

几种改进的灰色模型在广西年用水量预测中的比较研究

温彩云, 陈国丽, 孙海霞, 涂火年*

广西财经学院, 中国 - 东盟统计学院, 广西 南宁

收稿日期: 2022年11月26日; 录用日期: 2022年12月21日; 发布日期: 2022年12月28日

摘要

文中基于广西壮族自治区2011~2020年的年用水量数据, 采用传统GM(1,1)模型预测, 存在精度不足的问题。为了提高预测的准确性, 对传统GM(1,1)模型采用四种改进方法, 分别是原始序列数据通过函数变换、对GM(1,1)模型预测后的残差修正、对初始序列的弱化算子处理后建模以及本文作者提出的对数变换后灰色建模。通过对比预测结果精度和操作难易程度, 发现预测精度四种方法分别为99.32%, 99.61%, 99.87%, 99.75%; 从操作难易程度上看, 弱化算子法需要对原始数据进行一阶、二阶弱化处理后, 再进行一次累加生产法构建模型, 整个流程相对复杂, 函数变换法和修正误差法操作难度也高于对数变换法, 但低于弱化算子处理法, 对比之下, 对数变换后灰色建模法便有了价值, 对于精度要求不是最高的, 同时又希望快速估算结果的时候就很有价值。

关键词

用水量, 灰色预测模型, 函数改进, 残差修正, 弱化算子, 对数变换

Comparative Study of Several Improved Grey Models in Forecasting Annual Water Consumption in Guangxi

Caiyun Wen, Guoli Chen, Haixia Sun, Huonian Tu*

China-ASEAN Institute of Statistics, Guangxi University of Finance and Economics, Nanning Guangxi

Received: Nov. 26th, 2022; accepted: Dec. 21st, 2022; published: Dec. 28th, 2022

Abstract

In the paper, based on the annual water consumption data of Guangxi Zhuang Autonomous Region

*通讯作者。

文章引用: 温彩云, 陈国丽, 孙海霞, 涂火年. 几种改进的灰色模型在广西年用水量预测中的比较研究[J]. 应用数学进展, 2022, 11(12): 9011-9016. DOI: 10.12677/aam.2022.1112950

from 2011 to 2020, the traditional GM(1,1) model was used for prediction, which had the problem of insufficient accuracy. In order to improve the accuracy of prediction, four improvement methods are used for the traditional GM(1,1) model, which are the original sequence data through function transformation, residual correction after prediction of the GM(1,1) model, modeling after processing of the weakening operator of the initial sequence, and gray modeling after logarithmic transformation proposed by the authors of this paper. By comparing the accuracy of prediction results and the ease of operation, it is found that the four methods of prediction accuracy are 99.32%, 99.61%, 99.87%, and 99.75%, respectively; in terms of the ease of operation, the weakening operator method requires first-order and second-order weakening processing of the original data and then a cumulative production method to construct the model, and the whole process is relatively complicated, and the operational difficulty of the function transformation method and the correction error method is also higher than the logarithmic transformation method, but lower than the weakening operator processing method. In contrast, the gray modeling method after logarithmic transformation is valuable, and it is valuable for those whose accuracy requirements are not the highest, and who also want to estimate the results quickly.

Keywords

Water Consumption, Grey Prediction Model, Function Improvement, Residual Correction, Weakening Operator, Logarithmic Transformation

Copyright © 2022 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

目前常见的城市用水量的预测方法有：可解释性法、多因素回归分析法、灰色预测法、时间序列分析法、神经网络法等。其中，运算方法简单、检验方法容易的灰色预测法对城市用水量的预测较为有效 [1]，并且它还具有所需样本数据少、中短期预测精度高、不要求有明显的分布规律等 [2] [3] [4] [5] 优点。GM(1,1)模型是灰色系统理论中最典型、最传统的模型，但是对于简单的预测使用传统的灰色模型，其预测结果误差较大，可靠性不高。

为了提高预测精度，通过查阅已有的文献，发现有函数变换法、残差修正法、弱化算子法，本文在参考已有文献的基础上提出对数变换思想，并对变换后的数据建立灰色模型预测。对结果对比之后，我们发现弱化算子和对数变换后模型的精度最高，分别为 99.87% 和 99.75%。在两种方法得到精度相近的基础上，由于对数变换法的操作更为简便，且在预测用水量方面更为适用，所以本文提出的对数变换法能得到更为合理的预测结果。

