

多元线性回归分析在北京城市生活需水量预测中的应用^{*}

张雅君 刘全胜 冯萃敏

提要 从多元线性回归分析的特点出发,探讨北京城市生活需水量的影响因素,并对选定的影响因素进行回归分析,确定了最终的预测方程,应用预测方程对北京市2010年城市生活需水量进行了预测。

关键词 生活需水量 多元线性回归 预测 北京市

0 前言

随着城市对水资源的需求日益提高,对城市需水量进行科学的分析和预测是满足城市需求、保证水资源可持续利用的前提。科学的供水规划源于对未来的科学分析与准确预测,需水量的长期预测即着眼于这一问题。在众多的城市需水量预测方法中,结构分析法是需水量长期预测的有效方法。结构分析法从研究客观事物与影响因素的关系入手,分析影响预测对象的各种主要因素,建立预测对象与影响因素之间的简单关系模型,通过研究影响因素的变化规律间接反映出预测对象的变化规律。

回归分析是结构分析法的主要分析方法,分为一元线性回归分析、多元线性回归分析和非线性回归分析。本文将采用多元线性回归分析的方法对2010年北京城市生活需水量(居住用水和公共生活用水)进行预测。

1 多元线性回归分析

1.1 一元线性回归分析

预测模型为: $\hat{y} = \beta_0 + \beta_1 x$ (1)

$$L_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2$$

$$L_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}$$

$$L_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2$$

则参数 β_0, β_1 的最小二乘估计 β_0, β_1 为:

$$\begin{cases} \beta_1 = L_{xy} / L_{xx} \\ \beta_0 = \bar{y} - \beta_1 \bar{x} \end{cases} \quad (2)$$

回归模型的显著性检验采用相关系数 R 来判断:

$$R = \frac{L_{xy}}{\sqrt{L_{xx}} \sqrt{L_{yy}}} \quad (3)$$

在给定可靠度 α 下,当 $|R| > R_{1-\alpha}$ 时,认为回归方程显著。

1.2 多元线性回归分析

预测模型为: $Y = X\beta$ (4)

其中 $Y = (y_1, y_2, \dots, y_n)^T$

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

$$\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_p)^T$$

式中 Y ——预测对象的历史观测向量;

X ——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 ——系数向量。

系数向量的最小二乘估计为:

$$\beta = (X^T X)^{-1} X^T Y \quad (5)$$

方程的显著性检验由两部分组成:一部分是对方程拟和误差的检验,用复相关系数 R 完成;另一部分是对方程回归系数的检验,用统计量 t 来完成^[1]。

2 北京生活需水量的多元回归分析

2.1 生活需水量影响因素即回归模型自变量选择

生活用水系统复杂,影响因素多种多样,自变量的选择从全面性、重点性、可量化及可控制的原则出发,尽可能包含有效变量、不包含无效变量。其主要

*北京市教委科技发展计划资助项目(99kj-45)。

影响因素包括城市非农业人口、第三产业状况、建筑面积和生活水平,此外,绿地和道路用水与当年绿地面积和当年降水量有关。由于本次预测是对城市生活需水量的预测,所以选用城市非农业人口作为影响因素来表示人口的因素;而第三产业产值和建筑面积是公共用水的主要影响因素,对于北京这种公共用水占生活用水量比重很大的城市,这两个因素应重点考虑。所以选择了第三产业产值和居住面积(建筑面积的数据难以获得,此处以居住面积代之)作为影响因素;人均居住面积在一定程度上反映了人民生活水平,一方面,人均居住面积的提高是人民生活水平改善的最直接的反映,与之伴随的是生活设施的改善,尤其是卫生设施的改善;另一方面,人均居住面积的增加也是人民经济水平上升的标志,相应地人们的第三产业用水消费也将增加,所以人均居住面积的增加必将导致人均用水定额的上升。因此,北京城市生活用水量的预测主要选择了非农业人口、第三产业产值、居住面积、人均居住面积及绿地因子(当年绿地面积与当年降水量之比)5个影响因素进行多元线性回归分析。

