Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Physica A 355 (2005) 46–53

PHYSICA A

www.elsevier.com/locate/physa

Prediction of energy consumption and risk of excess demand in a distribution system

Marko Thaler*, Igor Grabec, Alojz Poredoš

Faculty of Mechanical Engineering, University of Ljubljana, Aškerčeva 6, SI-1000 Ljubljana, Slovenia

Received 31 October 2004; received in revised form 28 December 2004
Available online 4 May 2005

Abstract

An empirical model for prediction of energy consumption in a distribution system is described. The model resembles a normalized radial basis function neural network whose neurons contain prototype joint data about the consumption process and the environment. A set of prototype patterns of consumption and environmental variables is formed from a record of a multi-component time series by a self-organized process. Prediction of energy consumption is performed by a conditional average estimator based upon known prototype patterns and given future values of environmental variables. Importance of these variables for the prediction is determined by a genetic algorithm. Prediction performance of the model is tested on a one-year-long consumption record of a gas distribution system. Prediction error is determined by the difference between predicted and actually observed consumption. Its value depends on time and amounts to a few percent of the actual consumption. The probability distribution of prediction error is estimated from a properly selected time interval of prediction. This distribution can be used to estimate the risk of energy demand beyond some prescribed value. For an optimization of the distribution process, a cost function that includes operation and control costs of a distribution system as well as penalties related to excess energy demand is proposed. Its minimum corresponds to an economically optimal energy distribution.

PACS: 05.45.Tp; 07.05.Mh

Keywords: Energy consumption prediction; Self-organization; Conditional average; Genetic algorithm; Risk of excess demand

*Corresponding author. Tel.: +386 1 47 71 446; fax: +386 1 25 18 567.
E-mail address: marko.thaler@fs.uni-lj.si (M. Thaler).

0378-4371/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2005.02.066

1. Introduction

Energy distribution systems exhibit appreciable fluctuations of energy consumption over shorter and longer time intervals. These fluctuations are a consequence of the changing needs of clients that are related to the population activity and environmental conditions. A distributor must satisfy the guaranteed energy demand of clients and should not exceed the contractually determined energy supply of providers. Thus the basic task of a distributor is to minimize the difference between requested and provided energy amounts by control actions in a way to achieve maximal economic profit. For this purpose, a reliable prediction [1] of energy consumption [2] is needed. One of the basic tasks of prediction is the estimation of excess energy demand risk, since it is related with an excess cost.

Energy consumption is generally a high-dimensional complex process that depends on deterministic and random components. Deterministic components can be analytically described by characteristics of the distribution system and some variables that characterize the influence of the environment. Conversely, random components cannot be described analytically but only probabilistically. Reduction of the related uncertainty in system description can be achieved by properly accounting various environmental variables and their influence on energy consumption. A non-parametric regression [3] and genetic algorithm [4] are applied for this purpose.

2. Empirical information and self-organized formation of prototype vectors

A time series of energy consumption and environmental variables that can include data about day in the week, season, environmental temperature, wind speed, solar radiation, etc. must be jointly represented by data that are suitable for application in a statistical model. For this purpose, we form representative joint data vectors $\mathbf{z}_i$ from the given record of a multi-component time series as follows:

$$\underbrace{\mathbf{z}_i}_{\text{data vector}} = \left(\underbrace{x_{i,1}, x_{i,2}, \ldots, x_{i,N}}_{\text{environmental variables vector}}, \underbrace{y_{i,1}, y_{i,2}, \ldots, y_{i,M}}_{\text{energy consumption vector}} \right). \tag{1}$$

This vector represents recorded values of environmental and energy consumption variable at the ith point of the time series. N is the number of considered environmental variables and M the number of consumption values around the selected point. Environmental variables at a particular time can also be represented by vectors similarly, as the energy consumption variable. The set of data vectors is a source of empirical information about relations between environmental variables and consumption. It represents a basis for the consumption prediction if the future values of environmental variables are provided. The number of data vectors that are extracted from the time series record can be reduced by a self-organized adaptation of a smaller number of prototype data vectors $\mathbf{q}_p = (\mathbf{x}_p, \mathbf{y}_p)$. Here $\mathbf{x}_p$ represents the vector of environmental variables and $\mathbf{y}_p$ the consumption variable vector in pth