2. 研究背景

纵观广西近几年的发展动态，随着广西区城市规模的持续扩大和经济的迅速发展，人民的生活水平和流动性也在逐步提高，这使得广西年用水量迅速增长，导致供水的复杂性和不确定性日益凸显。精准的城市用水量对于城市水资源的管理、调度、决策有重要的作用，但是目前的城市水资源的管理还使用着传统的水资源管理方法，使得城市的水资源存在着的许多问题。

关于城市用水量的预测方法，国内外的学者对此作了大量的研究，并取得了不少的成果。传统的预测方法包括：递增率法、生长曲线法、双向差分法等，以上方法具有良好的预测准确率，但需要大量的

历史数据资料，且对数据的完整性和可靠性有很高的要求，而广西的用水量序列相对较短，可靠性低，采用上述方法得到的预测结果较为不合理。

灰色预测模型[6]具有需要的历史数据资料较少、运算简便等优势，更加适合用水量序列的预测工作，但传统灰色模型本身存在较多缺陷，主要体现在当数据较为离散时精度差。基于此，有学者针对灰色 GM(1,1) 模型进行了算法改进，Li Zhibin 等[7]利用改进的灰色模型进行了养殖水水质的预测；赵宇哲，武春友[8]等先后通过平移变换和几何平均变换方法弱化序列振动幅度，提高了 GM(1,1)模型的预测精度；韩卫红[9]等使用两种残差对预测数据进行了修正的模型，得到了较高精度的预测；蒋知廷[10]等利用灰色预测模型对矿井涌水量进行了预测；伍德权[11]等利用灰色系统理论对贵阳市城市用水总量进行了预测研究。

对前人的研究进行归纳和学习，本文通过总结并综合运用改进措施，以广西年用水量为研究对象，指标准则选用预测结果的平均相对误差最小值，选取出具有较高精度的改进模型，进行实际的应用研究。

3. 数据收集

本文采用广西区 2011 年~2020 年年用水量数据，数据来源于中国统计局年度数据板块，详细数据见表 1。

Table 1. Annual water consumption data of Guangxi from 2011 to 2020

表 1. 广西区 2011~2020 年年用水量数据

年份	2011 年	2012 年	2013 年	2014 年	2015 年	2016 年	2017 年	2018 年	2019 年	2020 年
用水量/ 10^8 m^3	301.81	303.02	308.17	307.60	299.30	290.60	284.90	287.80	283.40	261.10

4. 灰色预测模型

4.1. 传统灰色 GM(1,1)模型

当前对灰色模型的实际应用中，传统的 GM(1,1)模型的使用最常见，最凸显的优点是简单的建模原理和易于操作。传统的 GM(1,1)模型的本质是使用不同方法对原始数据进行简单处理，使得不稳定的原始数列转换成有规律性的序列，传统的 GM(1,1)模型建立完毕后，进行残差检验、级比检验[12]，检验模型预测值是否准确合理可靠。

4.2. 函数变换改进的灰色 GM(1,1)模型

由于很难通过传统 GM(1,1)模型来判断序列是否呈现随机波动的变化，并且该模型对有波动变化大的序列预测精度较低。通过查阅相关文献，发现有学者提出了利用函数变换的方法对 GM(1,1)模型进行改进。函数变换就是先利用平均变换和几何变换的方式对灰色振荡序列进行预处理[8]，使得序列数据趋于集中，减小数据间的差距，在此基础上建立 GM(1,1)模型。

4.3. 残差修正改进的灰色 GM(1,1)模型

虽然传统的 GM(1,1)对于有明显发展趋势和规律的序列往往有较好预测效果，但实际应用中很难找到有明显规律的样本数据。为了有效的保证 GM(1,1)模型的预测精度，查阅大量文献后发现研究者提出了用残差建立 GM(1,1)模型，使得原模型得到进一步的修正[13]。方法有一阶残差修正、二阶残差修正[14]，前者是使用预测值残差与原始数据进行一次修正，后者是利用原始数据的累加序列，和灰色模型微分方程解的残差做一次修正。

4.4. 弱化算子处理后的 GM(1,1)模型

当数据的产生过程受到外部冲击的干扰时，其内部蕴含的机制会一定程度上发生偏移进而失真，此时，建立常规的灰色预测模型便无法正确描述并预测，为此，刘思峰和邓聚龙共同提出了平均弱化缓冲算子，建立了相应的公理系统，在弱化缓冲算子公理范围下，许多学者从不同角度构建一系列缓冲算子，如杜懿和麻荣永将弱化算子的思想引入到灰色预测模型中[15]，对原始用水量数据进行一阶、二阶弱化处理后，再进行一次累加生产法构建模型。

