

一元回归分析在污水处理厂中的应用

Application of Single Regression Analysis on Wastewater Treatment Plant

宫艳萍¹, 王 劼^{1,2}

(1. 国电东北环保产业集团有限公司, 辽宁 沈阳 110014;

2. 东北大学资源与土木工程学院, 辽宁 沈阳 110004)

摘 要: 以沈阳某污水处理厂的进水水质为例, 用 SPSS 软件对 COD 和 BOD 进行回归分析。结果表明, COD 和 BOD 之间具有线性关系, 其一元线性回归方程为 $COD=35.76+2.26*BOD$ 。利用该回归方程测算 BOD, 对 BOD 预测值和实际值进行 t 检验, 发现两者之间不存在显著性差异, 从而表明利用回归方程测算 BOD 方便快捷, 有助于生产运行管理。

关键词: 一元回归; 污水处理厂; SPSS 软件

中图分类号: TU992.3

文献标识码: A

Abstract: Taking the influent water quality of a wastewater treatment plant in Shenyang as example, this paper analyzed the relationship between COD values and BOD values by SPSS software. The results indicate that values of COD and BOD are in linear relation. The linear regression equation is $COD=35.76+2.26*BOD$. Compared the BOD value estimated by regression equation and the actual value with t test, there is no significant difference. So, using regression equation to estimate BOD value is convenient and helpful for management.

Key words: Single Regression; Wastewater Treatment Plant; SPSS Software

CLC number: TU992.3

BOD₅和COD_{Cr}是污水处理厂每天必测的重要参数^[1], 但是在污水厂实际水质监测中, 一般均以COD_{Cr}为主, 这主要是因为监测COD_{Cr}较易, 用哈希或其他先进仪器检测, 很短的时间就可以得出结果, 而测定BOD₅耗时长, 无论是国标法或是其他测试系统, 一般要在5 d以后才得到数据, 因此快速测定BOD₅显得非常重要。实际上, 在污水的水质组成中, 各组分之间存在着某种关联, 有些组分之间存在线性关系, 如能找出它们之间的线性回归方程, 就可以很容易推算出其他较难测定的数据, 从而便于生产管理。关于COD_{Cr}和BOD₅之间的关系国内有过报道^[2,3], 但对具体的污水厂, 这些公式则存在很大差异。文章

主要通过对沈阳市某污水处理厂COD_{Cr}和BOD₅之间的关系研究, 利用一元回归分析揭示其中的规律, 从而通过COD_{Cr}来快速测算BOD₅这一重要数据, 以期更好指导水厂生产运行。

1 一元线性回归概述

一元线性回归是指一个变量与另一个变量之间的线性关系。设 X 和 Y 来表示两个随机变量, 自变量 x 的统计值为 $x_1, x_2, \dots, x_n$, 因变量 y 为 $y_1, y_2, \dots, y_n$, 则它们之间的线性关系, 可用下述线性函数来表示: 即

$$y_i = a + bx_i + \varepsilon_i \quad (i=1, 2, \dots, n)$$

式中 x_i —称自变量, 即影响预测变量的因素

收稿日期: 2012-10-20

作者简介: 宫艳萍 (1971-), 女, 高级工程师。研究方向: 污水处理生产管理。E-mail: scglb666@sina.cn

变量; y_i —称因变量, 实际是依赖于自变量 x_i 的随机变量;

n —称样本容量, 即观测数据点个数;

a, b —待求的回归方程的参数;

ε_i —称回归剩余项(或称随机干扰项), 由于总体回归模型的参数: a, b, c 都是未知的, 可以利用试验观测值对它们进行估计, 得到相应的估计的回归方程^[4]:

$$\hat{y}_i = \hat{a} + \hat{b}x_i \quad (i=1, 2, \dots, n)$$

2 一元线性回归在污水厂数据分析中的具体应用

沈阳某污水处理厂的2011年全年进水数据的平均值见表1。

表1 沈阳某污水处理厂的进水数据 $\text{mg} \cdot \text{L}^{-1}$

$t/\text{月}$	COD_{Cr}	BOD_5
1	108	46
2	132	48
3	140	44
4	160	55
5	164	50
6	166	54
7	170	56
8	200	77
9	202	75
10	212	70
11	220	75
12	234	94

对表中是数据用SPSS软件进行分析^[5], 得到散点见图1。

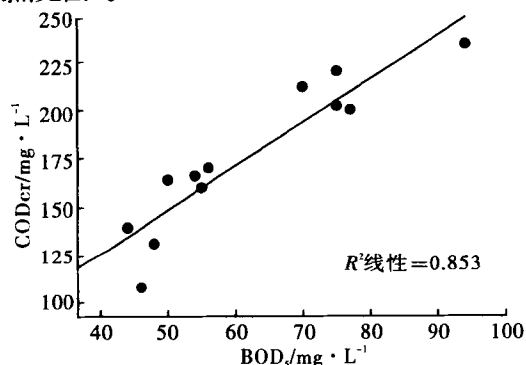


图1 沈阳某污水处理厂COD和BOD散点图

从图1中可以看出, $R^2=0.853$, COD和BOD之间线性关系较好。沈阳某污水处理厂进水数据方差分析见表2。

表2 沈阳某污水处理厂进水数据方差分析

统计量	自由度	SS	均方	F 检验	显著性差异
回归分析	1	13851.12	13851.12	58.01	0
残差	10	2387.55	238.76		
总计	11	16238.67			

