Sediment-carrying capacity prediction using extreme learning machine

Xianglong Wei1,2, Chang Li1, Wei Huang1,2, Jiyi Gu1, Jing Liu1, Mingcheng Zhu1,*

1State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China
2State Key Laboratory of Hydraulics and Mountain River Engineering, College of Water Resource and Hydropower, Sichuan University, Chengdu 610065, China
*Corresponding author’s e-mail: mchzhu@nhri.cn

Abstract. Sediment-carrying capacity prediction is an important part in the river sediments movement simulation. In this paper, a novel machine learning approach was proposed to predict the sediment-carrying capacity by using the Extreme Learning Machine (ELM). The evaluation indexes and the average errors of these prediction models were analysed in detail. The result shows that the ELM model has a good generalization performance for the sediment-carrying capacity prediction, and the ELM approach could be an effective tool for the sediment movement simulation.

1. Introduction
Sediment-carrying capacity prediction plays a key role in the river sediments movement simulation. Previous studies have showed that the traditional way to predict sediment-carrying capacity in the river was using the formulas built by the experiment data and the measured data in the field. Recent years, machine learning approaches were widely used in the areas of river dynamics and hydraulics, such as regression[1] and classification[2]. Some researchers proposed the machine learning approaches to predicting sediment-carrying capacity. A novel prediction model of sediment-carrying capacity using the BP neural network was proposed by Chen and Tang [3], the prediction data and the measured data showed a good agreement. Li, Xie, Zhang and Li [4] proposed a Support Vector Machine (SVM) model to predict the sediment-carrying capacity, the data used for training the SVM model was the same data used for training the BP model of Chen and Tang [3], and the result in the study of Li, Xie, Zhang and Li [4] showed that the SVM model had a better performance comparing with the BP model. The Generalized Regression Neural Network(GRNN) approach was also proposed[5]. While with the development of machine learning algorithms, more advanced and practical algorithms could be used for solving this problem, and up to now, there are no studies about the applications of Extreme Learning Machine (ELM) in the sediment-carrying capacity prediction have been published.

The Extreme Learning Machine (ELM) is a new and robust machine learning algorithm based on the Single-hidden Layer Feedforward Network (SLFN), which was firstly proposed by Huang, Zhu and Siew [6]. Previous studies have showed that the ELM could be used for classification[7-10] and regression[11-14], and showed a good generalization performance at extremely fast learning speed[15]. The purpose of this study is to develop a novel sediment-carrying capacity prediction model based on the ELM, which could accurately predict the sediment-carrying capacity. The data for training and testing the ELM model were the same data in the studies of[3, 4]. This is the first study on the application
of ELM in the sediment-carrying capacity prediction. Therefore, the findings in this study should make a contribution to the river sediments movement simulation.

The rest of this paper is organized as follows. The data and the method will be introduced in Section 2. It will then go on to discuss the application of the ELM approach in the sediment-carrying capacity prediction. The main findings of this paper will be summarized in Section 3.

2. Data and method

2.1. Data collection

The data used for training the ELM models in this paper was collected from the study of Chen and Tang [3], which were 30 sets of data for training and 4 sets of data for validation. The surface velocity, slope and settling velocity measured in the experiments were imported into the ELM models as the training data, the sediment-carrying capacity was imported into ELM models as the training label in the model training process. In the next section, the fundamental of the extreme learning machine will be introduced.

2.2. Extreme learning machine

The Extreme Learning Machine was proposed by Huang, Zhu and Siew [6], which was a new single-hidden layer feedforward network. The ELM algorithm was widely used in solving the regression and classification problems [16]. Considering a single-layer feed-forward neural network (SLFN) with \( n \) neurons in the input layer, \( l \) neurons in the hidden layer and \( m \) neurons in the output layer, the general structure of SLFN could be expressed as Figure 1:

![Figure 1. General structure of a single-layer feed-forward neural network](image)

The weight \( \mathbf{w} \) between the neurons in the input layer and the neurons in the hidden layer is randomly generated, so is the bias \( \mathbf{b} \) in the hidden layer. The weight between the neurons in the hidden layer and the neurons in the output layer could be expressed as \( \mathbf{\beta} \).