2.2 选定自变量的多元线性回归分析

对出自《北京市统计年鉴》的1986~1999的生活用水量相关数据归纳见表1^[3]。

分别选城市生活需水量作响应变量 Y ,非农业人口数 X_1 ,第三产业产值 X_2 ,居住面积 X_3 ,人均居

表1 生活用水回归分析数据

年份	生活用水量 Y / 万 m^3	非农业人口 X_1 / 万人	第三产业产值 X_2 / 万元	居住面积 X_3 / 万 m^2	人均居住面积 X_4 / m^2 /人	绿地因子 X_5
1986	46 474.9	621	99.97	3 749.4	6.46	5.663 61
1987	45 744.1	637	119.92	4 054.8	6.82	5.467 174
1988	50 904.4	650	151.88	4 361.4	7.17	6.050 795
1989	51 428.7	664	165.2	4 653	7.45	9.384 894
1990	52 852.8	673	194.55	4 916.2	7.72	6.037 573
1991	56 619.5	683	261.84	5 165.5	8.01	5.721 888
1992	58 376.6	692	314.52	5 424	8.31	7.780 24
1993	62 571.9	707	395.17	5 661.9	8.51	8.786 461
1994	67 295.2	725	509.42	5 963.2	8.73	6.418 193
1995	70 540.5	815	698.28	6 317	8.87	8.762 445
1996	69 860.5	829	849.13	6 661.1	9.17	7.343 986
1997	75 573.5	826	986.68	7 021.3	9.49	12.550 48
1998	78 731.5	838	1 137.9	7 428.5	9.87	8.679 787
1999	82 715.5	854	1 246.75	8 046.9	10.43	24.192 58

住面积 X_4 ,绿地因子 X_5 作自变量进行完全多元回归分析。以每个自变量可建立5个一元回归方程,以任意2个自变量可建立10个二元回归方程,以任意3个自变量可建立10个三元回归方程,以任意4个自变量可建立5个四元回归方程,以5个自变量可建立1个五元回归方程,共得出31个多元回归方程(由于篇幅所限,不能一一列出)。

2.3 回归方程的分析及最终预测方程的确定

选择回归方程须遵循以下标准进行。

(1) R^* 值要足够大, R^* 反映了回归拟合的好坏,是历史数据与回归方程符合的程度,如果 R^* 太小,以至于不能通过方程显著性检验,则不能给方程以足够的“信任”。

(2)回归方程的各系数都须通过 t 检验。经计算,在95%的置信度的条件下,通过回归方程显著性检验的 R^* 值和 t 值见表2。

表2 生活用水量回归方程 R^* 和 t 检验

自变量数	1	2	3	4	5
R^*	0.286	0.420	0.527	0.617	0.698
t	2.178 8	2.201 0	2.228 1	2.262 2	2.306 0

(3)在用水系统中,数据的相关性并不能完全反逻辑的相关性,对于同时满足表2中数据的回归方程,还须考察其自变量是否在逻辑上与回归变量的关系更加密切。

经过比较31个生活需水量的回归方程,选择多元回归方程为:

城市生活需水量预测方程

$$Q_{\text{生活}} = 47.566\,299\,64X_1 + 6\,992.364\,413X_4 - 31\,004.224\,60 \quad (6)$$

式中 $Q_{\text{生活}}$ ——北京城市生活需水量,万 m^3 ;

X_1 ——全市非农业人口,万人;

X_4 ——人均居住面积, m^2 /人。

从逻辑上看,生活用水量的增加主要由两部分引起:一方面由于用水人口的增加导致用水规模的增加,引起用水量的增加;另一方面,由于人民生活水平的提高,导致人均用水量的增加,从而导致总用水量的增加,所以,在回归方程中能反映此两方面的因素是比较合理的。上述方程中非农业人口反映了用水人口的状况,而人均居住面积反映了人民居住

