

时间序列特性驱动的 城市供水量预测方法及应用研究

Research and Application on Water Supplies Forecasting Based on
Feature-Driven of the Time Series

白云 王圃 李川 著



中国财经出版传媒集团



经济科学出版社
Economic Science Press

时间序列特性驱动的 城市供水量预测方法及应用研究

Research and Application on Water Supplies Forecasting Based on
Feature-Driven of the Time Series



王圃 李川 著

中国财经出版传媒集团



经济科学出版社
Economic Science Press

图书在版编目 (CIP) 数据

时间序列特性驱动的城市供水量预测方法及应用研究/
白云, 王圃, 李川著. —北京: 经济科学出版社, 2018. 4
ISBN 978 - 7 - 5141 - 9300 - 8

I. ①时… II. ①白…②王…③李… III. ①城市供水 -
预测 - 研究 IV. ①TU991

中国版本图书馆 CIP 数据核字 (2018) 第 093106 号

责任编辑: 李 雪

责任校对: 杨 海

责任印制: 邱 天

时间序列特性驱动的城市供水量预测方法及应用研究

白云 王圃 李川 著

经济科学出版社出版、发行 新华书店经销

社址: 北京市海淀区阜成路甲 28 号 邮编: 100142

总编部电话: 010 - 88191217 发行部电话: 010 - 88191522

网址: [www. esp. com. cn](http://www.esp.com.cn)

电子邮件: [esp@ esp. com. cn](mailto:esp@esp.com.cn)

天猫网店: 经济科学出版社旗舰店

网址: [http: //jjkxcbs. tmall. com](http://jjkxcbs.tmall.com)

固安华明印业有限公司印装

710 × 1000 16 开 14.25 印张 200000 字

2018 年 5 月第 1 版 2018 年 5 月第 1 次印刷

ISBN 978 - 7 - 5141 - 9300 - 8 定价: 50.00 元

(图书出现印装问题, 本社负责调换。电话: 010 - 88191510)

(版权所有 侵权必究 举报电话: 010 - 88191586)

电子邮箱: [dbts@ esp. com. cn](mailto:dbts@esp.com.cn))

前 言

随着城市化进程的推进和社会经济的快速发展,人民生活水平不断提高,城市自来水公司的供水系统的规模迅速扩大,从而导致了企业的供水系统调度复杂性逐年提高,这一现象给传统调度方式带来了前所未有的挑战。而供水量预测作为企业供水系统调度工作的基础和前提,一直是城市供水企业和运行管理部门最为棘手的问题之一。通过供水量预测,既可以为供水系统科学调度提供数据依据,又可以提高水资源利用效率,改善城市生态环境,促进社会和谐健康发展。

日供水量的预测可以保证用户在不同时间对水量和水压的要求,同时也能提高水厂的生产效率,减少生产成本,从而提高供水服务质量。月供水量的预测可以平衡水源与各水厂的供给量,提高区域调度能力,减少水资源的浪费。因此,笔者从日、月供水量预测两个方面出发,结合实例深入分析供水量时间序列特性,利用混沌、多尺度分析、信息融合、模型预测控制、智能预测等理论和方法,构造不同时间序列特性驱动的城市供水量预测模型。

全书共分8章:第1章,在对国内外研究现状和文献分析的基础上提出本书研究框架;第2章,按照建模的一般步骤概述了本书预测建模所涉及的关键理论和技术;第3章,基于混沌相空间重构理论定性和定量分析供水量时间序列的混沌特性,从而证明其可预测性;第4章,介绍并比较几种常用供水量时间序列预测模型,并利用 Matlab

数值分析软件实现其实例建模；第5章，鉴于日供水量时间序列的局部特性差异性，提出一种多尺度二乘支持向量回归的预测模型；第6章，鉴于日供水量时间序列的时变动态性，提出一种变结构最小二乘支持向量回归的动态预测模型；第7章，鉴于月供水量时间序列的趋势、周期和随机干扰耦合特性，提出一种特性加法模型；第8章，总结。

本书可供系统工程、管理科学与工程、市政工程、环境工程等领域的科研人员、规划设计人员、工程技术人员阅读，也可对进行时间序列分析、预测建模及应用研究的有关学者、高校师生提供参考。

