An expert system with radial basis function neural network based on decision trees for predicting sediment transport in sewers
Isa Ebtehaj, Hossein Bonakdari and Amir Hossein Zaji

ABSTRACT
In this study, an expert system with a radial basis function neural network (RBF-NN) based on decision trees (DT) is designed to predict sediment transport in sewer pipes at the limit of deposition. First, sensitivity analysis is carried out to investigate the effect of each parameter on predicting the densimetric Froude number \( (Fr) \). The results indicate that utilizing the ratio of the median particle diameter to pipe diameter \( (d/D) \), ratio of median particle diameter to hydraulic radius \( (d/R) \) and volumetric sediment concentration \( (CV) \) as the input combination leads to the best \( Fr \) prediction. Subsequently, the new hybrid DT-RBF method is presented. The results of DT-RBF are compared with RBF and RBF-particle swarm optimization (PSO), which uses PSO for RBF training. It appears that DT-RBF is more accurate \( (R^2 = 0.934, \text{MARE} = 0.103, \text{RMSE} = 0.527, \text{SI} = 0.13, \text{BIAS} = -0.071) \) than the two other RBF methods. Moreover, the proposed DT-RBF model offers explicit expressions for use by practicing engineers.

Key words | bed load, decision trees (DT), limit of deposition, pipe channel, radial basis function (RBF), sediment transport

INTRODUCTION
Optimized economic performance of sewer systems is crucial in urban drainage engineering. One of the mechanisms affecting sewer system operation is sediment transport. Due to washed-out solid matter along the flow path, the flow through sewer systems often contains suspended solids. If the flow velocity is not sufficiently high to transport sediments without deposition, sedimentation occurs. The presence of sediments deposited at the bottom of pipes causes greater pipe roughness and lower cross-sectional flow area, which then leads to reduced transport capacity. The simplest method of designing sewer systems is to use minimum velocity or minimum shear stress. The exact recommended values for shear stress and velocity were presented by Vongvisessomjai et al. (2010). Inefficient methods do not consider hydraulic parameters and channel characteristics, which have significant impact on determining the minimum velocity. Hence, sediments deposit on the bottom of pipes led to several problems, such as combined sewer overflow.

Numerous experimental studies have been conducted on sediment transport analysis and understanding the parameters affecting sediment transport. Consequently, a number of regression-based equations have been proposed (Novak & Nalluri 1975; May et al. 1989; Ab Ghani 1993; Ackers et al. 1996; Ota & Nalluri 2003; Banasiak 2008; Almedeij & Almohsen 2010; Almedeij 2012; Ota & Perrusquía 2013; Bonakdari & Ebtehaj 2014). Considering the many equations proposed to determine the minimum velocity using linear regression, as well as the method’s weakness in nonlinear systems, using these equations may cause higher or lower velocity estimation, thus leading to uneconomical or inefficient design, respectively (Ebtehaj et al. 2014). Owing to the high capacity of artificial intelligence (AI) to solve various problems in complex and non-linear systems, AI methods have been extensively used in the hydraulic and hydrology fields, as well as sediment transport (Ebtehaj & Bonakdari 2015, 2014a). Ebtehaj & Bonakdari (2016) surveyed the performance of three different gradient algorithms, namely resilient backpropagation, variable learning rate and Levenberg–Marquardt, in training a multi-layer perceptron neural network (MLP-NN) to predict sediment transport in clean pipes. The authors found that
A new hybrid technique comprising an radial basis function neural network (RBF-NN) based on decision trees (DT-RBF) is proposed in this study. The model’s performance in predicting the Froude number is compared with the respective RBF and RBF-particle swarm optimization (PSO) performance. The superiority of the RBF-PSO over the simple RBF lies in using an evolutionary algorithm to determine the most appropriate RBF adjustments. However, DT-RBF has a different procedure. In this method, the DT classification is combined with the RBF in order to reallocate the RBF power to suitable data subsets. The functional equation suggested by Ebtehaj & Bonakdari (2014b) as the best model for predicting the Fr and RBF neural networks, is utilized to carry out sensitivity analysis and to understand the parameters’ effect. The best model is ultimately used to forecast the Fr using DT-RBF.

