Geophysical Research Abstracts
Vol. 17, EGU2015-6127, 2015
EGU General Assembly 2015

Mapping of Estimations and Prediction Intervals Using Extreme Learning Machines

Michael Leuenberger and Mikhail Kanevski

Institute of Earth Surface Dynamics (IDYST), University of Lausanne, Switzerland (Michael.Leuenberger@unil.ch, Mikhail.Kanevski@unil.ch)

Due to the large amount and complexity of data available nowadays in environmental sciences, we face the need to apply more robust methodology allowing analyses and understanding of the phenomena under study. One particular but very important aspect of this understanding is the reliability of generated prediction models. From the data collection to the prediction map, several sources of error can occur and affect the final result. Theses sources are mainly identified as uncertainty in data (data noise), and uncertainty in the model. Their combination leads to the so-called prediction interval. Quantifying these two categories of uncertainty allows a finer understanding of phenomena under study and a better assessment of the prediction accuracy.

The present research deals with a methodology combining a machine learning algorithm (ELM - Extreme Learning Machine) with a bootstrap-based procedure. Developed by G.-B. Huang et al. (2006), ELM is an artificial neural network following the structure of a multilayer perceptron (MLP) with one single hidden layer. Compared to classical MLP, ELM has the ability to learn faster without loss of accuracy, and need only one hyper-parameter to be fitted (that is the number of nodes in the hidden layer).

The key steps of the proposed method are as following: sample from the original data a variety of subsets using bootstrapping; from these subsets, train and validate ELM models; and compute residuals. Then, the same procedure is performed a second time with only the squared training residuals. Finally, taking into account the two modeling levels allows developing the mean prediction map, the model uncertainty variance, and the data noise variance. The proposed approach is illustrated using geospatial data.

References

Efron B., and Tibshirani R. 1986, Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical accuracy, Statistical Science, vol. 1: 54-75.

Huang G.-B., Zhu Q.-Y., and Siew C.-K. 2006, Extreme learning machine: theory and applications, Neurocomputing, vol. 70: 489-501.