水平及生活水平。

3 北京市 2010 年城市生活需水量的多元回归预测

3.1 非农业人口的预测

引起一个地区人口变化的因素主要有两个,即人口的自然增长和机械增长。对于相对封闭的地区,机械增长可以近似看作零,人口的增长主要由自然增长决定;而对于流动人口较大的地区,人口的机械变动不能忽略,甚至起着决定的作用。北京是人口高度流动的城市,人口的自然增长已经接近于零,北京市人口自然变动情况见图 1。

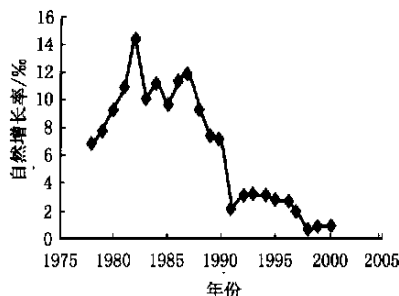


图 1 人口的自然变动

由图 1 可以看出,北京的人口增长率已经低于千分之一,随着计划生育政策的继续实行,人口老龄化的进一步发展,人口的自然增长率将继续下降,甚至出现负值,然而这与北京市城市非农业人口的变化情况是不一致的,北京市城市非农业人口的变化情况见图 2。

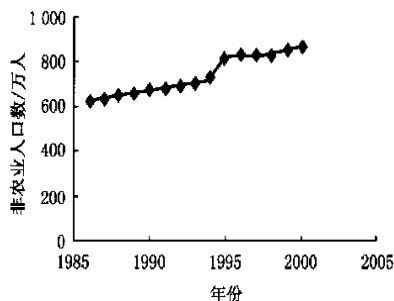


图 2 非农业人口曲线

由图 2 可以看出,北京市非农业人口的实际增长速度维持在 2% 左右,这主要是由于两方面的原因造成的:

- (1) 北京吸引人才政策的落实及首都的吸引力,外地人口进京造成较大的机械增长;
- (2) 随着城市化进程的加快,原来的农业人口

转化为城市人口。

进入 1994 年以后,北京市人口的增长明显进入了具有减速增长特征的对数增长期,增长速度有下降的趋势,在今后 10 年内,随着自然增长率的继续下降,城市非农业人口将继续呈现对数增长趋势,对 1994 年以后的数据进行对数曲线拟合,见图 3 所示。

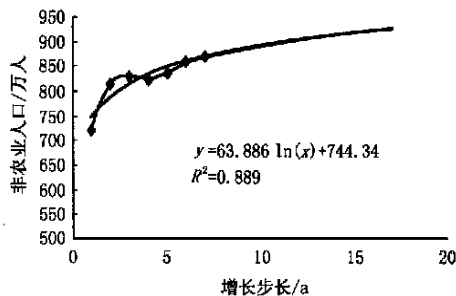


图 3 非农业人口拟合(1994 年为起始年)

3.2 人均居住面积的预测

人均居住面积的变化是人民经济水平提高的直接结果,在过去的 16 年里,人均居住面积基本上呈线性增长,在今后一个相当长的时间内,这种趋势不会有大的改变,这主要是由于经济的发展,住房改革的深入,使更多的人购买商品住房,改善住房水平;一些先富起来的家庭已经开始了住房消费由安居型向舒适型转变,这也将促进人均居住面积的提高。对人均居住面积曲线进行线性拟合,见图 4 所示。

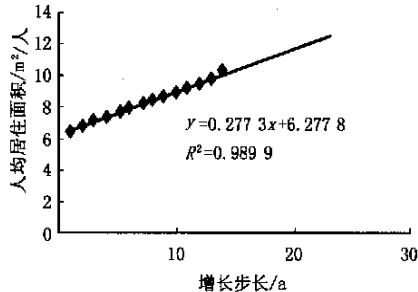


图 4 人均居住面积拟合(1986 年为起始年)