Data presentation

Field and experimental studies on sediment transport in sewer pipes (May et al. 1980; Ab Ghani 1993; Vongvisessomjai et al. 2010) indicate that the minimum velocity required to prevent solid matter deposition at the bottom of pipes depends on pipe diameter (D), flow depth (y) or hydraulic radius (R), specific gravity of sediments (s = ρs/ρ), density of water (ρ), density of sediments (ρs), gravitational acceleration (g), median diameter of particles (d), overall friction factor of sediment (λ) and volumetric sediment concentration (CV). Therefore, the functional relationship between the minimum velocity and parameters affecting it is considered as follows:

\[
V = f(C_V, y, R, D, d, s, g, \lambda_s)
\]  

(1)

Different models including movement (Fr = V/\sqrt{(g(s−1)/d)}, transport (CV), sediment (Dgr = d/(g(s−1)/\nu^2)^{1/3}), d/D, where \( \nu \) is the kinematic viscosity of fluid and Dgr is the dimensionless particle number, transport mode (d/R, D^2/A, R/D) and flow resistance (\( \lambda_s \)) are presented through different dimensionless parameters and classifying them into dimensionless groups. They are used to predict the minimum velocity required to prevent sediment deposition in sewer pipes (Ebtehaj & Bonakdari 2014a, 2014b). Ebtehaj & Bonakdari (2014b) presented different models by considering the effect of four groups, ‘sediment,’ ‘transport,’ ‘transport mode,’ and ‘flow resistance’ on determining the minimum velocity that is expressed as Fr (the movement group). Regarding the fact that the ‘transport’ (CV) and ‘flow resistance’ (\( \lambda_s \)) groups contain only one parameter, these two parameters were considered constant in all groups. In total, six different models are presented as follows, where three parameters are in the ‘transport mode’ group (d/R, D^2/A, R/D) and two are in the ‘sediment’ group (Dgr, d/D).

Model 1: 
Fr = f (CV, Dgr, d/R, \( \lambda_s \))

Model 2: 
Fr = f (CV, Dgr, D^2/A, \( \lambda_s \))

Model 3: 
Fr = f (CV, Dgr, R/D, \( \lambda_s \))

Model 4: 
Fr = f (CV, d/D, d/R, \( \lambda_s \))

Model 5: 
Fr = f (CV, d/D, D^2/A, \( \lambda_s \))

Model 6: 
Fr = f (CV, d/D, R/D, \( \lambda_s \))

A recent study by Ebtehaj & Bonakdari (2014b) showed that using the d/R parameter as part of the ‘transport mode’ group led to the finest performance compared with D^2/A and R/D. This is displayed in Models 1 and 4. Using the d/D parameter in the ‘sediment’ group (Model 4) led to a more accurate model than the other models. Thus, according to the fixed parameters CV and \( \lambda_s \), d/D, parameters (sediment group) and d/R (transport mode group) can be considered the most effective choice of parameters.

Model 4 (Fr = f (CV, d/D, d/R, \( \lambda_s \))) performed best among all above-mentioned models (Ebtehaj & Bonakdari 2014b). Therefore, sensitivity analysis was conducted using RBF-NN in this study in order to examine the effect of each dimensionless parameter presented in Model 4. The following models were obtained based on Model 4. After using RBF-NN to select the best model amongst 4A to 4D, Fr was predicted using DT-RBF.

Model 4-A: 
Fr = f (CV, d/D, d/R, \( \lambda_s \))

Model 4-B: 
Fr = f (CV, d/D, d/R)

Model 4-C: 
Fr = f (CV, d/D, \( \lambda_s \))

Model 4-D: 
Fr = f (CV, d/R, \( \lambda_s \))

Model 4-E: 
Fr = f (d/D, d/R, \( \lambda_s \))

The sediment transport at limit of deposition was obtained from a dataset of 218 samples collected from studies by Ab Ghani (1993), Ota & Nalluri (1999) and Vongvisessomjai et al. (2010), which were conducted in open channel condition for storm sewers. The datasets were collected in different investigation circumstances, with varying pipe diameters and lengths, and a wide range of hydraulic conditions.

In this study, the entire dataset was divided into two groups: training and testing. Thirty percent of the whole
dataset was selected using random sampling for testing and the rest of the dataset was used for training.

