

自来水用户用水模式的确定方法研究

侯煜堃^{1,2}, 赵春会¹, 俞国平², 李福平¹

(1. 郑州市自来水总公司, 河南 郑州 450013; 2 同济大学 环境科学与工程学院, 上海 200092)

摘 要: 针对管网建模过程中用水量变化规律的复杂性,提出了根据实际用水量数据进行聚类分析,合理地对用户进行分类的方法,并通过主成分综合模型确定了每一类用户典型的用水模式曲线。该方法借助电磁水表实时远传的水量数据,能够及时更新用户的用水模式曲线,为提高建模的精确性打下了坚实的基础。

关键词: 给水管网; 用水模式; 聚类分析法; 主成分分析法

中图分类号: TU991 **文献标识码:** C **文章编号:** 1000 - 4602(2008)15 - 0063 - 04

Study on Determination of Water Consumption Mode for Drinking Water Customers

HOU Yu-kun^{1,2}, ZHAO Chun-hui¹, YU Guo-ping², LI Fu-ping¹

(1. Zhengzhou Water Supply Corporation, Zhengzhou 450013, China; 2 School of Environmental Science and Engineering, Tongji University, Shanghai 200092, China)

Abstract: Aimed at the complexity of water consumption variation during modeling of water distribution networks, a methodology based on actual water demand data is presented to categorize water consumers in conjunction with cluster analysis. The typical water consumption mode curves for every category of consumers are determined by taking advantage of integrated model of principal component analysis. By using the real-time remote data transmitted from electromagnetic meters, the method can update the curves of water consumption mode in time, which lays a solid foundation for improving the accuracy of network modeling.

Key words: water distribution network; water consumption mode; cluster analysis; principal component analysis

在建立给水管网微观水力模型的过程中,需要对用水量进行空间和时间分配。用水量空间分配通过建立水表与节点的对应关系来实现,用水量时间分配则是通过掌握各类用户的实际用水量变化规律,寻求每一类用户的用水模式来实现^[1]。

在恒定流模拟过程中,基本用水量不随时间变化。但在多工况延时模拟时,模型既需要基本水量数据,又需要水量随时间变化的信息,即用水模式。根据用水模式将用水量按不同时段分配后,就可进行连续的动态模拟。由于每个城市的水量变化各有

其特点,所以用水模式一般通过现场测试获得。过去,通过现场抄表的方法获取各类用户的用水模式,这既耗费大量的人力资源,又不能做到对用水模式的动态更新。除了在模型校验时采用这种方式外,一般情况下很难准确掌握用户的用水模式。

1 水量数据的获取

郑州市自来水总公司于 2006 年对部分大用户安装了 20 余台 DXL 系列安信电磁水表,将用户的水量数据以短消息的形式及时发送到营业收费部门(每天于 8:00 与 20:00 各发送一次)。这样,营业

收费部门就从过去的一个月抄表一次,转变为每天均可实时监控用户的水量变化情况。

DXL 系列安信电磁水表采用 3 6 V 锂电供电,不需外接电源,计量可达 6 年以上,数据远传 2 年以上;能显示和传送累计流量和瞬时流量;计量范围宽,可提供用水量与水表口径匹配特性分析;没有机械转动和传动机构,灵敏度高,性能优于国家 B 类水表。对郑州市 2007 年 3 月 1 日—4 月 30 日的 18 个用水大户的实测数据进行处理,去掉缺失数据太多的两个用户,并剔除其中数据显示不完整或显示错误的时段,针对剩余的 16 个大用户、共 22 d 的流量数据进行分析研究。

2 用户类型的确定

用户类型是在新户申请报装时根据用水性质确定的,同一类型的用户享有相同的水费费率,具体分类如表 1 所示。

表 1 未聚类时的用户类型

Tab 1 Categories of water consumers before cluster analysis

D 号	户名	水表口径	用户类型
N10081	华北水利水电学院 (龙子湖区 U1 东 3)	DN200	生活
N10091	华北水利水电学院 (龙子湖区 U1 东 2)	DN200	
N20011	中国人民解放军信息工程大学	DN200	
N30011	郑州大学后勤集团公司	DN200	
N30041	河南省体育中心	DN150	
N30051	华北水利水电学院(水科路)	DN150	
N30061	郑州牧业工程高等专科学校 学校伙食科	DN150	行政
N40031	河南工业大学	DN150	
N20041	郑州市第一中学	DN150	
N30071	郑州师范高等专科学校	DN150	
N40011	郑州铁路局中心医院	DN150	
N40081	武警部队河南总队医院	DN150	服务
N40041	郑州铁路局办公室	DN150	
N50031	郑州卷烟总厂	DN200	生产
N20051	郑州新力电力有限公司	DN150	
N10071	郑州市四十七中学	DN100	

