THE PREDICTION OF GROUNDWATER LEVEL ON TIDAL LOWLANDS RECLAMATION USING EXTREME LEARNING MACHINE

1,2 NURHAYATI, 2INDRATMO SOEKARNO, 1IWAN K. HADIHARDAJA, 2M. CAHYONO

1Civil Engineering Department, University of Tanjungpura, Pontianak, Indonesia
2Civil Engineering Department, Institut Teknologi Bandung, Bandung, Indonesia

E-mail: 1nurhayatiuntan@gmail.com, 2indratmosoekarno@ftsl.itb.ac.id, 2iwankridasantausa@yahoo.co.id, 2mcahyono2003@yahoo.uk

ABSTRACT

Groundwater level in the tidal lowlands fluctuates according to space and time. It is influenced by local rainfall or tidal. Tidal lowlands development for agriculture particularly food crops requires proper water management strategies so that the productivity of land and the production of food crops can be optimized. Proper water management must be supported by an accurate prediction system. This research aims to apply extreme learning machine (ELM) which can be used to make a prediction system of groundwater level in tidal lowlands. ELM is a feed forward artificial neural network with a single hidden layer or commonly referred to as single hidden layer feed forward neural networks (SLFNs). ELM has the advantage in learning speed. The result of the ground water level prediction using ELM was better than that using BPANN. Based on these results, the ELM can be used to predict the ground water level in order to assist decision makers in determining water management strategies and the determination of appropriate cropping patterns in the tidal lowlands reclamation.

Keywords: Prediction, Ground Water Level, Back Propagation, Artificial Neural Network, Extreme Learning Machine

1. INTRODUCTION

Water management is one of the keys to succeed in farming systems in tidal lowlands for agriculture particularly food crops. Water management in tidal lowlands faces some obstacles including water level fluctuations in farm land. Groundwater level information is obtained by making observations in the field but it takes time, cost and a lot of energy [1]. An accurate prediction system is required to obtain water level information.

There are many prediction methods used by the experts to get the correct prediction, one of which is a method based on artificial intelligence (AI). The most popular AI method used and the one that has been applied by experts [2], [3], [4] in this prediction is Back propagation of Artificial Neural Network (BPANN). BPANN has an advantage in control area, prediction and pattern recognition. Many researchers have concluded that BPANN is better than the conventional prediction methods [5].

AI method used in this research is the extreme learning machine (ELM). ELM is a new learning method of neural network. ELM is a feed forward neural network with a single hidden layer or commonly referred to as single hidden layer feed forward neural networks (SLFNs) [5], [6], [7].

ELM has the advantage in the learning speed. Hence, by applying the ELM, it is expected to produce an accurate and effective prediction.

It is expected that the outcomes of groundwater level prediction research using ELM are better than those using BPANN and they can be used appropriately in determining water management strategies and the determination of cropping patterns on tidal lowlands reclamation in order to increase agricultural land productivity and crop productivity.

2. ARTIFICIAL NEURAL NETWORK (ANN)

ANN is an information processing system on specific characteristics of the object associated with biological neural networks [4]. ANN is built based on the generalization of mathematical models on
human or biological neuron and it is based on the following assumptions: 1) Information processing occurs on many elements called neurons, 2) Signals run between neurons connected by network, 3) Each of connection network is associated with a weight, which is in a particular neural network, double the signal transmission, and 4) Each neuron uses activation function the (non-linear) to its input network to determine the output signal.

NN is characterized by: 1) The pattern is connected between neurons (called architecture), 2) Method to determine the weights of the connections (called training or learning algorithm of groundwater level), and 3) Activation function.

NN consists of a lot of processing elements called neurons, units, cells or nodes. Each neuron is connected to another by direct network communication tools and each is connected by weights. Weight portrays information used by the network to solve the problem. NN can be used to store, retrieve data or patterns repeatedly, classify pattern, perform general mapping from input patterns to output patterns, group similar patterns, or find an optimal solution for a problem.

Each neuron has an internal part called activation or activation level which function is to receive input. In particular, a neuron sends activation in the form of a signal to other neurons.

### 3. BACKPROPAGATION (BP)

BP is a descent gradient method to minimize the total square error of the output computed by the network. BP trains network to get a balance between the ability of the network to recognize the patterns used during the training process and the ability of the network to provide the correct response to the similar (but not identical) input pattern.

1). Architecture

BP has several units located in one or more hidden layers.