To study the sediment transfer at limit of deposition, Ab Ghani (1995) conducted experimental tests with three pipe diameters (0.154, 0.305 and 0.450 m) and 20.5 m length. The minimal and maximal pipe diameters (0.154 and 0.45 m) were exploited using a smooth bed but the pipe with 0.305 m diameter was employed for smooth and rough beds. Ota & Nalluri (1999) employed 18 m long pipes with 0.305 m diameter and considered sediment gradation. The authors performed 24 experimental tests at limit of deposition in uniform \( (d = 0.71–5.61 \text{ mm}) \) and non-uniform \( (d = 2 \text{ mm}) \) states. Vongvisessomjai et al. (2010) performed different tests on two pipes with 0.1 and 0.15 m diameters and 16 m long. The authors used diverse slopes (0.002, 0.004, 0.006) and sediments \( (d = 0.2, 0.3, 0.43 \text{ mm}) \). The mean velocity was computed as the average of velocities close to the bed, at intermediate depth and at the current surface. The hydraulic parameter ranges were: \( 0.237 < V \text{ (m/s)} < 1.216; 1 < C_v \text{ (ppm)} < 1.280; 0.072 < d \text{ (mm)} < 8.3; 0.005 < R \text{ (m)} < 0.136; 0.153 < y/D < 0.84 \text{ and } 0.1 < D \text{ (m)} < 0.45. \)

**METHODS**

**Hybrid DT-RBF method**

The hybrid DT-RBF method introduced in this study comprises an RBF-NN regression method and DT classification algorithm.

RBF is constructed from three layers. The first layer is called the input layer, it sets the input variables to the model and transfers them to the next layer. The second layer is identified as the hidden layer, which lessens the experiments’ dimensionality using non-linear input variables. This projection is done by real-value functions called RBFs (Buhmann 2003). Following hidden layer projection, the hidden neurons are transferred to the output layer. The output layer evaluates the RBF-NN results by using a linear regressor. More details on the RBF-NN are given by Poggio & Girosi (1990). Setting up an optimum RBF-NN model requires determining the correct values of two parameters: the number of hidden layer neurons and the spread amount. These values are commonly determined through the trial and error method (Kisi 2008).

The DT-RBF method uses the DT algorithm (Breiman et al. 1993) to increase the RBF-NN performance by optimizing the regression power allocation. DT-RBF classifies the dataset using the input variables and then uses the small RBF-NN models on the divided dataset. The flowchart of the DT-RBF method is shown in Box 1, which illustrates the following procedure being implemented in the DT-RBF method: first, the DT algorithm is trained with the training dataset. In this study, the DT is programmed to divide the entire dataset into four classes. The important objective in the training procedure is to evaluate the classification accuracy. High classification accuracy may result in a large, impractical tree as well as an overfitting problem. Therefore, various classification error percentages should be tried to find the most suitable one. In the present study, the classification accuracy is controlled using the minimum parent size amount that is adjusted by the first for loop. After that, the dataset of each class is provided using the second for loop. The RBF-NN is split into four smaller models in order to model each DT class. Selecting the maximum allowable number of hidden neurons is done using trial and error to find an accurate RBF model and avoid overfitting. Overfitting occurs when the RBF’s performance in the training dataset is far better than for the testing dataset. Here, the maximum number of primal RBF-NN hidden neurons is

---

**Box 1 | DT-RBF algorithm**

**Starts**

- for MPS = Initial to max
  - Train the DT
- for Class = 1 to maximum number of classes
  - Provide the dataset for the present class
  - for RBF hidden neurons = Initial to max
  - for Spread = Initial to max
  - Train the RBF
  - Spread = Spread + Δ Spread
  - Hidden neurons = Hidden neurons + Δ Hidden neurons
- Class = Class + 1
- MPS = MPS + Δ MPS

**End**

- Save the optimum model that has the minimum error
- Combine the results of the classes’ RBF

**Finish**
12. Therefore, the sum of the maximum allowable number of smaller RBF-NNs is considered 12. The third and fourth for loops are used and the trial and error method is applied to determine the optimum number of hidden layer neurons and the spread amount of each developed RBF-NN model. Finally, the results of the class-based RBF-NN obtained are collected in order to acquire the final DT-RBF results.