4.5. 对数变换改进的灰色 GM(1,1)模型

对数变换改进的灰色模型相对于上述三种模型，具有简便实用、操作简单的特点，其原理使用一次累加生产法进行对数变换处理原始用水量数据。设 x_i 为原始的数据序列， x'_i 为对数变换后的新数据序列。公式如下： $x'_i = \log(x_i)$

其中 x 要求大于 0，即为原始用水量数据大于 0，如果有原始数据等于 0 的情况，可对原始数据每个值都加上一个正数进行处理。

当数据量级相差较大时，通常先对数据进行对数转换，然后再进行标准化。使用 \log 函数对数据进行转化可以减小量级差异对结果造成的影响，原理如下：

$$x'_i = \log(a * b) = \log(a) + \log(b)$$

$$x'_i = \log\left(\frac{a}{b}\right) = \log(a) - \log(b)$$

Log 函数转换还可以与 min-max 标准化结合在一起使用，公式如下：

$$x'_i = \frac{\log(x_i) - \log(x)_{\min}}{\log(x)_{\max} - \log(x)_{\min}}$$

本文预测研究中，取原始用水量数据序列为：

$$x_i = \{301.81, 303.02, 308.17, 307.6, 299.3, 290.6, 284.9, 287.8, 283.4, 261.1\}$$

对原始数据进行对数变换处理，得到以下新的数据序列：

$$x'_i = \{5.71, 5.71, 5.73, 5.73, 5.70, 5.67, 5.65, 5.66, 5.65, 5.56\}$$

最终结果见表 2 及表 3。

Table 2. Comparison of prediction results of three improved grey prediction models 10^8 m^3

表 2. 3 种改进灰色预测模型的预测结果比较 10^8 m^3

年份	用水量	传统 GM(1,1)		函数变换		残差修正		弱化算子	
		预测值	相对误差 /%	预测值	相对误差 /%	预测值	相对误差 /%	预测值	相对误差 /%
2011	301.81	301.81	0.00%	301.81	0.00%	301.81	0.00%	302.89	0.00%
2012	303.02	311.65	2.85%	299.89	1.03%	301.63	0.46%	308.40	-0.04%
2013	308.17	306.48	0.55%	313.73	1.81%	308.02	0.05%	307.00	0.07%
2014	307.60	301.40	2.01%	305.11	0.81%	307.35	0.08%	298.96	-0.20%
2015	299.30	296.40	0.97%	297.23	0.69%	297.87	0.48%	290.12	-0.11%

Continued

2016	290.60	291.49	0.31%	292.32	0.59%	292.45	0.64%	284.49	-0.17%
2017	284.90	286.66	0.62%	286.62	0.60%	286.98	0.73%	287.69	-0.14%
2018	287.80	285.90	0.66%	285.39	0.83%	286.37	0.50%	283.50	-0.04%
2019	283.40	280.23	1.12%	280.93	0.87%	285.62	0.78%	261.33	0.04%
2020	261.10	262.63	0.59%	262.53	0.55%	263.53	0.93%	302.89	0.09%

Table 3. GM(1,1) model prediction results after logarithmic processing 10^8 m^3
表 3. 对数处理后的 GM(1,1)模型预测结果 10^8 m^3

年份	真实值	预测值	相对误差/%
2011	301.81	301.81	0.00%
2012	303.02	302.93	-0.03%
2013	308.17	307.41	-0.24%
2014	307.60	308.86	0.41%
2015	299.30	299.23	-0.02%
2016	290.60	289.92	-0.23%
2017	284.90	285.29	0.14%
2018	287.80	286.69	-0.39%
2019	283.40	282.45	-0.34%
2020	261.10	259.69	-0.54%

表 2 表明, 传统 GM(1,1)模型、函数变换法、残差修正模型、弱化算子模型的相对误差都较小, 其中弱化算子处理的灰色模型比传统灰色模型以及函数变换模型的拟合效果更好。表 3 的结果表明, 对数变换灰色模型的平均相对误差 0.23%与弱化算子处理灰色模型十分接近, 表 4 给出了每种模型的平均相对误差和预测精度, 由此可知, 对数变换处理后的灰色模型以及弱化算子处理后的灰色模型, 精度都很高, 预测结果也合理可靠, 但是相对于弱化算子法, 对数变换法操作更简便, 在城市用水量预测中值得广泛使用。

Table 4. 5 Accuracy comparison of grey models

表 4. 5 种灰色模型精度对比

灰色预测模型	平均相对误差/%	预测精度/%
传统灰色预测	0.97	99.00
函数变换改进	0.78	99.32
残差修正改进	0.49	99.61
对数变换改进	0.23	99.75
弱化算子改进	0.11	99.87