由于笔者水平有限，文中有些观点的归纳和阐述难免有疏漏和不足的地方，还望各位读者海涵并提出宝贵意见。

白云

2018年3月

目 录

第1章 绪论 / 1

- 1.1 研究背景及意义 / 1
- 1.2 城市供水量预测方法研究现状 / 5
- 1.3 供水量序列特性简述 / 20
- 1.4 研究内容及章节安排 / 23

第2章 预测建模相关理论及技术概述 / 27

- 2.1 数据准备及预处理技术 / 28
- 2.2 数据信息特征学习理论 / 34
- 2.3 模型选择问题 / 46
- 2.4 性能评估指标 / 51

第3章 供水量时间序列的可预测性分析 / 52

- 3.1 相空间重构 / 53
- 3.2 供水量时间序列混沌分析 / 71
- 3.3 本章小结 / 80

第4章 常用供水量预测模型研究 / 81

- 4.1 整合自回归移动平均模型 / 81
- 4.2 误差反向传播神经网络模型 / 97
- 4.3 自适应模糊神经网络模型 / 108
- 4.4 最小二乘支持向量回归模型 / 118
- 4.5 模型预测比较 / 124
- 4.6 本章小结 / 130

第5章 日供水量多尺度最小二乘支持向量回归预测模型研究 / 131

- 5.1 局部建模 / 132
- 5.2 日供水量变化规律分析 / 133
- 5.3 多尺度分析 / 136
- 5.4 多尺度最小二乘支持向量回归预测模型 / 143
- 5.5 实例分析 / 145
- 5.6 本章小结 / 149

第6章 日供水量变结构最小二乘支持向量回归预测模型研究 / 150

- 6.1 动态建模 / 152
- 6.2 最大预测时间 / 153
- 6.3 变结构最小二乘支持向量回归预测模型 / 155
- 6.4 实例分析 / 162
- 6.5 本章小结 / 167

第7章 月供水量加法预测模型研究 / 168

- 7.1 混合建模 / 169

7.2 月供水量变化规律分析 / 170

7.3 加法预测模型 / 172

7.4 实例分析 / 178

7.5 本章小结 / 193

第8章 总结 / 194

8.1 结论 / 194

8.2 主要创新点 / 196

参考文献 / 198

第 1 章

绪 论

1.1 研究背景及意义

城市供水系统是城市建设的重要组成部分之一，完善的供水系统对促进城市工农业生产、保障人民身体健康，以及保护环境免遭污染等都具有积极作用。近年来，随着城市化进程的推进和社会经济的快速发展，居民生活水平的提高，城市用水量大大增加。根据中国统计年鉴^[1]（见表 1-1），全国人均用水量呈上升趋势（2014 ~ 2016 年人均用水量有所下降，但仍然比“十一五”时期高）。

表 1-1 2000 ~ 2016 年全国人均用水量统计

序号	年份	人均用水量（立方米/人）	序号	年份	人均用水量（立方米/人）
1	2000	435.4	4	2003	421.9
2	2001	437.7	5	2004	428.0
3	2002	429.3	6	2005	432.1

续表

序号	年份	人均用水量 (立方米/人)	序号	年份	人均用水量 (立方米/人)
7	2006	442.0	13	2012	454.7
8	2007	441.5	14	2013	455.5
9	2008	446.2	15	2014	446.7
10	2009	448.0	16	2015	445.1
11	2010	450.2	17	2016	438.1
12	2011	454.4			

资料来源：中国统计年鉴 <http://www.stats.gov.cn/tjsj/ndsj/>。

为了满足日益增加的生产、生活用水需求，城市供水系统的范围及规模在逐年扩大。根据住房城乡建设部会同国家发改委编制的《全国城镇供水设施改造与建设“十二五”规划及2020年远景目标》^[2]，“十一五”期间全国供水能力增加0.33亿立方米/日，管网长度增加22.21万公里，用水人口增加0.96亿人。截至2010年年底，全国城镇供水能力总计达到3.87亿立方米/日，用水人口6.30亿人，管网长度103.55万公里，年供水总量714亿立方米，与2000年相比，全国新增供水能力0.68亿立方米/日，增长26.67%，新增用水人口2.30亿人，增长85.50%。根据住房城乡建设部会同发改委编制的《全国城市市政基础设施规划建设“十三五”规划》^[3]，“十二五”期间设市城市（县城）供水能力达到2.3101亿立方米/日（0.4675亿立方米/日），与2010年相比，增幅达15%（21%），公共供水普及率分别达到93.1%（85.1%）。东部、中部、西部地区市政基础设施水平差异逐渐缩小，市政基础设施公共服务城乡统筹、区域共建共享有序推进。

