

包头市市区居民生活用水量预测分析

刘治学^{1,2}, 张鑫¹, 王颖华¹

(1. 西北农林科技大学 水利与建筑工程学院, 陕西 杨凌 712100; 2. 包头市供水总公司, 内蒙古 包头 014030)

摘要: 用水量预测对区域水资源规划、利用和管理提供重要依据。运用灰色关联度分析法分析包头市市区居民生活用水量影响因素的基础上, 分别建立多元线性回归模型、灰色 GM(1,1) 模型及灰色线性组合模型对该地区 2009 年和 2010 年的生活用水量进行预测分析, 同时比较了三个模型的预测精度。结果表明: 城市居民生活用水量与城市用水人口、人均居住面积和水价的关联度较高; 2009 年和 2010 年用水量的预测采用组合灰色模型精度最高, 相对误差分别为 13.6% 和 6.5%, 均方根相对误差为 10.7%。组合预测模型的预测精度明显优于单一模型, 使结果更加准确、合理, 符合实际情况。

关键词: 多元线性回归模型; 灰色 GM(1,1) 模型; 组合灰色模型; 灰色关联分析; 城市生活用水预测

中图分类号: TU991.31

文献标识码: A

文章编号: 1672-643X(2012)05-0067-04

Forecast of urban residential domestic water consumption in Baotou

LIU Zhixue^{1,2}, ZHANG Xin¹, WANG Yinghua¹

(1. College of Water Resources and Architecture Engineering, Northwest A & F University, Yangling 712100, China;

2. Baotou City Water Supply Company, Baotou 014030, China)

Abstract: The water demand forecast can supply vital basis for regional water resources plan, utilization and management. On the basis of analysis of the influence factors of residential domestic water consumption in Baotou City by gray correlation analysis, the paper forecasted the regional domestic water consumption in 2009 and 2010 by modeling a multiple linear regression model, gray model (1,1) and gray linear combination model, and compared their prediction accuracy. The results showed that the main factors to regional domestic water consumption were population of urban water, per capita living space and water price. Forecast of water consumption in 2009 and 2010 had highest accuracy by gray linear combination model; relative errors were 13.6% and 6.5% respectively; relative error of root mean square was 10.7%. Prediction accuracy of combination forecast model was better than that of single one and made results more accurate, reasonable and realistic.

Key words: multiple linear regression model; gray model GM(1,1); combination gray model; gray correlation analysis; forecast of urban domestic water consumption

城市生活用水是城市用水的重要组成部分。随着城市化进程速度加快, 城市人口越来越密集, 居民对水量及水质的要求也逐步提高, 必然增加城市供水部门的压力。面对日趋严重的用水形势, 一方面研究居民用水量的影响因素是控制当地生活用水量增长基础; 另一方面, 通过对未来用水量增长趋势的预测, 及时采取相应措施进行合理规划, 对解决供需水矛盾, 维持水资源平衡及利用具有重要现实意义。用水量变化^[1]是一个受到居民收入、水价、社会经

济等多种因素综合作用的随机非线性变化过程, 对不同的用水变化过程具有不同的用水预测方法^[2-4]。针对生活用水的随机波动性和非线性的特点, 本文首先利用灰色关联度分析法分析了包头市市区居民生活用水量的主要影响因素, 然后运用基于多元线性回归模型和灰色预测 GM(1,1) 模型的组合灰色预测模型, 对未来城市生活用水量变化趋势进行预测, 以期通过本文研究, 能够深入了解包头市生活用水变化规律。

收稿日期: 2012-06-06; 修回日期: 2012-07-03

基金项目: 国家“863”计划项目(14110209); “十一五”国家重大科技支撑计划项目(2006BAD11B05); 西北农林科技大学博士科研启动基金资助项目(01140504); 西北农林科技大学科研专项(08080230)

作者简介: 刘治学(1968-), 男, 内蒙古凉城人, 在读硕士, 高级工程师, 从事市政给水方面的技术研究与企业管理工作。

通讯作者: 张鑫(1968-), 男, 河南淅川人, 副教授, 博士, 主要从事水文水资源研究。

1 生活用水量影响因素分析

依据《中国城市统计年鉴》包头市市区 2000 - 2008 年实际统计数据资料,主要分析城市居民生活用水人口(代表城市化指标)、人均居住面积、人均可支配收入(代表城市居民生活水平)、第三产业产值(代表城市服务设施水平)和水价对当地居民生活用水量变化的影响。由于各影响因素的单位不一致,无法进行比较,所以首先需要进行无量纲化,即用各影响因素的各年值除第一年的值,然后对消除单位后的各影响要素绘制时间变化趋势曲线见图 1,从直观上看,城市用水人口、人均居住面积曲线、水价和居民生活用水量曲线在几何形状上较类似。为了进一步揭示居民生活用水量变化趋势的主要影响因素,运用灰色关联度来定量描述生活用水量与影响因素之间的关系。关联度值越大,表明特征序列与相关序列变化趋势越一致,该序列对特征序列的影响也越大。