查阅相关文献发现, 模型预测效果的评估指标一般为平均相对误差和预测精度, 公式为:

$$AARD = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) / y_i, A = \frac{1}{n} \sum_{i=1}^n 1 - (\hat{y}_i - y_i) / \hat{y}_i$$

式中 $AARD$ 为平均相对误差, \hat{y}_i 为预测值, y_i 为实际值, A 为预测精度。若得到的平均相对误差越小, 说明预测精度越大, 证明模型有很好的预测效果, 更贴合实际情况。

从表 4 可知, 5 种模型的平均相对误差均较小, 其中对数变换改进的灰色模型相对误差为 0.23%, 对数变换改进的灰色模型预测精度高达 99.75%。

5. 结论

1) 广西用水量预测研究中存在着可使用的历史数据资料少、信息不充分等问题, 针对上述问题选择传统灰色模型较为可靠合理。

2) 先对用水量的原始数据进行简单的函数变换, 再使用变换后的数据进行灰色建模, 这样能使原序列的随机性得到有效的弱化, 预测的准确性也得到了提高。

3) 当传统的灰色 GM(1,1)模型预测结果不明显时, 为了可以更加有效地凸显预测效果, 可利用残差修正方法处理原始序列。

4) 使用对数变换方法优化用水量的原始数据, 以此为基础进行灰色建模, 得到的预测结果平均误差为 0.23%, 精度为 99.75%, 这充分彰显了经优化后的灰色模型具有更理想更可靠的预测效果。

5) 对数变换处理法与弱化算子处理法改进得到的两种灰色模型预测精度都很高, 且两者平均误差相差较小, 但对数变换处理法的理论简单、操作更简便, 在城市用水量预测中值得广泛使用。

基金项目

统计学广西一流学科建设项目资助(桂教科研〔2022〕1号)。

参考文献

- [1] 李琳, 左其亭. 城市用水量预测方法及应用研究[J]. 水资源与水工程学报, 2005, 16(3): 6-10.
- [2] 赖一峰, 高亚静. 基于灰色预测改进模型的配电网可靠性指标预测[J]. 陕西电力, 2012, 40(4): 68-71.
- [3] 冯蕾, 王效华. 江苏省农作物秸秆资源量的灰色预测[J]. 能源研究与利用, 2010(4): 1-3.
- [4] 李又君, 孙培良, 孙青然, 等. 灰色预测模型及干旱预测[J]. 山东气象, 2012, 32(1): 6-11.
- [5] 宋强, 王立新, 李华. 加工中心主轴热误差的灰色残差修正模型研究[J]. 制造技术与机床, 2009(2): 27-30.
- [6] Xiao, X.P., Hu, Y.C. and Guo, H. (2013) Modeling Mechanism and Extension of GM(1,1). *Journal of Systems Engineering and Electronics*, **24**, 445-453. <https://doi.org/10.1109/JSEE.2013.00053>
- [7] Li, Z.B., Jiang, Y., Yue, J., et al. (2012) An Improved Gray Model for Aquaculture Water Quality Prediction. *Intelligent Automation & Soft Computing*, **18**, 557-567. <https://doi.org/10.1080/10798587.2012.10643265>
- [8] 赵宇哲, 武春友. 灰色振荡序列 GM(1,1)模型及在城市用水中的应用[J]. 运筹与管理, 2010, 19(5): 155-159.
- [9] 韩卫红, 李丽. 灰色残差修正模型在中长期负荷预测中的应用[J]. 科技资讯, 2010(7): 124-125.
- [10] 蒋知廷, 张润畦. 水泉矿基于灰色预测模型的矿井涌水量预测[J]. 华北科技学院学报, 2022, 19(1): 7-12.
- [11] 伍德权, 杨延梅, 付宇. 基于灰色系统理论的贵阳市城市用水总量预测[J]. 南昌工程学院学报, 2022, 41(1): 40-46.
- [12] 张鹏程, 张维江, 陆军, 等. 基于灰色系统理论的宁夏万元 GDP 用水量预测[J]. 安徽农业科学, 2011, 39(1): 384-385.
- [13] 陈秀锋, 孟红. 灰色残差修正模型在区域物流量预测中的应用研究[J]. 商业经济研究, 2011(32): 34-35.
- [14] 李秀珍, 孔纪名, 王成华. 灰色 GM(1,1)残差修正模型在滑坡预测中的对比应用[J]. 山地学报, 2007, 25(6): 741-746.
- [15] 杜懿, 麻荣永. 不同改进灰色模型在广西年用水量预测中的应用研究[J]. 水资源与水工程学报, 2017, 28(3): 87-90.