PSO

The PSO algorithm is a population-based optimization technique inspired by the social behavior of birds. The social behavior of particles is modeled using a d-dimensional space with two properties: position and velocity. The PSO search starts with a random population. This initial population repeatedly moves in the d-dimensional search space to find new solutions. For each particle that has two specification vectors, i.e. position \((x_i)\) and velocity \((V_i)\), the fitness function \((f)\) is calculated to select the best particle. The best position of each particle until the current moment and its \(j^{th}\) value are kept at \(P_i = (p_{i1}, p_{i2}, \ldots, p_{id})\) and \(P_{gi}\), respectively. The duration repeated time \((t)\), and the velocity and position updating are done using Equations (2) and (3) (respectively) as follows:

\[
\begin{align*}
V_{ij}(t + 1) &= wV_{ij}(t) + c_1r_1(p_{ij}(t) - x_{ij}(t)) + c_2r_2(p_{gi}(t) - x_{ij}(t)) \\
x_{ij}(t + 1) &= x_{ij}(t) + V_{ij}(t + 1)
\end{align*}
\]

where \(w\) is the inertia coefficient, \(r_1\) and \(r_2\) are random numbers in \([0,1]\) and \(c_1\) and \(c_2\) are weight factors.

RESULTS AND DISCUSSION

The performance of the models presented in this study, RBF and DT-RBF, is evaluated in this section with different statistical indexes: R-square \((R^2)\), mean absolute relative error \((MARE)\), root mean square error \((RMSE)\), scatter index \((SI)\) and BIAS:

\[
R^2 = \frac{\sum_{i=1}^{n} (Fr_{observed i} - Fr_{predicted i}) (Fr_{predicted i} - Fr_{observed i})}{\left(\sum_{i=1}^{n} (Fr_{observed i} - Fr_{observed i})^2 \sum_{i=1}^{n} (Fr_{predicted i} - Fr_{predicted i})^2\right)^{\frac{1}{2}}}
\]

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Fr_{observed i} - Fr_{predicted i})^2}
\]

\[
MARE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Fr_{observed i} - Fr_{predicted i}}{Fr_{observed i}} \right|
\]

\[
SI = \frac{RMSE}{Fr_{observed i}}
\]

\[
BIAS = \frac{1}{n} \sum_{i=1}^{n} \left( Fr_{observed i} - Fr_{predicted i} \right)
\]

where \(Fr_{observed i}\) and \(Fr_{predicted i}\) are the observed and predicted \(Fr\) for the \(i^{th}\) sample (respectively). Table 1