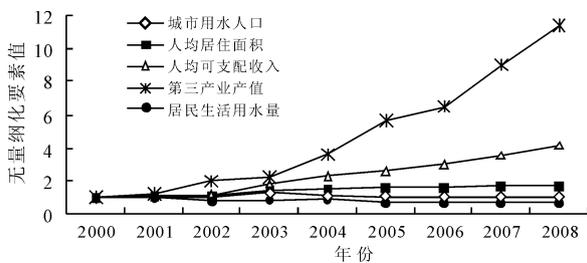


图 1 包头市市区居民生活用水量及其影响因素变化趋势

关联度分析法^[5]是分析灰色系统内所含各因素间相互关系密切程度的一种方法,与传统统计方法相比,它不考虑样本数目和分布特性,且方便计算,不会出现量化结果与定性分析结果不一致的现象,因而具有广泛的适用性。灰色关联度分析法的基本原理具体见参考文献[6]。

根据关联度计算公式求得: $r_{01} = 0.9621$, $r_{02} = 0.9289$, $r_{03} = 0.8204$, $r_{04} = 0.6402$, $r_{05} = 0.9124$ 。 $r_{01} > r_{02} > r_{05} > r_{03} > r_{04}$,说明包头市市区居民生活用水量与城市用水人口、人均居住面积和水价关联度较大,人均可支配收入次之,与第三产业产值关系最小。

2 生活用水量预测分析

2.1 多元线性回归模型应用

2.1.1 模型的建立 多元线性回归法是在预测对象和影响因素之间通过建立回归模型来进行预测,该模型多适用于长期预测。预测模型^[7]如下:

$$\hat{Y} = X\beta \quad (1)$$

其中: $Y = [y_1 \ y_2 \ y_3 \ \dots \ y_n]^T$

$$X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1p} \\ 1 & X_{21} & X_{22} & \dots & X_{2p} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix}$$

$$\beta = [\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_p]$$

式中: Y 为预测对象的历史观测向量; X 为影响因素的历史观测矩阵,是收集到的 p 个 y 的影响因素的 n 次观测值; β 为系数向量。

系数向量的最小二乘估计为:

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

2.1.2 模型精度检验 根据上述方法,利用包头市市区 2000 - 2008 年统计数据,以城市生活用水量为因变量,以城市用水人口、人均居住面积、人均可支配收入、第三产业产值和水价为自变量,借助 SPSS 软件对模型采用进入法回归,建立模型如下:

$$Y = -1749.41 + 77.7X_1 - 83.58X_2 + 0.29X_3 - 0.0004X_4 - 3273.23X_5$$

式中: Y 为包头市市区居民生活用水量,万 m^3 ; X_1 为包头市城市用水人口,万人; X_2 为包头市市区人均居住面积, m^2 ; X_3 为人均可支配收入,元; X_4 为第三产业产值,万元; X_5 为水价,元/吨。

模型的复测定系数 $R^2 = 0.69$,对模型进行 F 检验, $F = 8.65$ 。查 F 分布表得,自由度为(5,5)的临界值 $F_{0.05} = 5.06$,显然 $F > F_{0.05}$,表明该模型回归总体上显著。利用模型对包头市市区 2000 - 2008 年居民生活用水量的模拟结果见表 1。由表 1 可知,模拟误差的平均值为 -1.65%,模型模拟精度较高。

表 1 多元线性回归模型模拟结果 万 m^3 ; %

年份	实测数据	模拟数据	相对误差
2000	3923	3880	-1.10
2001	2874	2820	-1.89
2002	3193	3134	-1.84
2003	3205	3372	5.20
2004	3021	4074	34.85
2005	4742	3886	-18.06
2006	4808	4128	-14.15
2007	3610	2941	-18.54
2008	2841	2860	0.66
平均值	3580	3455	-1.65

2.2 GM(1,1)模型应用

2.2.1 模型的建立 灰色 GM(1,1)模型是基于随

机性的原始时间序列,经按时间累加后所形成的新的时间序列呈现的规律可用一阶线性微分方程的解来逼近的关于数列预测的灰色预测模型。模型进行预测所需原始数据量小,预测精度较高,既可进行宏观长期的预测,亦可用于微观短期的预测,具有广泛地适用性^[8-9]。GM(1,1)模型一般形式^[10]如下:

