Hydrol. Earth Syst. Sci. Discuss., 4, S1022–S1024, 2007
www.hydrol-earth-syst-sci-discuss.net/4/S1022/2007/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Prediction of littoral drift with artificial neural networks" *by* A. K. Singh et al.

Anonymous Referee #1

Received and published: 21 September 2007

The manuscript addresses the issue of quantification of sand moving parallel to coast line. Such sand transport is known as littoral drift. The quantification of littoral drift is important for coastal, harbour and port engineering. The existing methods to predict littoral drift are empirical in nature. The authors present an ANN approach to predict the sand transport using measured data. The approach was applied with limited data to predict sand transport for a short stretch in the western coast of India. The authors report superiority of their approach over the existing methods.

The manuscript is conceptually poor and is not well written. Main comments are:

1. The number of data points used in the study is limited (circa 61 in training and 20 in testing). The complexity of the ANN model is too high for the limited available data. Determining the unknown weights for a 5-6-1 ANN model with 61 training data points

to provide a generalization capacity of the ANN seems to be ambitious. The authors have made the ANN model more complex by having another ANN model. With the introduction of another ANN model the number of unknown weights increases. Some references on ANN may be useful (eg Haykin, 1999).

2.Second ANN: The purpose of using the second ANN is not at all clear. The authors are using the measured littoral drift as the target output of the first ANN model. The target output of the second ANN model is also the same. The logic behind this selection may please be elaborated. Why the second model would "fine tune" the prediction of the first ANN model? Any available literature on this aspect may also be referred. In the ANN literature sometimes the error of a model is used in another model to "fine tune" the result of the first model, however, the authors are not following the same principle.

3. Selection of inputs: The authors made an exhaustive search for the input variables by iterating selection of new set of input variables, training a model and comparing the model's performance on the testing dataset containing 20 data points. Probably the authors also could have chosen the input variables based on the physics of the process and/or making a data analysis involving correlation analysis or average mutual information.

4. Data used: a) It will be useful to know whether the authors did any data analysis or not. The analysis of the data to detect any spurious data should also be reported. This is particularly important as the authors have also reported that collecting data "in fierce oceanic condition is a difficult task". b) More information on the data collection procedure may be useful. c) Due to the absence of any cross-validation dataset how did you prevent over-training? d) Statistics of the training and testing datasets will be useful. As the data is of limited size so data-plots could be provided. e) Results on the training dataset can additionally be presented.

5. Linear regression and non-linear regression models: These models do not have the same input variables as the ANN model. The reason may please be elaborated.

HESSD

4, S1022–S1024, 2007

6. The manuscript does not contain sufficient description of the physical process. Description of the existing formulae to compute littoral drift and their applicability elsewhere are not reported well.

7. Page 2501 line 21: if you know some variables are secondary why do you choose them particularly when the data is limited.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 4, 2497, 2007.