Figure 1 represents the architecture of BP with n inputs (plus a bias), a hidden layer consisting of p units (plus a bias), as well as m output units. Where $u_{ij}$ is the weight of a line from the input unit $A_i$ to hidden layer unit $B_j$ ($u_{i0}$ is the weight of the line connecting bias in the input to the hidden layer unit $B_j$). $v_{jk}$ is the weight from hidden layer unit $B_j$ to output unit $C_k$ ($v_{k0}$ is the weight of the line connecting bias in the hidden layer unit to the output unit $B_j$).

![Figure 1. Back Propagation Architecture](image)

2). Activation Function

Activation function for BP network has several important characteristics. Continuous activation function can be derived and it does not decrease monotonically. The derivative of the activation function is also easy to calculate for efficiency calculations. The most common activation function used is the value of derivatives (on the specific value of the independent variable) which can be expressed in terms of the value of the activation function (the value of the independent variable).

The activation function most often used is the binary sigmoid function while another function that is also quite often used is bipolar sigmoid function. Explanation of the two functions is given as follows:

a. A binary sigmoid function is for the interval $(0,1)$

$$f(x) = \frac{1}{1 + e^{-x}}$$

$$f'(x) = f(1 - f(x))$$

b. A bipolar sigmoid function is for the interval $(-1,1)$

$$f(x) = \frac{2}{1 + e^{-x}} - 1$$

$$f'(x) = \frac{1}{2} [1 + f_1(x)][1 - f_1(x)]$$

3). Algorithm of Back propagation water level

**Step 0** : Initialize weights of all weights with small random numbers

**Step 1** : If the termination condition is not met, do step 2-9

**Step 2** : For each pair of training data, do steps 3-8

**Step 3** : Each input unit receives signals and
Step 4: Calculate all output in hidden unit
Step 5: Calculate all output in output unit
Step 6: Calculate $\delta$ units of output based on the error in each output unit $C_k$
Step 7: Calculate $\delta$ hidden units based on error in each hidden unit $B_j$
Step 8: Calculate all the weight changes
Step 9: Test the termination conditions

4. EXTREME LEARNING MACHINE (ELM)

Extreme learning machine is a new learning method of neural network. ELM is a feedforward neural network with a single hidden layer or commonly referred to as single hidden layer feedforward neural networks (SLFNs) [5].

ELM learning method was created to overcome the weaknesses of feedforward artificial neural networks especially in terms of learning speed. There are two reasons why the feedforward artificial neural network has a low learning speed [5] [6] [7] [8] [9]: (1) Using a slow gradient based learning algorithm to do the training. All parameters on the network are determined iteratively by using that method of learning. (2) Learning using conventional gradient based learning algorithms, such as backpropagation (BP) and its variants, Lavenberg Marquadt (LM). All parameters of the feedforward artificial neural networks must be manually specified [6].

The parameters meant are the input weight and hidden bias. These parameters are also interconnected between one layer with another. Thus, it requires a long learning speed and it is often stuck in local minima [7], [9]. Input weight and hidden bias on the ELM are randomly selected so that ELM has a fast learning speed and it is able to produce good generalization performance. Figure 2 is the structure of the ELM.

ELM methods have a different mathematical model from feedforward artificial neural networks. The mathematical model of the ELM is simpler and more effective.

Mathematical model of the ELM to N number of different samples ($X_i, T_i$).

$$X_i = [X_{i1}, X_{i2}, \cdots, X_{in}]^T \in \mathbb{R}^n$$  \hspace{1cm} (5)

$$T_i = [T_{i1}, T_{i2}, \cdots, T_{im}]^T \in \mathbb{R}^m$$  \hspace{1cm} (6)

SLFNs standard with N number of hidden nodes and $g(x)$ activation function can be described mathematically as follows:

$$\sum_{i=1}^{N} \beta_i g(x_j) = \sum_{i=1}^{N} \beta_i [W_{i1}X_i + b_1] = o_j$$  \hspace{1cm} (7)

Where: $j = 1, 2, \ldots, N$; $W_i = (W_{i1}, W_{i2}, \ldots, W_{in})^T$ is the vector of weight connecting i to the hidden nodes and input nodes; $\beta_i = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{im})^T$ is the vector of weight connecting i to the hidden nodes and output nodes; $b_i$ = threshold of i to the hidden nodes; $W_iX_j$ = inner product of $W_i$ and $X_j$.