For the given training samples \( \mathbf{X} \) and the output matrix \( \mathbf{Y} \):

\[
\mathbf{X} = \begin{bmatrix}
  x_{11} & \cdots & x_{1Q} \\
  \vdots & \ddots & \vdots \\
  x_{n1} & \cdots & x_{nQ}
\end{bmatrix}_{n \times Q} \quad (1)
\]

\[
\mathbf{Y} = \begin{bmatrix}
  y_{11} & \cdots & y_{1Q} \\
  \vdots & \ddots & \vdots \\
  y_{m1} & \cdots & y_{mQ}
\end{bmatrix}_{m \times Q} \quad (2)
\]

Assuming that the activation function in the hidden layer was \( g(x) \), then the output \( \mathbf{T} \) is:

\[
\mathbf{T} = [t_1, t_2, \ldots, t_Q], \quad t_j = \begin{bmatrix}
  t_{1j} \\
  \vdots \\
  t_{mj}
\end{bmatrix} = \begin{bmatrix}
  \sum_{i=1}^{l} \beta_{ij} g(\mathbf{w}_i x_j + b_i) \\
  \vdots \\
  \sum_{i=1}^{l} \beta_{ij} g(\mathbf{w}_i x_j + b_i)
\end{bmatrix} \quad (j=1, 2, 3, \ldots, Q) \quad (3)
\]

where \( \mathbf{w}_i = [\omega_{i1}, \omega_{i2}, \ldots, \omega_{im}] \), \( x_j = [x_{1j}, x_{2j}, \ldots, x_{nj}]^T \).

The above equation can be rewritten in the following form:

\[
\mathbf{H} \mathbf{\beta} = \mathbf{T}'
\]

where \( \mathbf{T}' \) is the transposed matrix of \( \mathbf{T} \). \( \mathbf{H} \) is the hidden layer output matrix of the neural network, which is as follows:
The minimum norm least-squares solution of
$$\min_{\beta} \| H\beta - T' \|$$
is unique [6, 17], which is:
$$\hat{\beta} = H^+ T'$$  (6)
where $H^+$ is the Moore-Penrose generalized inverse of matrix of $H$.

The randomly generated weight $w$ and the bias $b$ are constant in the training process, the weight $\beta$ could be determined by Eq.(6), then the ELM model is established with the selected activation function.

In this paper, two ELM models with sigmoid function and sin function were established to predict the sediment-carrying capacity.

3. Results and discussion

3.1. The influence of the number of hidden neurons and the activation function

The activation function used in this paper were the sigmoid function and the sin function. In order to evaluate the predicted performance of these models, the Correlation Coefficient ($CC$) and the Index of agreement ($I$) are introduced as follows:

$$CC = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 \sum_{i=1}^{N}(y_i - \bar{y})^2}}$$  (7)

$$I = 1 - \frac{\sum_{i=1}^{N}(y_i - x_i)^2}{\sum_{i=1}^{N}[(y_i - \bar{y}) + (x_i - \bar{x})]^2}$$  (8)

where $x_i$ are the measured values, and their average is $\bar{x}$. $y_i$ are the predicted values, and their average is $\bar{y}$. $N$ is the number of observations.

The Correlation Coefficient and the Index of agreement is a number no more than 1, if these values were closer to 1, then the performances of the models were much better. The performances of ELM models with sigmoid function was shown in Figure 2. The hidden nodes of these models were set from 1 to 30. It can be seen in Figure 2, the evaluation indexes of these models were significantly increasing following the increase of the hidden neurons number. Except the model with 1 hidden neuron, the minimum $CC$ of the models was 0.68 and the minimum $I$ was 0.78. What stands out in Figure 2 is that, for an ELM model with a randomly selected number of hidden neurons from 11 to 30, the index values of $CC$ and $I$ were no less than 0.9, and this shows that the ELM algorithm has a good generalization performance for the sediment-carrying capacity prediction.

![Figure 2. Predicted performance of ELM models with sigmoid function.](image)
The predicted performances of ELM models with sin function were shown in Figure 3. The same findings could also be found in this figure. An increasing trend of CC and I could be found in Figure 3 following the increase of the hidden neurons number. For an ELM model with a randomly selected number of hidden neurons from 16 to 30, the index values of CC and I were no less than 0.9.

\[ \text{Figure 3. Predicted performance of ELM models with sin function.} \]

3.2. The predicted performance for the validation data
4 sets of data were set as the validation data for these ELM models. The error between the predicted sediment-carrying capacity and the measured sediment-carrying capacity of each data set was calculated, and the average error of the validation data of the ELM models was shown in Figure 4. The average errors of the ELM models were not steady with the increase of the hidden neurons. The maximum average error was 112.4%, and the minimum average error was 7.1%. The performance of the ELM models with the sigmoid function was much better than that of the ELM models with the sin function.