设有变量 $x^{(0)} = \{x^{(0)}(i), i = 1, 2, \dots, n\}$ 为某一预测对象的非负单调原始数据列,为建立灰色预测模型:首先对 $X^{(0)}$ 进行一次累加(1-AGO, Accumulated Generating Operator)生成一次累加序列:

$$x^{(1)} = \{x^{(1)}(k), k = 1, 2, \dots, n\} \quad (2)$$

式中: $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) = x^{(1)}(k-1) + x^{(0)}(k)$

$z^{(1)}$ 是 $x^{(1)}$ 的紧邻均值生成序列:

$$z^{(1)} = \{z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)\} \quad (3)$$

其中, $z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1); k = 2, 3, \dots, n$ 。

对 $x^{(1)}$ 可建立下述一级白化微分方程:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (4)$$

称为 GM(1,1)模型。式中, a 为发展系数, b 为灰色作用量。

参数向量 a, b 可通过最小二乘法求得:

$$\hat{a} = (B^T B)^{-1} B^T Y \quad (5)$$

式中:

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \dots & \dots \\ -z^{(1)}(n) & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) & 1 \\ x^{(0)}(3) & 1 \\ \dots & \dots \\ x^{(0)}(n) & 1 \end{bmatrix}$$

将计算求得参数 a, b 代入公式(3),取 $x^{(0)}(0) = x^{(0)}(1)$,即得灰色 GM(1,1)模型:

$$\hat{x}^{(0)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a} \quad (6)$$

由式(5)得到的是依次累加值 $x^{(1)}(k+1)$ 的模拟值,再由依次累减得真实的预测值:

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \quad (7)$$

其中: $k = 1, 2, \dots, n$ 。

2.2.2 模型精度检验 依据上述方法,令 $x^{(0)}(0) = x^{(0)}(1) = 3923$,建立 GM(1,1)模型为:

$$\begin{aligned} \hat{x}^{(0)}(k+1) &= \left[x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a} \\ &= 908204.3343e^{0.0037k} - 904281.3343 \end{aligned}$$

模型中的发展系数 $a = -0.003737, b = 3379.65209$,满足 $|a| < 2$,表明该模型符合建模要

求,可以用于预测。利用包头市市区 2000-2008 年居民生活用水量实测数据使用灰色 GM(1,1)模型模拟,结果见表 2。由表 2 可知,模型模拟误差的平均值为 1.11%,比多元线性回归模型的模拟精度提高。

表 2 GM(1,1)模型模拟结果 万 m³, %

年份	实测数据	模拟数据	相对误差	年份	实测数据	模拟数据	相对误差
2000	3923	3923	0.00	2005	4742	3452	-27.21
2001	2874	3400	18.31	2006	4808	3464	-27.94
2002	3193	3413	6.89	2007	3610	3477	-3.67
2003	3205	3426	6.89	2008	2841	3490	22.86
2004	3021	3439	13.82	平均值	3580	3498	1.11

2.3 组合灰色预测模型应用

2.3.1 模型的建立 灰色线性回归模型改善了原线性回归模型中没有指数增长趋势和灰色 GM(1,1)模型中没有线性因素的不足,因此该组合模型更适用于既有线性趋势又有指数增长趋势的序列^[6],以达到更优地预测效果。参考文献^[11]建模过程如下:

设 $X_1^{(0)}(t)$ 和 $X_2^{(0)}(t)$ 是同一现象 $X^{(0)}(t)$ 采用不同预测模型的两种预测值,则称

$$\hat{X}^{(0)}(t) = \rho X_1^{(0)}(t) + (1 - \rho) X_2^{(0)}(t) \quad (8)$$

为两项 GM 组合模型。对 $X_1^{(0)}(t)$ 和 $X_2^{(0)}(t)$ 中至少有一个用 GM 方法确定。

当 $0 < \rho < 1$,可通过下式确定:

$$\rho = \frac{D(e_2)}{D(e_1) + E(e_2)} \quad (9)$$

式中: e_1 为多元线性回归模型的预测值; e_2 为 GM(1,1)模型的预测值; $E(e)$ 为方差; ρ 为 GM 组合模型的系数。

2.3.2 模型精度检验 采用上述方法求得组合灰色模型如下:

$$\hat{X}^{(0)}(t) = 0.07X_1^{(0)}(t) + 0.93X_2^{(0)}(t)$$

用该模型模拟包头市市区居民生活用水量结果见表 3。

表 3 组合 GM 模型模拟结果 万 m³; %

年份	实测数据	模拟数据	相对误差	年份	实测数据	模拟数据	相对误差
2000	3923	3920	-0.08	2005	4742	3482	-26.56
2001	2874	3359	16.89	2006	4808	3510	-26.99
2002	3193	3393	6.28	2007	3610	3439	-4.72
2003	3205	3422	6.78	2008	2841	3446	21.29
2004	3021	3483	15.31	平均值	3580	3495	0.91

