Expert Systems with Applications 36 (2009) 4248–4252

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Prediction of aeration efficiency on stepped cascades by using least square support vector machines

Davut Hanbay [a], Ahmet Baylar [b,*], Murat Batan [b]

[a] Firat University, Electronic and Computer Science Department, Elazig, Turkey
[b] Firat University, Civil Engineering Department, Elazig, Turkey

ARTICLE INFO

Keywords:
Support vector machine
Stepped cascade
Flow condition
Aeration efficiency

ABSTRACT

It is important to predict aeration efficiency in stepped cascades because they are used in most water treatment applications for re-oxygenation. The flow conditions on stepped cascades have been classified into nappe, transition and skimming flows. Due to the different mechanisms of air entrainment in the nappe, transition and skimming flow conditions, the aeration efficiencies of the three flow conditions differ significantly from each other. In this paper, two intelligent models were created to predict flow condition and aeration efficiency in stepped cascades using critical flow depth, step height and channel slope information. Least square support vector machine (LS-SVM) was used as intelligent tool. The performances of LS-SVM models were evaluated by 3-fold cross validation test method. The correlation between observed and predicted flow condition is 0.99 and the correlation between measured and predicted aeration efficiency is 0.89. The test results indicated that the LS-SVM can be used successfully in predicting flow condition and aeration efficiency in stepped cascades.

1. Introduction

Water quality and its enhancement have a close connection with the presence of dissolved oxygen. In fact, the oxygen concentration in surface waters is a prime indicator of the water quality for human use as well as for the aquatic biota. The physical process of oxygen transfer or oxygen absorption from the atmosphere acts to replenish the used oxygen. This process has been termed re-aeration or aeration. Aeration enhancement by macro-roughness is well-known in water treatment, and one form is the aeration cascade. Re-aeration cascades are used in water treatment for re-oxygenation, denitrification or volatile organic compounds (VOC) removals. In the treatment of drinking water, cascade aeration may be used to remove chlorine and to eliminate or reduce offensive taste and odor. In summary, stepped cascades are very efficient means of aeration because of the strong turbulent mixing, the large residence time and the substantial air bubble entrainment (Toombes & Chanson, 2000).

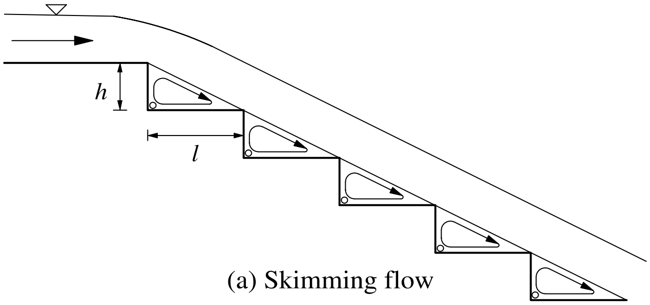
Stepped flows can be classified into skimming flow, transition flow, and nappe flow. For narrow steps or larger discharges such as the design discharge the water skims over the step corners and recirculating zones develop in triangular niches formed by the step faces and the pseudo-bottom, as shown in Fig. 1a. In skimming flow the water flows as a coherent stream over the pseudo-bottom formed by the step corners. For a range of intermediate discharges, a transition flow condition takes place. The dominant feature is stagnation on the horizontal step face associated with significant splashing and a chaotic appearance (Fig. 1b). For nappe flow the steps act as a series of overfalls with the water plunging from one step to another (Fig. 1c). Generally speaking nappe flow is found for low discharges and wide steps (Chanson, 2001).

The hydraulic design of stepped cascades, and in particular, characteristics of nappe flow, transition flow and skimming flow over stepped cascades has been studied experimentally by a number of investigators. Recently, Baylar and Emiroglu (2003), Emiroglu and Baylar (2003), Baylar and Emiroglu (2004), Baylar and Emiroglu (2005), Baylar, Emiroglu, and Bagatur (2006), Emiroglu and Baylar (2006) and Baylar et al. (2007a, 2007b, 2007c, in press) did some detailed studies on the aeration efficiency of stepped cascades. In recent years, the developments in intelligent methods make them possible to use in complex systems modeling. Support vector machines (SVMs) introduced by Vapnik and his co-workers are a particular instance of this. SVMs are a powerful methodology and have been studied extensively for classification, regression and pattern recognition. The basic idea of SVMs is mapping the input data points to a high-dimensional feature space and finds a hyperplane. The least square (LS) version of the SVM was described in Suykens and Vandewalle (1999). LS-SVM is widely used in complex system studies for modeling, regression or parameter prediction. However, its applications to the hydraulic systems are very limited. In this study, LS-SVM based classification and regression were

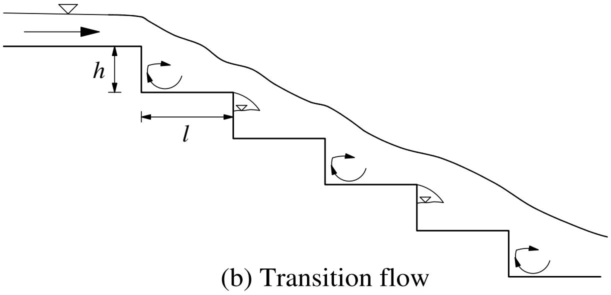
---

* Corresponding author.
E-mail address: abaylar@firat.edu.tr (A. Baylar).