<table>
<thead>
<tr>
<th></th>
<th>Train</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RBF</td>
<td>Model 4-A</td>
<td>0.850</td>
<td>0.130</td>
<td>0.749</td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td>Model 4-B</td>
<td>0.872</td>
<td>0.112</td>
<td>0.692</td>
<td>0.171</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Model 4-C</td>
<td>0.819</td>
<td>0.147</td>
<td>0.822</td>
<td>0.203</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Model 4-D</td>
<td>0.859</td>
<td>0.126</td>
<td>0.727</td>
<td>0.179</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Model 4-E</td>
<td>0.790</td>
<td>0.177</td>
<td>0.886</td>
<td>0.218</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>RBF-PSO</td>
<td>0.837</td>
<td>0.092</td>
<td>1.081</td>
<td>0.281</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>DT-RBF</td>
<td>0.943</td>
<td>0.090</td>
<td>0.461</td>
<td>0.114</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RBF</td>
<td>Model 4-A</td>
<td>0.807</td>
<td>0.171</td>
<td>0.918</td>
<td>0.227</td>
</tr>
<tr>
<td></td>
<td>Model 4-B</td>
<td>0.842</td>
<td>0.128</td>
<td>0.842</td>
<td>0.208</td>
<td>-0.176</td>
</tr>
<tr>
<td></td>
<td>Model 4-C</td>
<td>0.767</td>
<td>0.186</td>
<td>0.993</td>
<td>0.245</td>
<td>-0.140</td>
</tr>
<tr>
<td></td>
<td>Model 4-D</td>
<td>0.815</td>
<td>0.165</td>
<td>0.903</td>
<td>0.223</td>
<td>-0.217</td>
</tr>
<tr>
<td></td>
<td>Model 4-E</td>
<td>0.668</td>
<td>0.216</td>
<td>1.176</td>
<td>0.291</td>
<td>-0.090</td>
</tr>
<tr>
<td></td>
<td>RBF-PSO</td>
<td>0.827</td>
<td>0.119</td>
<td>1.612</td>
<td>0.361</td>
<td>-0.127</td>
</tr>
<tr>
<td></td>
<td>DT-RBF</td>
<td>0.934</td>
<td>0.103</td>
<td>0.527</td>
<td>0.130</td>
<td>-0.071</td>
</tr>
</tbody>
</table>
represents the performance of the three different RBF neural networks, i.e. RBF (Models 4A to 4E), DT-RBF (Model 4B) and RBF-PSO (Model 4B), in predicting the Fr for sediment transport without deposition. The effect of each of the four parameters presented in Model 4 selected by Ebtehaj & Bonakdari (2014b) as the best model, is initially examined through sensitivity analysis in this table using RBF (Models 4A–4E). The best model is then selected and will also be modeled through DT-RBF. This table indicates that not using the parameter related to the ‘flow resistance’ ($\lambda_s$) group increases Fr prediction by RBF. Except for the SI index, which is equal for all models, the other indexes displayed in Table 1 are better for all models. The comparison between model testing and training, indicate that model 4B ($R^2 = 0.842$) performs better than model 4A ($R^2 = 0.807$). Therefore, using $Fr = f (C_v, d/D, d/R)$ as the input combination for predicting the Fr results in a decreasing $MARE$ value from 0.13 (Model 4A) to 0.112 (Model 4B) in training and from 0.171 (Model 4A) to 0.128 (Model 4B) in testing mode. Unlike model 4A ($R^2 = 0.807$), which considers the $d/R$ parameter from the ‘transport mode’ dimensionless group an influential parameter in predicting $Fr$ using RBF, Model 4C ($R^2 = 0.767$) does not. This decreases $Fr$ prediction accuracy, such that the $MARE$ index increases by approximately 6% for Model 4C compared with Model 4A. Table 1 indicates that the $RMSE$ and $SI$ index values that present the root mean square error as absolute and relative (respectively), increased for Model 4C compared with the reference model (Model 4A). The $d/R$ parameter (transport mode group) is known to affect $Fr$ prediction, and not using it decreases the RBF performance in predicting the $Fr$. Not considering $d/D$ (sediment group) as a parameter affecting $Fr$ prediction decreases the modeling performance, similar to the $d/R$ parameter. Its effect is, however, slighter such that the values of most indexes presented in Table 1 are better for Model 4D ($R^2 = 0.815$) than Model 4C ($R^2 = 0.767$), indicating the superiority of Model 4D over Model 4C. It could be reasoned that for all 218 samples used in this study, the change in pipe diameter is insignificant in relation to the change in hydraulic radius ($R$). Also, using the hydraulic radius parameter takes into account the effect of both flow depth and pipe diameter factors simultaneously, while $D$ only considers the pipe diameter. The comparison between Model 4A and Model 4E indicates that volumetric sediment concentration ($C_v$) significantly affects $Fr$ prediction, as this model ($R^2 = 0.668$, $MARE = 0.216$, $RMSE = 1.176$, $SI = 0.291$, $BIAS = -0.09$) performed the weakest among all models. Therefore, the $C_v$ parameter, which considers the direct effect of sediments in flow on their transport, is essential for predicting the $Fr$.

With regard to the presented explanations, the $\lambda_s$ parameter has the least effect on predicting the $Fr$ (Model 4) among the four parameters; thus, not considering this parameter as part of the input combination increases the model’s performance (Model 4B). The $C_v$ parameter has the greatest effect on $Fr$ prediction. Therefore, Model 4B, which performs best among all models predicted using RBF, is predicted using a hybrid of RBF based on DT (DT-RBF).