从表2中可以看出, 方差分析是把总离散平方和及其自由度进行分解, 利用 F 统计量检验两个变量之间线性相关显著性的一种方法。从上面的方差分析结果可以看出, 分子自由度为1, 分母自由度为10的 F 值为58.014, 其相应 $P=0$, 小于0.05, 这说明, COD和BOD之间具有线性关系。可以做回归分析, 见表3。

表3 沈阳某污水处理厂进水数据一元回归系数

系数	标准误差	T 检验	概率值
35.76	18.903	1.892	0.088
2.26	0.296	7.617	0.000

从表3中可以看出, 沈阳某污水处理厂的进水 $\text{COD}=35.76+2.26 \times \text{BOD}$ 。该一元回归方程中的35.756为常数项, 2.257为偏回归系数, 它表示在COD每改变一个单位, BOD平均改变2.26个单位。

利用该回归方程, 对BOD进行预测, 预测值和实际值见表4。

表4 BOD预测值和实测值比较

次数	预测值/ $\text{mg} \cdot \text{L}^{-1}$	实际值/ $\text{mg} \cdot \text{L}^{-1}$	误差/%
1	115.9	111.2	4.06
2	70.5	70.4	0.14
3	83.2	87.7	-5.41
4	79.8	74.6	6.52
5	73.4	66.2	9.81
6	75.2	72.8	3.19
7	83.5	89.3	-6.95
8	79.2	72.9	7.95
9	59.8	60.1	-0.50
10	63.6	64.1	-0.79
11	81.8	84.1	-2.81
12	97.5	94.3	3.28
13	69.2	67.0	3.18
14	62.4	57.1	8.49
15	78.3	75.2	3.96

将BOD预测值和实际值进行 t 检验, 结果见表5。 (下转第28页)

中污水厂的事故排污频率较大,由于径流水少且污水在河道内滞流,形成局部的黑臭河道。

5 结语

(1) 蒲河是沈阳地区具有北方特点的中型天然河流,具有水资源欠缺与受控、受城市排污影响和富营养化的代表性特征。该类中小河流的污染控制与水生态修复对于大河流的水质安全具有重要影响。

(2) 以蒲河为代表的北方中小河流已形成由污染控制型向生态修复与管理型转变的带状人工河湖,必须以水资源保护、水体健康与污染控制和生态河道建设为主要建设内容。

(3) 现阶段城市河流的污染控制仍是生态修复的重要内容,特别是要提高污水的收集处理

率和处理效率、控制各类污染事故的发生。

(4) 城市中小河流最优生态平衡的建立要依托流域绿色经济的发展,应研究建立环境受体与承载能力的最佳平衡点,保证经济发展和水生态修复目标的和谐和统一。

参考文献

[1]荆 勇.蒲河流域污染状况与生态修复技术对策的研究[J].环境保护科学, 2011, 37(6): 29-32.
[2]刘卓也,李 崇,王 哲.蒲河流域水生态系统评价及生态修复研究[J].东北水利水电, 2011(04): 41-43.
[3]荆 勇.沈阳市蒲河水水质污染特征及水质改善途径的研究[J].环境保护科学, 2012, 38(4):33-36.
[4]杨东溟.沈阳要将蒲河流域打造成生态“大公园”[N].辽宁日报, 2010-04-01 (002).
[5]魏雁飞.蒲河生态廊道建设开展仪式在沈阳市档案馆隆重举行[J].兰台世界, 2012(28):44.

(上接第14页)

表5 实测值和预测值之间成对数据t检验

对数	比较	成对差异					t	自由度	显著性
		平均	标准差	标准误	置信区间				
					下限	上限			
Pair 1	BOD实际值 -BOD预测值	-1.75	3.90	1.01	-3.91	0.40	-1.74	14	0.103

从表5中可以看出, BOD预测值和实际值配对检验为 $p=0.103$, 大于0.05, 不具有统计学意义, 可以认为, 预测值和实际值不存在显著差异, 利用线性回归方程预测的BOD值和真实值比较接近。

3 讨论

对于某个特定的污水处理厂而言, 由于其汇水面积一定, 在该汇水区域内, 工厂从原料, 加工过程、生产工艺流程等相对固定, 反应到工业废水方面, 其水质就会表现出某种规律性的现象; 而对于生活污水而言, 由于在一定时期, 该区域的居民人口数量较为稳定、居民的生活习惯、生活条件、食物结构变化不大, 因此, 该区域排出的工业废水和生活污水中的各种有机物和无机物的组成相对稳定, 从而使 BOD_5 和 COD_{Cr} 的

比值围绕一定值波动, 可以近似地用一元回归方程表示: $COD=35.76+ 2.26*BOD$ 。

在污水厂水质的实际检测过程中, 测定BOD操作复杂, 需要5天甚至更长的时间, 而测定COD相对较易。利用该回归方程, 就可以很容易地根据COD推算出BOD的数值, 这对于水厂生产运行的快速诊断具有重要意义。

参考文献

[1]国家环境保护总局, 水和废水监测分析方法编委会. 水和废水监测分析方法(第四版)[M]. 北京: 中国科学出版社, 2011, 210-220, 227-234.
[2]司慧敏, 张景馨, 宋雅范, 等. 水质监测中 BOD_5 与 COD 值相关性研究[J]. 化学分析计量, 2005, 14 (2): 50-51.
[3]凌小芹. 工业废水 BOD_5 与 COD 相关关系以及 BOD_5 测定的简化法[J]. 中国环境监测, 2003(3): 30.
[4]杜荣赛. 生物统计学[M]. 北京: 高教出版社, 1999.07.
[5]罗应婷, 杨钰娟. SPSS统计分析从基础到实践[M]. 北京: 电子工业出版社, 2007:06.