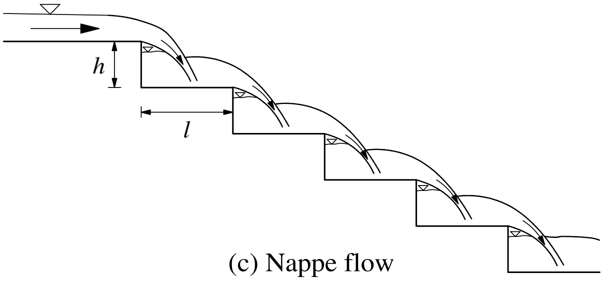
0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.03.003



(a) Skimming flow



(b) Transition flow



(c) Nappe flow

Fig. 1. Flow conditions over stepped cascades.

realized. The flow condition and aeration efficiency in stepped cascades were predicted based on LS-SVM. The organization of the paper is as follow; at first theoretical background of the stepped cascades and LS-SVM are presented. Then realized study is described in details. At last obtained results are concluded.

2. Aeration efficiency

The physical process of oxygen transfer or oxygen absorption from the atmosphere acts to replenish the used oxygen. This process has been termed re-aeration or aeration. Aeration efficiency, E may be defined as (Gulliver, Thene, & Rindels (1990))

$$E = \frac{C_d - C_u}{C_s - C_u} \tag{1}$$

where C is dissolved oxygen (DO) concentration, C_s is saturation concentration and u and d are subscripts indicating upstream and downstream locations.

A transfer efficiency value of 1.0 means that the full transfer up to the saturation value has occurred at the structure. No transfer would correspond to $E = 0.0$. The saturation concentration in distilled, deionized water may be obtained from charts or equations. This is an approximation because the saturation DO concentration for natural waters is often different from that of distilled, deionized water due to the salinity affects.

Comparative evaluations of oxygen uptake at hydraulic structures require that aeration efficiency is corrected to a reference temperature. To provide a uniform basis for comparison of measurement results, the aeration efficiency is often normalized to a 20 °C standard. Gulliver et al. (1990) proposed the following equation to describe the influence of temperature

$$1 - E_{20} = (1 - E)^{1/f} \tag{2}$$

where E is transfer efficiency at actual water temperature, E_{20} is transfer efficiency for 20 °C and f is exponent described by

$$f = 1.0 + 2.1 \times 10^{-2}(T - 20) + 8.26 \times 10^{-5}(T - 20)^2 \tag{3}$$

where T is water temperature.

In this study, the aeration efficiency was normalized to 20 °C using Eq. (2).

3. Least square support vector machine (LS-SVM)

The SVM is a supervised learning technique which is developed by Vladimir Vapnik and co-workers at AT&T Bell Laboratories in 1995. It is applicable to both classification and regression. The SVMs are based on the principle of structural risk minimization (Cortes & Vapnik, 1995).

Consider a given training set $\{x_k, y_k\}_{k=1}^{N}$ with input data $x_k \in R^n$ and output data $y_k \in R$ with class labels $y_k \in \{-1, +1\}$ and linear classifier

$$y(x) = \operatorname{sign}[w^T x + b] \tag{4}$$

When the data of the two classes are separable one can say

$$\left\{ \begin{array}{ll} w^T x_k + b \geqslant +1, & \text{if} \quad y_k = +1 \\ w^T x_k + b \leqslant -1, & \text{if} \quad y_k = -1 \end{array} \right\} \tag{5}$$

These two sets of inequalities can be combined into one single set as follows:

$$y_k[w^T x_k + b] \geqslant 1, \; k = 1, \ldots, N \tag{6}$$

SVM formulations are done with in a context of convex optimization theory. The general methodology is to start formulating the problem as a constrained optimization problem, next formulate the Lagrangian and then take the conditions for optimality, finally solve the problem in the dual space of Lagrange multipliers. With resulting classifier

$$y(x) = \operatorname{sign}\left[\sum_{k=1}^{N} \alpha_k y_k x_k^T x + b\right] \tag{7}$$

This linear SVM classifier was extended to non-separable case by Cortes and Vapnik (1995). It is done by taking additional slack variable in the problem formulation. One modifies the set of inequalities into

$$y_k[w^T x_k + b] \geqslant 1 - \xi_k, \quad k = 1, \ldots, N \tag{8}$$

The SVM has been used for linear and nonlinear function estimation too. The details of them can be found in literature. The least square version of the SVM classifier was described by Suykens and Vandewalle (1999). The LS-SVM considers equality type constraints instead of inequalities as in the classic SVM approach. This reformulation greatly simplifies a problem such that the LS-SVM solution follows directly from solving a set of linear equations rather than from a convex quadratic program. For a LS-SVM classifier, in the primal space it takes the form,

$$y(x) = \operatorname{sign}[w^T x + b] \tag{9}$$

where b is a real constant. For nonlinear classification, the LS-SVM classifier in the dual space takes the form,

$$y(x) = \operatorname{sign}\left[\sum_{k=1}^{N} \alpha_k y_k K(x, x_k) + b\right] \tag{10}$$

where α_i are positive real constants and b is a real constant, in general, $K(x_i, x) = \langle \phi(x_i), \phi(x) \rangle$, $\langle \bullet, \bullet \rangle$ is inner product, and $\phi(x)$ is the