3.3 生活需水量预测结果

根据拟合关系式,预测北京市今后 10 年的非农业人口和人均居住面积,并根据式(6)预测北京市未来 10 年的生活需水量预测结果见表 3 所示。

通过预测,北京市 2005 年城市生活需水量约为 9.5 亿 m^3 , 2010 年约为 10.5 亿 m^3 。

城市给排水工程规划水量规模的确定

吴兆申 皇甫佳群 金家明

提要 确定城市给水排水工程规划水量规模十分重要, 提出从总体规划、专业规划、详细规划到工程实施等各规划阶段预测水量规模的依据及需注意的问题。

关键词 给水 排水 规划水量规模 预测 依据

城市给水排水工程从各规划阶段到具体项目实施, 确定其水量规模是首要内容, 规模预测是否符合发展趋势和实际需要, 将对水资源的合理利用、工程总体布局、实施步骤和工程费用产生重大影响。因而合理确定水量规模十分重要。

城市给水排水工程从总体规划、专业规划、详细规划阶段, 到工程实施, 其水量规模的确定是逐步深化和完善的过程, 各阶段有不同的规范、标准、指标作指导。

1 总体规划阶段给水水量预测

城市总体规划阶段的给水工程规划是根据城市发展目标、用地、人口规模, 空间布局安排和水资源状况, 提出各取水水源、供水系统的规划期内工程水量、水质目标和设施布局。

国家标准《城市给水工程规划规范》(GB50282—98)(以下简称“给水规范”)是测算城市总用水量规模的主要依据。

(1)“给水规范”所提指标适用于城市总体规划期内(一般为20年)的水量预测, 并按此控制水资源和提出总水量规模。由于城市用水有逐步增长的过程, 因而近期指标要大幅下降。

(2)“给水范围”所提指标是全国通用指标, 选用时不能简单按照城市规模类别和分区进行套用, 必须先对城市现状指标进行测算研究, 按照发展趋势确定规划期所采用的指标。由于编制“给水规范”所提指标是依据1991~1994年统计资料, 该年段正处于用水高速增长期, 并按照逐年增长的概念来测算。近年来由于水资源紧缺, 节水措施的加强, 高耗水工业的更新换代和工厂外迁等因素, 城市供水量增长缓慢, 有些城市还有所下降, 使“给水规范”所提指标偏大。如南方某市1995年单位人口综合用水量指标为 $0.723 \text{ 万 m}^3/(\text{万人} \cdot \text{d})$, 至2000年下降至 $0.577 \text{ 万 m}^3/(\text{万人} \cdot \text{d})$, 2001年为 $0.603 \text{ 万 m}^3/(\text{万人} \cdot \text{d})$ 。在编制总体规划时, 确定近期(2005年) $0.65 \text{ 万 m}^3/(\text{万人} \cdot \text{d})$, 远期(2020年) $0.75 \text{ 万 m}^3/(\text{万人} \cdot \text{d})$, 而“给水规范”建议指标为 $0.8 \sim 1.2 \text{ 万 m}^3/(\text{万人} \cdot \text{d})$ 。同一城市的不同地区, 由于用地性质和供水条件不同, 应采用不同的指标。一些水资源不足的城市和供水距离较远的地区, 更应强调节约用水, 采用多种措施降低耗水量, 其综合用水量指标也应大幅下降。

(3)“给水规范”所指“人均”是指户籍人口, 未包

表3 生活需水量预测

年份	城市非农业人口 /万人	人均居住面积 /m ² /人	生活需水量 /万 m ³
2001	877.2	10.71	85 609.2
2002	884.7	10.99	87 923.8
2003	891.4	11.27	90 200.3
2004	897.5	11.55	92 448.3
2005	903.1	11.82	94 602.6
2006	908.2	12.10	96 803.1
2007	912.9	12.38	98 984.5
2008	917.3	12.66	101 151.7
2009	921.5	12.93	103 239.4
2010	925.3	13.21	105 378.0

参考文献

- 1 吴翊. 等. 应用数据统计. 北京: 国防科技大学出版社, 1995
- 2 北京市统计局. 北京统计年鉴. 北京: 中国统计出版社, 1985—2001

○作者通讯处: 100044 北京建筑工程学院城建系

电话: (010)68322115

E-mail: zhangyajun@sohu.com

修回日期: 2003-1-24