但是随着时间的变化,同一用户的用水量变化规律在不同的时期会有所不同,同一类型的不同用户之间,其用水量的变化规律也可能相差较大,这就造成了归纳某类用户典型用水模式的困难。为此,考虑根据用户实际的用水规律进行分类,并找出每一类用户典型的用水模式。

聚类分析主要用于辨识具有相似性的事物,并根据彼此不同的特性加以“聚类”,使同一类的事物具有高度的相似性。在此次研究中,笔者就根据用户 24 h 的水量变化对其进行聚类分析,从而判别其分类情况。

采用聚类分析的层次方法,对给定的数据对象集合进行层次分解。将 16 个大用户的流量数据组成聚类分析的样本矩阵,选择欧氏距离的平方来度量样本间的距离,并采用最远邻法度量类间的距离,其结果是把 16 个大用户样本分为 5 类,聚类后的结果如表 2 所示(参照聚类前的分类对新用户类型重新进行了命名)。

表 2 聚类后的用户类型

Tab 2 Categories of water consumers after cluster analysis

D 号	新用户类型
N10081、N10091、N30061、N40031、N30071	新生活
N30051、N10071、N20051、N40081、N20041、N40041、N40011	新行政
N30011、N30041	新服务
N20011	新军校
N50031	新生产

对比表 1 和表 2 可知,通过聚类分析对用户类型进行重新分类后,部分用户的类型有了改变,例如 N20011 和 N50031 这两个用户分别独立成一类,分析原因可能是:N20011 是实行军事化管理的学校,有严格的作息制度,与一般的生活类型用户有很大的差别;而 N50031 是郑州卷烟总厂,其生产特点比较特殊,故需对它们单独进行用水模式分析,建立各自的用水模式曲线。“新行政”类型的用户保留了原来的 3 个用户,原来的另 1 个用户并入“新生活”类型,同时将其其他符合这个类型用水特点的 4 个用户归并过来。

3 同一类型用户典型用水模式曲线的确定

分别采用用水模式系数法和主成分综合法确定同一类用户典型的用水模式曲线,并对原有方法做了适当改进。

3.1 基于模式系数法的用水模式曲线

首先通过现场测试得到各时段的用水量和平均时用水量,两者相除即为模式系数^[2],然后根据用水模式系数绘制用水模式曲线。由于用户在一周内的用水模式并非完全相同,故考虑对一段时间内的模式系数数据进行横向处理(将所有数据按照星期数进行分类),得到周一至周日的 7 个横向数据集

合。以“新生活”用水类型为例,根据模式系数法确定其周日的用水模式曲线见图 1。可以看出,“新生活”用水类型连续几个周日的用水模式曲线相似。

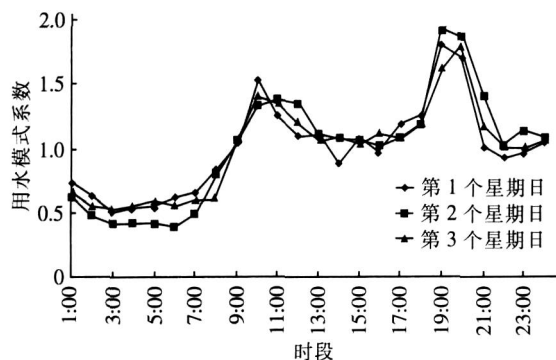


图 1 基于模式系数法的周日用水模式曲线

Fig 1 Water usage pattern curves on Sunday based on pattern factor method

3.2 基于主成分综合模型的用水模式曲线

在对用户进行科学分类的基础上,采用主成分分析法能确定某类用户的特征用户。该法可同时实现用水量的归纳和综合,所得序列彼此无关,少数几个用户序列即可包含全体变量的绝大部分信息,在此基础上可分别针对每一天的用水量建立“主成分综合”模型。

