其绘制成图如图 1 所示。

Table 3. Settlement data with equal time intervals**表 3.** 等时间间隔的沉降数据

期号	沉降量(mm)	期号	沉降量(mm)
1	0	9	0.88
2	-0.27	10	0.78
3	1.12	11	1.64
4	0.89	12	1.15
5	1.13	13	1.86
6	1.15	14	1.15
7	0.62	15	0.64
8	1.14		

对沉降数据利用 Eviews 软件进行 ADF 检验，ADF 统计量的值大于临界值，因此此序列不平稳，需要进行差分处理。从一阶差分开始，如果数据经过一阶差分后仍不平稳，则进行二阶差分，以此类推，对数据进行处理。经过 ADF 检验，二阶差分后的沉降数据符合平稳性要求，图 2 为二阶差分后沉降量变化曲线。

3.2. 模型识别

利用 Eviews 软件计算出序列的自相关函数和偏相关函数如图 3 所示。

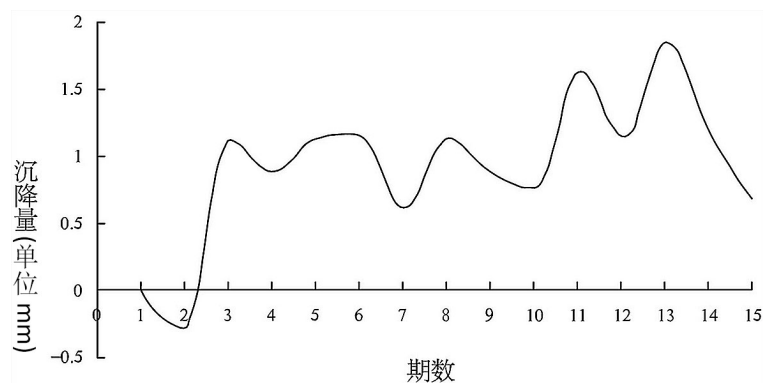


Figure 1. Settlement curve with equal time intervals

图 1. 等时间间隔的沉降量变化曲线

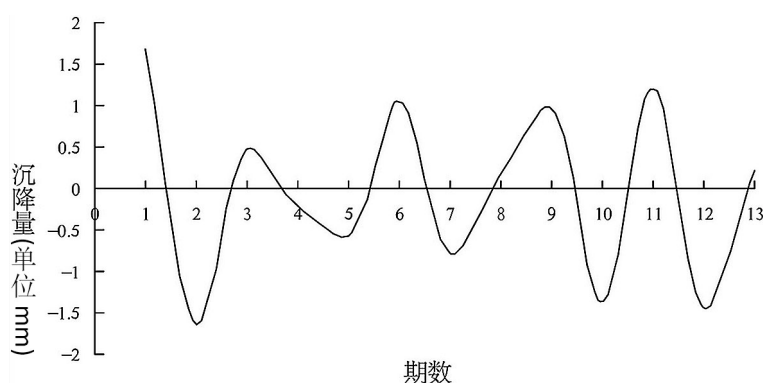


Figure 2. Variation curve of settlement after second order difference

图 2. 二阶差分后沉降量的变化曲线

Autocorrelation	Partial Correlation	AC	PAC	Q-stat	Prob	
		1	-0.700	-0.700	7.9572	0.005
		2	0.259	-0.452	9.1454	0.010
		3	0.123	0.171	9.4389	0.024
		4	-0.425	-0.300	13.353	0.010
		5	0.432	-0.228	17.907	0.003
		6	-0.225	0.053	19.314	0.004
		7	-0.085	-0.227	19.550	0.007
		8	0.326	-0.078	23.689	0.003
		9	-0.353	0.068	29.763	0.000
		10	0.316	0.249	36.243	0.000
		11	-0.194	-0.065	39.915	0.000
		12	0.026	0.075	40.050	0.000

Figure 3. Distribution of autocorrelation function and partial correlation function

图 3. 自相关函数和偏相关函数分布图

图 3 中黑色虚线为标准误差，左边为自相关函数的分布图，右边为偏相关函数分布图，分别用 AC 和 PAC 表示。可以看出自相关系数和偏相关系数除了延迟一阶在 2 倍标准差之外，其余全部在 2 倍标准差范围内波动，这表明序列有明显的短期相关。序列由显著非零的相关系数衰减为小值波动的过程相当持续和缓慢，因此自相关函数和偏相关系数均不截尾，由此可以初步判定模型为 ARMA (n, m) 模型。

对于 ARMA (n, m) 模型来说，图 3 中自相关系数在延迟一阶时显著不为 0，偏相关系数也在延迟一阶时显著不为 0，因此可以取 $n = 1, m = 1$ 来拟合 ARMA (n, m)，进行模型适用性检验。如果检验通过，则

确定 ARMA (n, m) 为适用模型。然而这样的判断有很大的主观性, 因此可以同时建立多个模型 ARMA (1, 1), ARMA (2, 1), ARMA (1, 2) 进行比较。

