Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 119 (2015) 798 – 802

Procedia Engineering

www.elsevier.com/locate/procedia

13th Computer Control for Water Industry Conference, CCWI 2015

Log logistic distribution to model water demand data

Seevali Surendran[1, 2]* and Kiran Tota-Maharaj[1]

[1]University of Greenwich, Faculty of Engineering & Science, Kent, ME4 4TB, United Kingdom.

[2]Environment Agency, Kings Meadow house, Kings Meadow Road, Reading, RG1 8DQ, United Kingdom

Abstract

There had been insufficient studies previously to conclude the suitability of the appropriate probability distribution functions in modelling water demand. The purpose of this study is to find an appropriate probability density function to apply in simulating water demand using real water consumption data. Daily water consumption data for four years obtained from a water company in UK and analysed using normal, log normal, log logistic and Weibull distributions and a comparison on the applicability of each distribution was assessed. Statistical modelling was performed using MINITAB. The Anderson Darling (AD) statistic was used as the goodness of fit parameter in the analysis.

Peer-review under responsibility of the Scientific Committee of CCWI 2015

Keywords: Water demand, stochastic nature, probability distribution function, log logistic distribution, Minitab, Anderson Darling

1. Introduction

The aim of this research is to study real water consumption data and establish a standard statistical distribution to use in water demand modelling. Water demand data for 4 years were obtained from a water company in UK for this study.

* Corresponding author. Tel.: +4401189533147
E-mail address: seevali.surendran@environment-agency.gov.uk

Peer-review under responsibility of the Scientific Committee of CCWI 2015
doi:10.1016/j.proeng.2015.08.940

Water demand varies with time of use, season and socio economic pattern of the consumers. It has never been same at any particular time and hence defined as a continuous random variable. Therefore incorporating variation of demands in modelling will lead to more realistic assessments of the performance of water distribution systems. However, few studies can be found in which the random variations of demands have been considered.
Goulter and Bouchart (1990), Xu and Goulter (1997, 1998, 1999) made an assumption that the demands have a normal distribution. Mays (1994) used randomly generated water consumption data using a range of distributions to study the sensitivity of the system's performance to changes in water consumption patterns. Khomsi et al. (1996) stated that the demand is behaving as having a normal distribution based on the Kolmogorov-Smirnov test. Surendran and Tanyimboh (2002), Tanyimboh and Surendran (2004) addressed the issue of the modelling of short term demand variations in a comprehensive way using UK water demand data and concluded, water demand data fit well in to log logistic distribution than a normal distribution.

AWWA Research foundation sponsored a study (Bowen et al.1993) in residential water demand use patterns in USA results, revealed that the demand data was not distributed normally. Several data transformations to improve the data analysis were investigated and it was found that the log transformation was only mildly effective in reducing the positive skewness of the frequency distributions of the data, making them more nearly normal.

2. Methodology

Daily water consumption data for 4 years from a UK water company was obtained and analysed. The data was collected using data loggers at 15 minutes interval. The data was obtained from April 2009 to April 2013.

In this research model selection was done in a series of applications. Data were screened and sorted to select the model. Raw demand data was drawn against time. This provided a quick reference to check the accuracy of data. If the points were homogeneously distributed and there was no negative points, this meant that the data is almost accurate. Similarly if there was any inconsistently in distribution, this would allow us to remove all abnormal data points.

2.1 Data Analysis

Following the sorting out of data, the data was then analysed (using MINITAB statistical package to fit into a probability distribution. Continuous distributions such as, normal, Log- normal, Weibull and log-logistic were applied) to find a suitable distributions.

There are various numerical and graphical methods used in estimating the parameters of a probability distribution. In this research, graphical methods were selected for the analysis along with the Maximum Likelihood method to draw the probability plots (Fig 1-4). The data were analysed using 95% confidence interval (5% significant level). Once the data was fitted to normal, lognormal and log logistic distributions, parameters of the particular distribution such as location, shape and scale were essential to describe the distribution. Table 1 shows the parameter estimates for the obtained data.