ELM with N hidden nodes and $g(x)$ activation function is assumed to be able to approximate with an error rate of 0 or it can be denoted as follows:

$$\sum_{j=1}^{N} \|o_j - t_j\| = 0 \quad \text{so that} \quad o_j = t_j$$  \hspace{1cm} (8)

$$\sum_{i=1}^{N} \beta_i g(W_{i1}X_j + b_i) = t_j$$  \hspace{1cm} (9)

Equation (9) can be written simply as

$$H\beta = T$$  \hspace{1cm} (10)

$$H = \begin{bmatrix} g(w_{11}x_1 + b_1) & \cdots & g(w_{1n}x_N + b_n) \\ \vdots & \ddots & \vdots \\ g(w_{n1}x_1 + b_1) & \cdots & g(w_{nn}x_N + b_n) \end{bmatrix} \in \mathbb{R}^{N \times N}$$  \hspace{1cm} (11)

$$H = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_N^T \end{bmatrix}$$  \hspace{1cm} (12)

$$T = \begin{bmatrix} T_1^T \\ \vdots \\ T_N^T \end{bmatrix}$$  \hspace{1cm} (13)

In equation (12) above is the hidden layer output matrix. $g(w_{ij}x_i + b_i)$ shows that the output of the hidden neurons related to the input of $x_i$, $\beta$ is the output weight matrix and $T$ matrix of the target or output.

Input weight and hidden bias in ELM are randomly determined. Then the output weight
associated with the hidden layer can be determined from equation (15).

$$\beta = H^T T$$ \hspace{1cm} (15)

5. IMPLEMENTATION OF THE PROPOSED METHOD TO THE SYSTEM

Several stages to go through in the ground water level prediction using ELM are as follows:

5.1. Data Collection and Data Processing

The types of data used in this research are secondary data. Data used include rainfall, evapotranspiration, water level in the canal, hydraulic conductivity, and drain spacing. The source of data research is taken from the project Land and Water Management of Tidal Lowlands (LWMTL) 2005 and Strengthening Tidal Lowlands Development (STLD) 2007 in the district of Banyuasin, South Sumatra, Indonesia from April 2006 to June 2008.

5.2. Data Distribution

Training and testing process are absolutely necessary in the prediction process using ELM. Training process was used to develop a model of the ELM while testing was used to evaluate the ability of ELM as forecasting tool. Therefore the data were divided into two, namely the training data and the testing data. Data were shared with the ratio of 70:30, i.e. 70% for training and 30% for testing.

5.3. ELM Training

ELM has to go through the training process first before it is being used as a tool to predict the groundwater level. The purpose of this process is to get input weight, bias and output weight with a low error rate.

5.4. Data Training Normalization

Data to be inputted to the ELM should be normalized so as to have a certain range of value. This is necessary because the activation function used will produce output with a range of data [0,1] or [-1,1]. Training data in this research are normalized so that they have the value range [-1,1]. The formula used in the normalization process.

$$X = \frac{2 \times (X_p - \min\{X_p\})}{\max\{X_p\} - \min\{X_p\}} - 1$$ \hspace{1cm} (16)

Where: $X$ = the value of the normalization result ranging between [-1,1]; $X_p$ = the value of the original data that have not been normalized; $\min\{X_p\}$ = minimum value in the data set, and $\max\{X_p\}$ = maximum value in the data set.
5.5. Determining Activation Function and The Number of Hidden Neuron

The number of hidden neurons and the activation function in the training process should be determined in advance. On this research, trials using the log sigmoid activation function. In addition, linear transfer function was also used since the data predicted were stationary. Linear transfer function has a weakness in the data pattern that has a trend [4].

5.6. Determining Weight Function, Bias of Hidden Neuron and Output Weight

The output of the ELM training process was the input and output weight as well as bias of the hidden neuron with a low error rate measured by the MSE and MAPE. Input weight was determined randomly while the output weight was the inverse of the matrix of hidden layer and output.

5.7. Denormalization Output

The output resulted from the training process was denormalized so that we obtained the predicted groundwater levels from training data. Denormalization formula used was:
\[
X = 0.5 \times \left( X_p + 1 \right) \times \left[ \max \{ X_p \} - \min \{ X_p \} \right] + \min \{ X_p \} + \min \{ X_p \}
\]
\[
(17)
\]
Where: \( X \) = the data value after denormalization; \( X_p \) = the output data before denormalization; \( \min \{ X_p \} \) = minimum data on the data sets before normalization; \( \max \{ X_p \} \) = maximum data on the data sets before normalization.