2016年是“十三五”规划的开局之年，是全面落实中央城市工作会议的第一年。根据2016年城乡建设统计公报^[4]，截至2016年年

底,城市(县城)供水综合生产能力达到3.03亿立方米/日(0.54亿立方米/日),比上年增长2.2%(减少6.0%),其中公共供水能力2.39亿立方米/日(0.46亿立方米/日),比上年增长3.4%(减少3.3%)。供水管道长度75.7万公里(21.1万公里),比上年增长6.5%(减少1.6%)。用水普及率98.42%(90.5%),比上年增加0.35个百分点(0.54个百分点)。

由于城市供水系统范围及规模的逐年扩大,供水复杂性逐年提升,从而导致与此相关的城市供水系统的调度筛选方案及调度复杂性提高。而当前多数中小城市供水调度决策依旧停留在传统经验的基础上,即根据日常经验判断或根据流量表与压力表数据判断。前者盲目依靠经验判断,后者缺乏预知性(滞后性),容易造成:①供水富余,消耗大量电能;②供水过度造成管网压力偏高,增加水资源漏损或水管爆裂风险的概率;③供水不足,造成供需失衡的不良后果。《全国城市市政基础设施规划建设“十三五”规划》^[3]中也明确指出了“十三五”时期所面临和解决的主要问题之一,即现有供水企业缺乏专业化、规范化、规模化的建设和运营管理,同时供水设施监管信息化水平普遍偏低,智慧化管理仍有较大差距,导致供水设施运行效率、服务质量较低。

所以,传统供水系统管理方式面临着巨大的挑战,对城市供水系统的调度进行优化是必要的。而城市企业供水系统的优化调度主要包括三个环节:用水量预测、运行工况模拟和供水系统调度决策。在这三个环节中,用水量的预测工作是后两个环节的基础和前提,其准确度直接影响供水系统工况模拟结果的合理性,以及调度决策模型的针对性和可靠性,所以预测环节的研究显得尤为重要。由此可见,供水量的预测既能为水资源的规划和管理提供数据基础,同时是供水企业进行优化调度的重要组成部分,并有着十分重要的意义。

①供水系统的短期供水量预测可以保证各用水单元在不同时段对水量和水压的不同要求,保证了正常用水量的需求,提高了供水企业管理水平和供水服务质量。

②供水量预测是供水系统优化调度的第一环节,为供水系统的优化调度提供数据支持。自来水厂出水经泵站加压后送至各用水单元,这一过程需要电能消耗(供水企业主要经济支出)。通过供水量预测,可以指导泵站的优化调度工作、提高供水系统储备能力的利用率、降低生产能耗,从而可以实现供水系统在确保安全、稳定、优质的供水过程中节省能源开支。吕谋等^[5]研究表明,利用时用水量的预测结果优化了的调度系统运行费用比常规经验调度方式节约了3%~5%的能源开支(以电费为评价指标)。翟光日^[6]通过开展用水量预测、宏观水力模型建立、调度模型的建立和求解,得到24小时的调度压力,普遍比实际的调度压力低了1米左右,具有实际的节能意义。巴克尔等^[7]、贾科梅勒等^[8]均通过用水量预测优化供水泵组,成功削减了调度运行费用。而供水电耗的减少不仅提高了供水企业的经济效益,也缓解了国家能源紧张的局面。

③通过供水量预测,可以合理分配不同区域的输送水量,为各个供水厂配送水量提供依据,最大限度地降低了供水调度成本。

④由于生活用水对水质要求较高,不能过久储存,所以要求供水企业的生产、输送、分配、调度和用户用水同时进行。因此,通过供水量预测可以实现产水、输水、用水的供需平衡,降低水质安全风险。

⑤城市供水系统的中长期供水量预测可以为市政供水基础设施的规划、投资、建设及改造,水资源评估及利用等方面提供科学的数据依据,从而避免了资源的浪费。

因此,城市供水企业和运行管理部门对供水量的预测尤为关注,

特别像重庆市如此严重的缺水城市（全国 400 个缺水城市之一，属于中度缺水地区，其中 12 个区县属于重度缺水地区）。另外，我国供水设施专业管理水平偏低，缺乏有效监管，水资源利用方式粗放。所以，供水量的预测一方面可以为相关部门水资源管理提供技术支持，促进管理和决策的合理化，提高水资源利用的效率；另一方面可以对该领域预测方法进行有效的探索和检验，并对进一步研究起到一定的推动作用。