Figure 1 shows the $Fr$ values predicted using the RBF and DT-RBF methods in both testing and training stages. RBF presents a similar process in both testing and training modes of the model, in that it overestimates the majority with a large mean relative error. Some of the predictions are underestimated as the $Fr$ value increases. RBF always predicts the $Fr$ with errors greater than 10%, while 40% of the predicted values have a relative error greater than 10% as seen in Figure 2. DT-RBF resolves most problems with RBF, whereby the majority of predicted values have a relative error less than 10% at high $Fr$. Although it is less accurate in terms of small $Fr$, it obviously presents better results for both testing and training modes of the models than RBF.

Table 1 serves to quantitatively examine the RBF and DT-RBF methods. According to this table, DT-RBF predicts the $Fr$ in both testing and training modes with a relative error of almost 10%, while the value of this index is approximately 12.8% for RBF in the test mode. The $SI$ index, which is the $RMSE$ value that has been made dimensionless, indicates that the root mean square error is equal to 0.208 for RBF ($SI = 0.208$) – more than 1.5 times the value of this index for DT-RBF ($SI = 0.13$). Figure 2 shows that the DT-RBF predicted 50% of the data with less than 6% relative error, while RBF made only 40% of the predictions with less than 6% relative error. Moreover, the largest relative error regarding $Fr$ prediction was approximately 80% when made by RBF, while this value was less than 30% for DT-RBF. Therefore, it is clear that using DT and combining it with RBF (DT-RBF) significantly reduces the weakness of RBF in predicting the $Fr$. The results obtained from predicting the $Fr$ with DT-RBF are compared with an RBF neural network that was improved for predicting the $Fr$ using a PSO algorithm (RBF-PSO). This comparison is presented in Table 1 and Figure 2. Table 1 shows that with a mean relative error of approximately 12%, RBF-PSO is less accurate than DT-RBF with a mean relative error of almost 10%. The $RMSE$ index is equal to 1.612 for RBF-PSO, which is nearly three times the value presented for DT-RBF ($RMSE = 0.527$). The $BIAS$ index also indicates that the values predicted by DT-RBF were estimated by an average of 0.07 less than
the actual values. This is a relatively small value, while RBF-PSO predicted the $Fr$ with an average of 0.127 less than its actual value. The maximum $Fr$ prediction error by RBF-PSO was nearly 40% while the maximum value predicted by DT-RBF had less than 30% relative error.

P11 L59: instead of ‘…by presenting the following equation using DT-RBF. According to the manuscript, the proposed DT-RBF method is used to …’ should read ‘…by presenting the equation using DT-RBF, which is used to …’

It is evident with regard to the presented explanations that in addition to significantly reducing the RBF weaknesses in $Fr$ prediction, the method proposed in this study (DT-RBF) outperforms the RBF-PSO hybrid (an evolutionary algorithm). Another flaw of RBF and RBF-PSO is that they do not present a specific equation for design engineers to use in practice. This flaw has been eliminated in this work by presenting the equation using DT-RBF, which is used to divide the whole dataset into four smaller datasets. Accordingly, the main RBF is divided into smaller datasets. As such, four smaller equations should be obtained in order to describe

**Figure 1** | Scatter plot of RBF and DT-RBF in the training and testing stages.

**Figure 2** | Error distribution of the proposed method (DT-RBF), RBF-PSO (Ebtehaj & Bonakdari 2014b) and RBF in predicting $Fr$ at the limit of deposition.
the model. Equation (9) represents the main RBF neural network equation that uses the \( \mathbf{IN} \) matrix (Equation (9.1)) as input variables. It is obvious that the other matrices of \( \mathbf{LW} \), \( \mathbf{IW} \), \( b_1 \), and \( b_2 \) are different for each model. Hence, these matrices are defined in Equations (10) to (13) for each class.

\[
\mathbf{u^*} = \mathbf{LW} \times \exp \left( -(\sum ((\mathbf{IW} - \mathbf{IN})^{0.5} \times b_1)^2 + b_2 \right) 
\]

\[
\mathbf{IN} = \begin{bmatrix} \mathbf{C} & d/D & d/R \end{bmatrix} 
\]

(9.1)

If \( \mathbf{C}_V < 4.195e - 5 \) and \( \frac{d}{R} < 0.0282 \)

\[
\mathbf{LW} = [6.578 \ 0.109 \ 0.023 \ -6.71] \times 10^8 
\]