\[ \text{Figure 4. Predicted error of the ELM models with different hidden neurons.} \]

4. Conclusion
In this paper, a novel machine learning approach which is using the Extreme Learning Machine was proposed to predict the sediment-carrying capacity. The evaluation indexes of these models show that the ELM model have a good generalization performance for the sediment-carrying capacity prediction, and the average error of the ELM models with specific number of hidden neurons could be small enough for the sediment-carrying capacity prediction. The ELM approach could be an effective tool for the sediment movement simulation.
Acknowledgments
This research was funded by the Postgraduate Research & Practice Innovation Program of Jiangsu Province and the Monitoring and Research Project for the Soft Mattress on the Typical Reach of the 6m Deep-Water Channel Training Project on the Yangtze River from Wuhan to Anqing Reaches.

References
[1] Goldstein, E. B., Coco, G. (2014) A machine learning approach for the prediction of settling velocity. Water Resources Research, 50: 3595–3601.
[2] Wei, X., Lu, Y., Wang, Z., Liu, X., Mo, S. (2018) A Machine Learning Approach to Evaluating the Damage Level of Tooth-Shape Spur Dikes. Water, 10: 1680.
[3] Chen, X. B., Tang, H. W. (2004) Research on sediment-carrying capacity based on theory of BP neural network. Journal of Sediment Research, 1: 29-34.
[4] Li, Z. Z., Xie, Y. B., Zhang, J., Li, X. L. (2010) Sediment Carrying Capacity Prediction Based on Chaos Optimization Support Vector Machines. Proceedings of SPIE - The International Society for Optical Engineering, 7749: 380-382.
[5] Peng, J., Yao, W., Yan, J., Xiao, P., Shen, Z., Yang, C. (2012) Experimental study on model of sediment transport capacity of slope runoff based on GRNN. Science of Soil & Water Conservation, 10: 25-31.
[6] Huang, G. B., Zhu, Q. Y., Siew, C. K. Extreme learning machine: a new learning scheme of feedforward neural networks. In: IEEE International Joint Conference on Neural Networks, 2004. Proceedings. 985-990 vol.982.
[7] Rong, H. J., Ong, Y. S., Tan, A. H., Zhu, Z., Neucom, J. (2008) A fast pruned-extreme learning machine for classification problem. Neurocomputing, 72: 359-366.
[8] Huang, G. B., Ding, X., Zhou, H. (2010) Optimization method based extreme learning machine for classification. Neurocomputing, 74: 155-163.
[9] Mahesh, Aaron, E., WARNER, Timothy, A. (2013) Kernel-based extreme learning machine for remote-sensing image classification. Remote Sensing Letters, 4: 853-862.
[10] Wei, L., Chen, C., Su, H., Qian, D. (2015) Local Binary Patterns and Extreme Learning Machine for Hyperspectral Imagery Classification. IEEE Transactions on Geoscience & Remote Sensing, 53: 3681-3693.
[11] Cheng, L., Zeng, Z., Wei, Y., Tang, H. (2014) Ensemble of extreme learning machine for landslide displacement prediction based on time series analysis. Neural Computing & Applications, 24: 99-107.
[12] Abdullah, S. S., Malek, M. A., Abdullah, N. S., Kisi, O., Yap, K. S. (2015) Extreme Learning Machines: A new approach for prediction of reference evapotranspiration. Journal of Hydrology, 527: 184-195.
[13] Deo, R. C., Şahin, M. (2015) Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia. Atmospheric Research, 153: 512-525.
[14] Ćojbašić, Ž., Petković, D., Shamshirband, S., Tong, C. W., Ch, S., Janković, P., Dučić, N., Baralić, J. (2016) Surface roughness prediction by extreme learning machine constructed with abrasive water jet. Precision Engineering, 43: 86-92.
[15] Huang, G. B., Zhu, Q. Y., Siew, C. K. (2006) Extreme learning machine: Theory and applications. Neurocomputing, 70: 489-501.
[16] Huang, G., Huang, G. B., Song, S., You, K. (2015) Trends in extreme learning machines: A review. Neural Networks, 61: 32-48.
[17] Huang, G.-B., Zhu, Q.-Y., Siew, C.-K. (2006) Extreme learning machine: Theory and applications. Neurocomputing, 70: 489-501.