由表 3 可知,综合单个模型优点,组合灰色模型模拟相对误差的平均值为 0.91%,并且每一年预测

值的相对误差降低。

2.4 生活用水量预测分析

根据已经建立的多元线性回归模型、灰色 GM(1,1) 模型及组合灰色模型分别对 2009 年和 2010 年包头市市区居民生活用水量进行用水量预测,并采用包头市市区 2009 和 2010 年居民生活用水量实测数据对预测精度进行检验,结果见表 4。

选择相对误差和均方根相对误差作为判断三种

方法的预测误差大小的依据。即:

$$\text{设残差 } e(k) = x^{(0)}(k) - \hat{x}^{(0)}(k), \text{ 则相对误差为:}$$

$$\zeta(k) = [e(k)/x^{(0)}(k)] \times 100\%$$

$$(k = 1, 2, \dots, n) \quad (10)$$

均方根相对误差^[4]为:

$$\xi_{MSE} = \sqrt{\frac{1}{n}(\zeta(k))^2} \times 100\% \quad (11)$$

表 4 三种模型用水量预测精度检验

万 m³, %

年份	用水量			相对误差			
	实测数据	预测数据			多元线性回归模型	GM(1,1)模型	组合灰色模型
		多元线性回归模型	GM(1,1)模型	组合灰色模型			
2009	3023	2521	3504	3435	-16.6	15.9	13.6
2010	3268	2990	3517	3480	-8.5	7.6	6.5
均方根相对误差					13.2	12.5	10.7

3 结 语

(1)近 10 年来,随着包头市节水工作的深入,阶梯式水价制度的实施,节水型用水器具的普及和广大市民节水意识的提升,人均日生活用水量开始出现回落,2000 年至 2010 年城市居民生活用水量由 3923 万 m³ 减少到 3268 万 m³,减少约 17%。同时,城市生活用水量受各种因素综合影响而变化,灰色关联度分析表明城市用水人口、人均居住面积和水价关联度较高,人均可支配收入和第三产业产值关联度较低。

(2)分别建立多元线性回归、灰色 GM(1,1) 模型及组合灰色 GM 三种模型对包头市市区居民生活用水量进行预测分析,其中组合 GM 模型的 2009 年和 2010 年生活用水量的预测相对误差分别为 13.6% 和 6.5%,均方根相对误差为 10.7%,模型模拟精度和预测精度均表明组合灰色 GM(1,1) 模型为最优模型。

(3)城市用水量具有随机非线性特点,受人口、经济收入、产值及气候等诸多因素的影响,在进行预测模型的选择时要综合考虑上述因素,改进单一模型的不足,提高预测精度,使结果更加符合实际情况。

参考文献:

- [1] 崔惠珊,邓逸群.居民用水量的影响因素研究评述[J].水资源保护,2009,25(1):83-85.
- [2] 张雄,党志良,张贤洪,等.城市用水量预测模型综合研究[J].水资源与水工程学报,2005,16(4):24-28.
- [3] 李琳,左其亭.城市用水量预测方法及应用比较研究[J].水资源与水工程学报,2005,16(3):6-10.
- [4] 张宏伟,岳琳,王亮.基于径向基函数的城市日用水量预测方法[J].天津大学学报,2006,39(4):86-89.
- [5] 拜存有,赵翔翔,冯旭,等.灌溉用水量变化趋势的灰色关联分析[J].人民黄河,2004,26(7):33-34.
- [6] 刘思峰.灰色系统与建模[M].北京:科学出版社,2004.
- [7] 杨崇豪,张川云,吴文学.郑州市未来 15 年城市需水量预测研究[J].人民黄河,2006,28(11):30-32.
- [8] 景亚平,张鑫,罗艳.城市需水量组合预测[J].西北农林科技大学学报(自然科学版),2011,39(7):229-234.
- [9] 邓丽娟,魏光辉.基于灰色模型的城市用水量预测[J].地下水,2011,33(1):100-101.
- [10] 梁学玉,张鑫,孙天青.组合灰色预测模型在城市用水量预测中的应用[J].人民黄河,2010,32(4):79-80.
- [11] 王涛,张宏伟,牛志广.GM 组合模型用于城市生活用水量预测[J].水资源保护,2007,23(5):28-30.