(10.1)

\[
\mathbf{IW} = \begin{bmatrix} 0 & 0.0063 & 0.0016 \\ 0 & 0.0071 & 0.0013 \\ 0 & 0.0259 & 0.0029 \\ 0 & 0.0064 & 0.0016 \end{bmatrix} 
\]

(10.2)

\( b_1 = [1.8501] \)

(10.3)

\( b_2 = [-614.4515] \)

(10.4)

If \( \mathbf{C}_V < 4.195e - 5 \) and \( \frac{d}{R} > 0.0282 \)

\[
\mathbf{LW} = [7.5914 \ -8.2196] \times 10^5 
\]

(11.1)

\[
\mathbf{IW} = \begin{bmatrix} 0 & 0.0299 & 0.0089 \\ 0 & 0.0351 & 0.0089 \end{bmatrix} 
\]

(11.2)

\( b_1 = [0.0833] \)

(11.3)

\( b_2 = [62825] \)

(11.4)

If \( \mathbf{C}_V > = 4.195e - 5 \) and \( \frac{d}{R} < 0.1058 \)

\[
\mathbf{LW} = [-0.0011 \ -0.0003 \ 7.8005 \ -7.7996] \times 10^5 
\]

(12.1)

\[
\mathbf{IW} = \begin{bmatrix} 0.0001 & 0.0176 & 0.0030 \\ 0.0004 & 0.1053 & 0.0273 \\ 0.0001 & 0.0167 & 0.0020 \\ 0 & 0.0167 & 0.0020 \end{bmatrix} 
\]

(12.2)

\( b_1 = [16.6511] \)

(12.3)

\( b_2 = [33.5993] \)

(12.4)

If \( \mathbf{C}_V > = 4.195e - 5 \) and \( \frac{d}{R} > = 0.1058 \)

\[
\mathbf{LW} = [46.0562 \ -45.1797] 
\]

(13.1)

\[
\mathbf{IW} = \begin{bmatrix} 0.0002 & 0.1063 & 0.0187 \\ 0.0001 & 0.1102 & 0.0272 \end{bmatrix} 
\]

(13.2)

\( b_1 = [3.469] \)

(13.3)

\( b_2 = [1.6865] \)

(13.4)

**CONCLUSION**

Sediment transport in sewer systems was predicted using a hybrid RBF based on decision trees (DT-RBF) in this study. This method was compared with other existing methods. First, the functional relation presented by Ebtehaj & Bonakdari (2014b) was used and the parameters affecting the prediction of the minimum velocity required to prevent sediment deposition \((\mathbf{Fr})\) was considered as \( \mathbf{Fr} = f(CV, d/D, d/R, \lambda_s) \). The effect of each parameter on \( \mathbf{Fr} \) prediction was evaluated with sensitivity analysis and RBF neural networks. The results indicate that the \( \mathbf{C}_V \) parameter had the most significant role in predicting \( \mathbf{Fr} \) and not using it as an input parameter in the network would dramatically decrease the model’s prediction accuracy. In contrast, not considering the \( \lambda_s \) parameter as an effective parameter in predicting \( \mathbf{Fr} \) leads to enhanced model performance. Therefore, the best model obtained from the sensitivity analysis is \( \mathbf{Fr} = f(CV, d/D, d/R) \) \((R^2 = 0.842, \text{MARE} = 0.128, \text{RMSE} = 0.842, SI = 0.208, BIAS = -0.176) \). Subsequently, because RBF is not characterized by high accuracy, the minimum \( \mathbf{Fr} \) needed to prevent sediment deposition on channel beds was predicted using DT-RBF and the input combination obtained from sensitivity analysis (Model 4B). According to the results, DT-RBF \((R^2 = 0.934, \text{MARE} = 0.103, \text{RMSE} = 0.527, SI = 0.13, BIAS = -0.07) \) outperformed RBF. The DT-RBF performance was also compared with that of RBF-PSO and the results demonstrated that using the RBF and DT hybrid is much better than using the RBF and PSO hybrid.
REFERENCES


Ota, J. J. & Nalluri, C. 1999 Graded Sediment Transport at Limit Deposition in Clean Pipe Channel. 28th International Association for Hydro-Environmental Engineering and Research, Graz, Austria.


First received 22 April 2015; accepted in revised form 30 March 2016. Available online 22 April 2016.