1.2 城市供水量预测方法研究现状

在我国供水事业迅猛发展的今天，衡量一个供水企业的管理是否走向现代化、自动化的显著标志之一就是城市供水量预测的水平。与此相关的理论研究一直没有中断过，方法也很多。一般的，预测工作是通过数据库中的历史供水量数据进行分析，得到数据背后的信息量（这些数据的整体特征及其发展趋势），从而建立预测模型。但是由于此类数据没有明显的模式化结构，使得用传统统计学模型或方法来分析和预测未来时段供水量或者不经济（数据收集工作量大、大数据运算负荷大、有效数据少等），或者在某些情况下不能获得有实际意义的模型（统计学模型是建立在数据序列具有特定性质假设的基础上），从而出现了数据丰富、信息短缺的现象。因此，此类问题亟须功能强大的数据分析处理技术来解决。通过利用各种先进分析方法从大量数据中发现数据序列内在特性，包括趋势、周期、季节性影响、随机等各种特性，从而较为准确的建立预测模式和数据间的关系。学者们通过长期对数据库技术的研究和开发，在对历史数据进行查询和处理技术提高的同时，能够发现历史数据序列之间的内在联系，从而

对未来信息进行有效预测。本章节从传统预测方法和基于新技术的方法两个方面来对研究现状进行描述和讨论。

1.2.1 传统预测方法

传统预测方法是基于统计学理论发展而来的,主要建模方法有回归分析法、指数平滑法、趋势外推法、移动平均法等。此类方法使用时要对数据序列性质进行假设,若假设条件合理,得出的结果较为理想;反之,则预测模型将会严重失真。尤其是,指数平滑法和趋势外推法更加依赖于数据特性,故使用范围较小。本书仅对使用较为广泛、效果良好的回归分析法、趋势外推法和混合自回归滑动平均模型进行介绍。

①回归分析法。回归分析法是在调研大量历史数据的基础上,利用数理统计方法建立因变量与自变量之间的回归方程(回归关系函数表达式)。在回归分析中,当研究的因果关系只涉及一个自变量时,称为一元回归分析;当研究的因果关系涉及两个或两个以上自变量时,称为多元回归分析。根据回归方程的表达形式,回归分析可以分为线性回归和非线性回归,其中非线性回归的一般解决方式是通过变量变换,将非线性回归转变为线性回归,然后对因果关系进行线性回归分析。而线性回归方法是按最小二乘法原理来求出回归系数值,从而得到预测模型^[9]。

张雅君等^[10]利用多元线性回归分析方法对城市生活需水量的影响因素(非农业人口数量、第三产业产值、人均居住面积和绿地因子)进行回归分析,确定了最终的预测方程(非农业人口数量和人均居住面积为主要因素),并对北京市2010年城市生活需水量进行了预测。龙德江^[11]选用总人口数、固定资产、工业单位数量、国内生产总

值、人均国民生产总值、人均日生活需水量和供水能力7个因素做主成分分析,最后建立总人口数、国内生产总值和供水量之间的多元线性回归模型。鞠佳伟等^[12]构造了以温度、天气和假日为变量的多元线性回归预测模型,通过测试发现重大节日对供水量预测误差影响较大,其中春节前后预测结果误差最大,对模型进行优化后预测误差显著降低。

亚萨尔等^[13]建立了月平均水费、总人口数、大气温度、相对湿度、降雨量、全球太阳辐射、日照时间、风速、气压与供水量的多元非线性回归模型,通过多步预测,得出最符合土耳其亚达纳城的供水量预测模型。梅斯等^[14]建立了中长期用水量与水价、人口、居民人均收入、年降雨量4个相关因子间的对数和半对数回归模型,并成功应用于美国得克萨斯州中长期用水量的预测。布雷克等^[15]采用逐步回归法进行城市供水量预测,缩短了多元回归模式中用于常规趋势分析和单位水量需求分析的时间,减少了统计人员的工作量,提高了建模效率和精度。