5.8. Testing ELM

Based on the input weight and output weight obtained from the training process, the next step is to predict the groundwater level. The input data were normalized in advance using the same range and the same normalization formula with the training data. Automatically, the output of this process should also be normalized.

5.9. The Analysis Prediction Results

After going through various stages as described above, the prediction value of groundwater level was obtained. The results were then analyzed whether they had small error rates (MSE and MAPE). If the resulting error rate is still relatively large, the steps that have been made will be re-evaluated, starting from the process of training and testing (predicting) until optimal results are obtained.

Mathematical formula of the mean square error (MSE) and mean absolute percentage error (MAPE).

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} e_i^2
\]
\[
(18)
\]
Where: \( e_i = X_i - F_i \)

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{PE_i}{X_i} \right|
\]
\[
(19)
\]
Where: \( PE_i = \left( \frac{X_i - F_i}{X_i} \right) \)

5.10. Comparing Prediction Results of BPANN and ELM

The values of MSE and MAPE would be obtained after the analysis of the results of ELM.
was done. The next step was to compare between the results of prediction of groundwater level with the result of the groundwater level as a result from the field observations. This was done to determine the accuracy level of the result of the groundwater level prediction using ELM with the groundwater level as a result from the observations in the field. Flowchart of training and prediction using ELM is shown in Figure 5.

6. RESULT AND ANALYSIS

The following are the prediction results of groundwater levels on tidal lowlands using BP-ANN and ELM. The activation function used on these two methods was sigmoid.

The data parameters used on BP-ANN are as follows:
- Epoch = 100
- Goal = 0.001
- Max performance incremental = 1.05
- Learning rate (Lr) = 0.01
- Lr incremental = 1.05
- Lr-decremental = 0.7
- Momentum = 0.9

The testing results (predictions) using ELM are as follows:
- Epoch = 100
- Goal = 0.001
- Max performance incremental = 1.05
- Learning rate (Lr) = 0.01
- Lr incremental = 1.05
- Lr-decremental = 0.7
- Momentum = 0.9

The testing results (predictions) using ELM are as follows:

Figure 6 shows that the groundwater level OT4.1 as a result from training using ELM has a relatively small error rate value. It meant that the result of the training was the same as the result of the observation. Therefore, it can be concluded that the training on groundwater level was successful. The values of the error rate as a result of the training were as follows: MSE = 0.00027466, and MAPE = 0.9670%.

Figure 7 shows that the groundwater level OT4.2 as a result from training using ELM has a relatively small error rate value. It meant that the result of the training was the same as the result of the observation. Therefore, it can be concluded that the training on groundwater level was successful. The values of the error rate as a result of the training were as follows: MSE = 0.000068223, and MAPE = 0.5516%.

Figure 8 shows that the groundwater level OT4.3 as a result from training using ELM has a relatively small error rate value. It meant that the result of the training was the same as the result of the observation. Therefore, it can be concluded that the training on groundwater level was successful. The values of the error rate as a result of the training were as follows: MSE = 0.00025623, and MAPE = 0.9045%.
Figure 9 shows that the groundwater level OT4.4 as a result from training using ELM has a relatively small error rate value. It meant that the result of the training was the same as the result of the observation. Therefore, it can be concluded that the training on groundwater level was successful. The values of the error rate as a result of the training were as follows: MSE = 0.000052688, and MAPE = 0.4461%.

Figure 10 shows that the groundwater level OT4.5 as a result from training using ELM has a relatively small error rate value. Therefore, it can be concluded that the training on groundwater level was successful. The values of the error rate as a result of the training were as follows: MSE = 0.00030391, and MAPE = 1.0206%.

Figure 11 shows that the groundwater level OT4.6 as a result from training using ELM has a relatively small error rate value. Therefore, it can be concluded that the training on groundwater level was successful. The values of the error rate as a result of the training were as follows: MSE = 0.000059577, and MAPE = 0.4905%.

