Geophysical Research Abstracts
Vol. 13, EGU2011-10535, 2011
EGU General Assembly 2011

Support vector machine based model for water content in soil interpolation

Milan Cisty, Zbynek Bajtek, and Juraj Bezak

Slovak University of Technology Bratislava, Department of Land and Water Resources Management, Bratislava, Slovakia (milan.cisty@stuba.sk)

Modeling and simulation of hydrological processes in the unsaturated zone of soil is useful decision support tool in agriculture or water management. Because the water content in soil is affected by many factors which are not easily quantified, it is difficult to achieve accurate results of soil moisture forecasting. Searching for a new and better modeling methods to refine the forecast results is motivation of this work. The authors are describing soil water content model utilizing support vector machine (SVM), which belongs to the data-driven classification and regression tools. Accuracy of SVM application to this regression problem - predicting soil moisture in the five underground depths - is compared with the model created on the basis of artificial neural network (multi-layer perceptron).

Models developed are proposed for interpolation of initially obtained values of soil moisture, which were measured in two-week interval, to the daily values. This is special case of forecasting, which is used in the phase of building more complex, hybrid model with physically based and data driven based parts (SVM is intended for data preparation).

For both methods (MLP and SVM) have been used same training and testing input data sets. The data used to construct these models were measured in the investigated zone (Báč, south Slovakia), where soil moisture was measured at five underground depth horizons, and data from near meteorological station (Gabčíkovo) were used as inputs, namely: data average daily temperatures (T), daily precipitation totals (Z), 20-days precipitation totals (U), 20-days average temperature (T_{20}) and antecedent precipitation index also for 20-days (API). Following combination of data were chosen for alternative models developed for both methodologies (SVM and MLP):

Model 1: $\theta_t = f(T_{t-1}, T_{t-2}, T_{t-3}, T_{t-4}, T_{t-5}, Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-4}, Z_{t-5}, U_{20}, T_{20})$

Model 2: $\theta_t = f(T_{t-1}, T_{t-2}, T_{t-3}, T_{t-4}, T_{t-5}, Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-4}, Z_{t-5})$

Model 3 $\theta_t = f(T_{t-1}, T_{t-2}, ..., T_{t-20}, Z_{t-1}, Z_{t-2}, ..., Z_{t-20})$

Model 4 $\theta_t = ff(T_{t-1}, T_{t-2}, T_{t-3}, T_{t-4}, T_{t-5}, Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-4}, Z_{t-5}, API_{20}, T_{20})$

Model 5 $\theta_t = f(Z_{t-1}, Z_{t-2}, Z_{t-3}, Z_{t-4}, Z_{t-5}, U_{20}, T_{20})$

Where θ_t is predicted value of soil moisture.

The results of both models were evaluated using statistical characteristics. Correlation coefficients for the calculations using MLP ranged from 0.55622 to 0.88913, and the average value is 0.735. SVM correlation coefficients obtained have values from 0.73924 to 0.88636 with an average value of 0.828. The results obtained by SVM are stable, unlike the MLP, which sometimes fall into local minima.

This method is intended as an alternative to the standard physical and mathematical modeling and has advantage of a smaller number of required input data. Moreover in context of standard models it is often capable to provide more accurate result for this particular task and what if scenarios could be done with such a hybrid, more complex modeling architecture.

This work was supported by the Slovak Research and Development Agency under the contract No. LPP - 0319-09.