参照主成分分析的原理^[3],以“新生活”用水类型为例,建立主成分综合模型,得到一个“虚拟”用户一天的用水量,然后采用模式系数法确定用水模式。以周日为例,用水模式曲线如图 2所示。

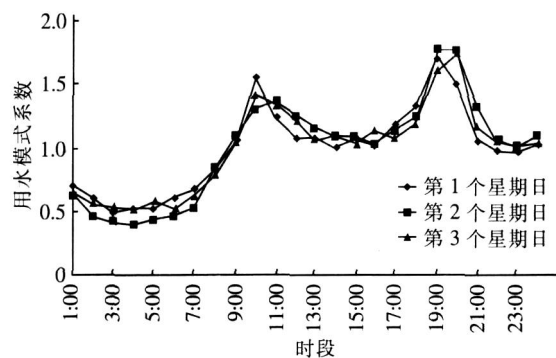


图 2 基于主成分综合法的周日用水模式曲线

Fig 2 Water usage pattern curves on Sunday based on integrated principal component analysis

3.3 两种方法的比较

标准差能反映出数据分散程度的差异,是衡量数据精密度大小的统计参数。根据图 1和图 2中的数据计算各时段多个用水模式系数的标准差,结果

如图 3所示。

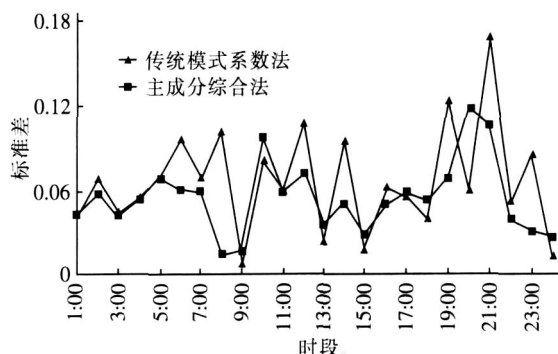


图 3 两种方法得到的用水模式标准差对比

Fig 3 Comparison of standard deviations of water usage pattern curves between two methods

由图 3可以看出,基于主成分综合模型得到的用水模式曲线的标准差较小,说明这种方法消除了某些时段的异常波动,更能代表“典型”的用水模式。通过对大量数据进行分析后发现,对于同一类用户,其包含的用户个数越多则利用主成分综合法得到的结果的趋同性越好。

若进一步对每一个数据集进行中位数处理,即针对每一时段,找出位于中位数的模式系数,就可得到某类用户“典型”的模式系数曲线。将一组数据按由小到大排列,用去掉两端逐步接近正中心的办法即可找到中位数。从统计的角度看,以中位数表达平均意念的最大好处是免受过高或偏低数据的影响,并可表示一组数据的“集中趋势”。

建立不同类型用户“典型”的用水模式曲线时,还可考虑针对每周 7 d分别建立用水模式曲线。按照这种方法建立的“新生活”类型用户一周的用水模式曲线如图 4所示。

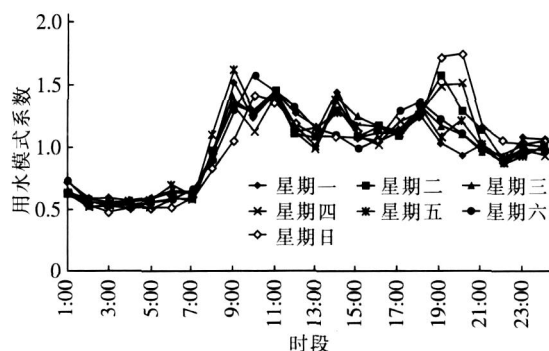


图 4 “新生活”类用户一周的用水模式曲线

Fig 4 New domestic water usage pattern curves during a week

(下转第 68页)

mg/L, 平均为 64.6 mg/L, 剩余氧化剂为 27.7% ~ 57.3%; 同期 COD 自动测定仪的测定值为 27.72 ~ 57.29 mg/L, 平均为 44.8 mg/L。