The result of the groundwater level prediction using ELM was as follows:

Figure 12 shows that the groundwater level as a result from prediction on OT4.1 using ELM has a relatively small error rate value. Therefore, it can be concluded that the prediction on groundwater level was successful. The values of the error rate as a result of the prediction were as follows: MSE = 0.00034573, and MAPE = 1.0238%.
Figure 13 shows that the groundwater level as a result from prediction on OT4.2 using ELM has a relatively small error rate value. It meant that the result of the prediction was the same as the result of observation. Therefore, it can be concluded that the prediction on groundwater level was successful. The values of the error rate as a result of the prediction were as follows: MSE = 0.000068923, and MAPE = 0.4653%.

Figure 14 shows that the groundwater level as a result from prediction on OT4.3 using ELM has a relatively small error rate value. It meant that the result of the prediction was the same as the result of observation. Therefore, it can be concluded that the prediction on groundwater level was successful. The values of the error rate as a result of the prediction were as follows: MSE = 0.00031283, and MAPE = 0.9647%.

Figure 15 shows that the groundwater level as a result from prediction on OT4.4 using ELM has a relatively small error rate value. It meant that the result of the prediction was the same as the result of observation. Therefore, it can be concluded that the prediction on groundwater level was successful. The values of the error rate as a result of the prediction were as follows: MSE = 0.000070310, and MAPE = 0.4628%.

Figure 16 shows that the groundwater level as a result from prediction on OT4.5 using ELM has a relatively small error rate value. It meant that the result of the prediction was the same as the result of observation. Therefore, it can be concluded that the prediction on groundwater level was successful. The values of the error rate as a result of the prediction were as follows: MSE = 0.00031845, and MAPE = 0.9870%.
Figure 17 shows that the ground water level as a result from prediction on OT4.6 using ELM has a relatively small error rate value. It meant that the result of the prediction was the same as the result of observation. Therefore, it can be concluded that the prediction on groundwater level was successful.

The values of the error rate as a result of the prediction were as follows: MSE = 0.000074588, and MAPE = 0.5062%.

Comparison of the results of the groundwater level prediction using ELM and BP ANN can be seen in Table 1.

**Table 1. Comparison Of The Results Of The Ground Water Level Using ELM And BPANN For OT4.1, OT4.3, OT4.5**

<table>
<thead>
<tr>
<th>Simulation</th>
<th>BPANN OT4.1, OT4.3, OT4.5</th>
<th>ELM OT4.1, OT4.3, OT4.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training time</td>
<td>19.5156</td>
<td>2.7813</td>
</tr>
<tr>
<td>MSE training</td>
<td>0.00035595</td>
<td>0.0002726</td>
</tr>
<tr>
<td>MAPE training</td>
<td>1.0212</td>
<td>0.9640</td>
</tr>
<tr>
<td>Testing time</td>
<td>0.1406</td>
<td>0.0781</td>
</tr>
<tr>
<td>MSE testing</td>
<td>0.0012</td>
<td>0.00032567</td>
</tr>
<tr>
<td>MAPE testing</td>
<td>1.8600</td>
<td>0.9918</td>
</tr>
</tbody>
</table>

Table 1 and Table 2 show that the comparison of the level of training error and prediction error on groundwater level prediction using ELM is smaller than that using BPANN. The time of the training process and the prediction process on groundwater level using ELM is faster than that using BPANN.

Table 1 and Table 2 show that the result of training and prediction using ELM and BP ANN is able to work well in recognizing input data given to the system because the error rate is relatively small. The result of training and prediction using ELM is better than those using BPANN.

### 7. CONCLUSION

ELM proposed in this paper is used to predict the ground water level in the tidal lowlands reclamation in Indonesia. The method used in this paper is ELM while BPANN is used as validation. The result shows that the training result and the groundwater level prediction using ELM are better than those using BPANN.

### ACKNOWLEDGEMENTS

The authors would like to thank the Government of Indonesia, especially Directorate General of Higher Education for Postgraduate Scholarship (BPPS), University Tanjungpura for University Decentralization Research on Doctoral Dissertation Grant during research that we receive, Water Resources Engineering Laboratory, Civil Engineering Faculty, Institut Teknologi Bandung (ITB), Bandung Indonesia for all the facilities provided for this research. In addition, the authors would like to express their gratitude to the project Land and Water Management of Tidal Lowlands (LWMTL) 2005 and Strengthening Tidal Lowlands Development (STLD) 2007 in the district of Banyuasin, South Sumatra, Indonesia for the willingness of giving out the data and permission for publication.

### REFERENCES:


[2] Jianhui Wu, Qi Ren, Yu Su, Sufeng Yin, Houjun Xu, Guoli Wang, Study and Application of Combination Prediction Model