3.3. 模型优化与参数估计

根据 AIC 对模型进行比较, 通过计算 AIC 值进行比较, AIC 值最小的那个模型就是相对最优模型。三个模型的 AIC 值计算结果见表 4:

Table 4. AIC calculation table

表 4. AIC 计算表

ARMA (n, m)	ARMA (1, 1)	ARMA (2, 1)	ARMA (1, 2)
AIC	3.323	3.452	3.406

从表 4 中可以看出, 当 $n = 1, m = 1$ 时模型的 AIC 值最小, 因此选用 ARMA (1, 1) 模型。利用最大似然法进行模型参数估计, 对应的模型表达式为公式(6):

$$\nabla^2 x_t = -0.374\nabla^2 x_{t-1} + 0.918a_{t-1} + a_t, \quad a_t \sim N(0, 0.581) \quad (6)$$

3.4. 模型适用性检验

将 ARMA (1, 1) 的 LB 统计量与置信概率 95% 的自由度为 m 的 χ^2 值 $\chi_{\alpha}^2(m)$ 进行比较, 可以看出 LB 统计量均小于 $\chi_{0.05}^2(m)$, 结果见表 5, 因此拟合模型显著有效。

Table 5. LB statistics

表 5. LB 统计量

延迟阶数	ARMA (1, 1) LB 统计量	$\chi_{0.05}^2(m)$
1	0.569	3.841
2	0.583	5.991
3	0.596	7.815
4	4.354	9.488
5	5.343	11.071
6	5.884	12.592
7	6.313	14.067
8	7.322	15.507
9	7.322	16.919

3.5. 模型预测

利用 ARMA (1, 1) 对建筑物沉降值进行预测, 并将结果还原为原观测的沉降值, 与实际值进行对比, 结果见表 6。

Table 6. Comparison between predicted value and actual value

表 6. 预测值与实际值对比

期数	预测值(mm)	实际值(mm)	误差绝对值(mm)
16	0.409	0.421	0.012
17	0.215	0.228	0.013
18	0.134	0.153	0.019

从预测结果上看, 误差绝对值都在 0.02 mm 以内, 说明模型的选择是合适的, ARMA 模型能够有效预测建筑物的沉降。

4. 结论

本文在建筑物沉降预测方面做了一些工作: 一是结合某住宅楼的实测沉降数据, 利用时间分析方法建立数学模型; 二是利用逆函数预报法在沉降数据有限的情况下进行预测。通过分析做出以下结论:

- 1、不等时间间隔的数据在进行时间序列分析时, 要将其转换为等时间间隔且平稳的数据来进行建模。
- 2、时间序列分析法能够对建筑物沉降进行预测, 且效果较为理想。
- 3、从预测结果上看, 利用逆函数预报法可以在样本数据有限的情况下进行预测, 且该方法计算简单, 短期预测精度高, 能够为建筑物沉降变形及时做出预警, 确保工程施工的安全和运营安全。

参考文献

- [1] 杨小虎, 朱庆伟, 沈宇恒, 李航. 改进灰色时序模型在建筑物变形监测中的应用[J]. 西安科技大学学报, 2020, 40(5): 919-926.
- [2] 潘纪顺, 李朋辉, 孙凯旋. 小波神经网络在水库流量预测中的应用[J]. 四川建材, 2020, 46(3): 187-189.
- [3] 周晓, 唐宇舟, 刘强. 基于卡尔曼滤波的道路平均速度预测模型研究[J]. 浙江工业大学学报, 2020, 48(4): 392-396.
- [4] 刘富强, 邓文彬, 尚海滨, 等. 多元回归分析法在地铁沉降监测中的应用[J]. 地理空间信息, 2019, 17(3): 105-108.
- [5] 曹嘉晟, 杨太华. 时间序列分析方法在上海地铁能耗预测中的应用[J]. 能源与环境, 2019 (3): 35-37, 40.
- [6] 罗芳琼, 吴春梅. 时间序列分析的理论与应用综述[J]. 柳州师专学报, 2009, 24(3): 113-117.
- [7] 董公何, 赵化放. 时间序列分析在建筑物基坑监测中的应用[J]. 测绘与空间地理信息, 2020(10), 191-194.
- [8] 陈侗. 时空序列模型在铁路构筑物沉降预测中的应用[J]. 铁道勘察, 2020(4): 37-40.
- [9] 李志强. 我国可转债市场弱式有效的实证研究[D]: [硕士学位论文]. 北京: 对外经济贸易大学, 2006.
- [10] 杨位钦. 时间序列分析, 顾岚. 时间序列分析与动态数据建模[M]. 北京: 北京工业学院出版社, 1986.
- [11] 赵仲荣, 袁树才, 张立群. 建筑物差异沉降的时间序列分析与预报[J]. 工程勘察, 2010, 38(1): 73-76.
- [12] 焦明连, 叶秋芹. 基于时序分析的建筑物沉降数据处理[J]. 北京测绘, 2009(1): 15-17+4.