5 结论

① 威海中融安信污水厂出水中的 Cl^- 浓度较高, 对 COD 的测定产生了较大干扰。

② 氯气校正法、低浓度氧化剂法的测定结果基本一致, 可排除高浓度 Cl^- 的干扰, 结果可信。相比较而言, 低浓度氧化剂法的操作更简便易行, 可作为该厂出水 COD 监测的推荐方法。

参考文献:

- [1] 国家环境保护总局. 水和废水监测分析方法 (第 4 版) [M]. 北京: 中国环境科学出版社, 2002
- [2] 丛俏. Cl^- 、 NH_4^+ 对化学需氧量测定的影响及消除的研究 [J]. 渤海大学学报 (自然科学版), 2005, 26 (3):

209 - 213

- [3] 杨俊仕, 谢翼飞, 李旭东. 高氯离子废水中低 COD 简易分析方法研究 [J]. 油气田环境保护, 2002, 12 (3): 22 - 25.
- [4] 尹洧. 水和废水监测分析方法指南 (上册) [M]. 北京: 中国环境科学出版社, 1990
- [5] 陶大钧, 张信华, 孙晓斌, 等. 高氯离子废水中低浓度化学需氧量分析技术 [J]. 环境监测管理与技术, 1999, 11 (3): 35 - 37.
- [6] 张月, 简森夫. COD 测定中氯离子干扰的消除方法 [J]. 工业用水与废水, 2004, 35 (4): 55 - 57.
- [7] 戴竹青, 申开莲, 林大泉, 等. 高氯离子废水化学需氧量分析方法的研究 [J]. 中国环境监测, 2002, 18 (3): 21 - 25.

电话: 13803042826

收稿日期: 2008 - 05 - 20

(上接第 65 页)

4 用水模式曲线的动态更新

在传统抄表模式下对用水模式进行动态更新是非常困难的。随着电磁水表、远传水表等新型计量设备的应用, 流量数据的实时传输与监控成为可能, 可方便地收集、分析用户 (特别是大用户) 的水量数据与用水变化规律。可按照一定的时间间隔 (例如一个月), 定期制定各类用户的典型用水模式曲线 (并且按每周 7 d 细分), 这样当下一个月的用水量数据更新时, 可同时更新各类用户的用水模式曲线。如果电磁水表、远传水表不能覆盖大多数的大用户, 那么可实行轮换制度, 按一定的周期和比例将普通水表更换为电磁水表或远传水表, 待收集了足够的水量数据后, 再将其轮换到其他大用户, 同样可以实现用水模式的科学归类、准确获取与动态更新。

5 结论与建议

① 当采用电磁水表或远传水表收集用户的水量数据并分析用水量变化规律时, 其时间间隔最好设置为 30 min 或更短一些, 这样得到的用水模式曲线更准确。

② 根据实际的用水量数据对用户进行聚类分析, 可以得到符合实际用水量变化规律的用户分类。

③ 基于主成分综合模型, 并结合模式系数法与中位数法得到的用水模式曲线可作为每一类用户

典型的用水模式曲线。

④ 对生活类用户来说, 建议根据一周 7 d 划分为七条不同的典型用水模式曲线; 对其他类用户, 也应结合实际水量数据判别是采用一条、两条 (平常日、周末或节假日) 还是多条典型的用水模式曲线。

⑤ 用水模式曲线应随着水量数据的定期更新而进行动态更新。在电磁水表、远传水表数量不足, 不能覆盖所有大用户的情况下, 可实行轮换制度, 分阶段、分批地收集众多大用户的水量数据, 然后进行用户用水规律与用水模式分析。

参考文献:

- [1] Thomas M W, Chase D V, Savic D A, et al. Advanced Water Distribution Modeling and Management [M]. Waterbury: Haestad Press, 2001.
- [2] 赵洪宾. 给水管网系统理论与分析 [M]. 北京: 中国建筑工业出版社, 2003.
- [3] 卢纹岱. SPSS for windows 统计分析 [M]. 北京: 电子工业出版社, 2004.

电话: (0371) 67695404 13838392286

E-mail: hou_yukun@126.com

收稿日期: 2008 - 02 - 20