

Ecological Informatics Applications in Water Management



6-7 November 2002

't Pand, Ghent University

Onderbergen
B-9000 Gent

Abstract book

Peter Goethals & Niels De Pauw

Prediction of class membership by means of support vector machines

Wies Akkermans¹, Piet Verdonshot², P.J. Goedhart¹ & Rebi Nijboer²

¹*Biometris, Wageningen UR, P.O.Box 16, 6700 AA Wageningen, the Netherlands,
email: L.m.w.akkermans@plant.wag-ur.nl.*

²*Alterra Green World Research, P.O.Box 47, 6700 AA Wageningen, the Netherlands*

Support Vector Machines (SVMs) are nonlinear predictors that can be used both for classification and regression. In this paper the emphasis is on classification.

In the SVM framework class boundaries are defined by so-called 'boundary observations' or support vectors. Hence support vectors are data points lying in some sense 'between' two classes. The search for these boundary observations is the main task of a Support Vector Machine.

It appears that SVMs can be expressed as quadratic programming problems, which implies that they have a well-behaved loss function and clearly defined stop criterion. This can be considered as an advantage of this method over Neural Networks.

In this paper a data set of freshwater sites is considered, which has been classified into cenotypes on the basis of macro-invertebrate species abundance. An SVM is fitted to these data in order to predict the cenotypes using only environment variables. The performance of the classifier is assessed by means of leave-one-out crossvalidation. The outcome is compared to results obtained with multinomial logistic regression.