Geophysical Research Abstracts
Vol. 12, EGU2010-2018-3, 2010
EGU General Assembly 2010

Improving Efficiency in Training of Artificial Neural Networks using Information-rich Data

Shailesh Kumar Singh (1), Sharad K. Jain (2), and András Bárdossy (1)

(1) Stuttgart University, Institute of hydraulic Engineering, Stuttgart, Germany (shailesh.singh@iws.uni-stuttgart.de), (2) Department of Water Resource Development & Management, Indian Institute of Technology Roorkee , Roorkee-247667 (Uttarakhand) India

Artificial Neural Networks (ANNs) are classified as a data driven technique which implies that their learning improves as more and more training data are presented to them. This observation is based on the premise that a longer time series of training samples will contain more events of different types and hence the generalization ability of the ANN will improve. However, a longer time series of training samples need not necessarily contain more information. If there is considerable repetition of the same type of information, the ANN may not become 'wiser' and one may be just wasting computational efforts and time. This study assumes that there are segments in a long time series which contain large quantum of information. If an ANN is trained using these segments rather than the whole series, the training would be the same or better and it will result in saving computational time and efforts. Here, data depth function was used as tool for identification of critical segments in a time series. Data with low half-space depth are considered as unusual or critical event. Different ANN architectures were trained using the whole time series data and using the data of only critical segments (identified by data depth function). A comparison of the results shows that the performance of the ANNs is only slightly worse then using all data if one uses the selected critical periods only. The result of this study can be generalise for identification of critical events in a time series.