本方法计算方便,适用于在系统没有发生重大变化、波动较小的年用水量的预测^[16],同时要求有更多类型的原始建模数据。对于数据波动较大的供水量预测问题,由于影响因素复杂,且影响因素未来值的准确预测较难,不宜采用该方法。由于该方法的预测精度直接取决于影响因素的选取、原始建模数据的准确性和选用模型的合理性,因此对上述问题在没有充分把握的情况下要慎重使用。根据回归理论,该方法主要用于内推预测,外推预测的精度无法保证。

②趋势外推法。趋势外推法是基于假设“未来趋势是过去和现在连续发展的结果”。换句话说,预测对象随时间推移呈现某种简单的渐进式变化趋势,没有明显的跳跃式发展规律,且能找到一个合适的函数或数学表达式来反映这种变化趋势,就可以利用趋势外推法来实

现预测。一般的,趋势外推法主要包括四个环节:选择预测参数、拟合曲线、趋势外推、预测验证。

周申蓓等^[17]提出一种计量抽样趋势外推法对工业用水量进行测算,相对误差控制在10%左右,对于技术和统计受限的行业用水量预测是一种比较理想的预测模型。魏光辉^[18]采用趋势外推法和用水定额法相结合的方式对研究区规划水平年各行业需水量进行预测。李稳等^[19]构造播种面积、产量与农业需水量变化的趋势面模型,具体分析了不同阶次趋势面模型的拟合程度,选择最佳阶次模型预测未来农业需水量。

此类方法用于水量预测在国外研究较少,主要是研究者考虑到其假设的局限性。除农业、工业等特定行业外(水量与生产过程参数有紧密关联性),水量时间序列具有一定的季节性、周期性等规律,无法找到合适的函数对其进行表征和拟合(常用函数表达式为线性模型、指数曲线、生长曲线、包络曲线等)。所以,趋势外推法的适用范围有限,对于复杂变化特性的时间序列无法得到满意的预测结果。

③混合自回归滑动平均模型。混合自回归滑动平均模型(auto-regressive integrated moving average, ARIMA)是自回归和滑动平均模型的综合,通过对时间序列的平稳性分析和模式识别,分别确定模型形式(有三种预测模式,AR、MA和ARMA)。ARIMA模型的基本思想是用数学模型(此模型由自回归项 p 、移动平均项 q 和差分次数 d 来确定)来近似描述一个随机序列(此序列是由预测对象随时间推移而形成的)。自从博克斯和詹金斯^[20]于1976年提出ARIMA后,该模型被广泛地运用于供水量预测。

赵凌等^[21]分解并剔除原始时间序列的长期趋势及季节变化后,构造其残差序列并进行识别,建立了城市月供水量的ARIMA模型,此模型成功用于成都市月供水量预测。练庭宏等^[22]利用自相关分析、

偏自相关分析等参数辨识模型阶次结构 (p, q, d) ，用来预测未来城市需水量趋势。吉乔伟等^[23]综合考虑时用水量的时变性、周期性、非线性、非平稳等特性，引入季节因子，改进 ARMA 模型，用于模拟时用水量趋势，结果表明改进 ARMA 季节模型预测精度较高且预测误差相对稳定。

列昂尼德等^[24]通过 ARMA 方法分解，实现对时供水量的模式识别，从而预测时供水量的趋势变化。蒙柏尼等^[25]将季节因素与 ARIMA 模型相结合，用于城市需水量预测。亚尔琴塔什等^[26]利用 ARIMA 模型构造了 2015 ~ 2018 年期间的时间序列供水和需求预测模型，进而计算供水量比、水损与需求比、水损与住宅需求比 3 个重要可持续指标，为城市供水系统可持续发展提供了管理建议。

以上模型预测结果表明，ARIMA 模型在日、月、季度等尺度上的预测精度均可接受，而在完善数据长度的条件下，预测精度将明显提高。此模型优点主要体现在模型简单、建模和预测速度快，可实现多步预测，并且对时、日、月和季供水量预测均有效（一般不用于年预测），在历史数据足够多且系统没有发生重大变化时可得到较高的预测精度，且可用于外推预测。然而 ARIMA 模型的预测精度取决于历史时间序列的长度，历史数据越多，预测精度越高，反之越低。由于实际供水量数据的收集难度较大，过程较为烦琐，所得数据的完整性难以保证，因此采用 ARIMA 模型的精确程度难以保证，并且缺乏适应性和灵活性。

1.2.2 新技术方法

根据目前研究成果，传统供水量预测方法存在以下缺点：一是预测模型大多基于全局建模的角度，而缺失了局部特征的代表性，影响