Modelling Sediment Trapping by Non-Submerged Grass Buffer Strips Using Nonparametric Supervised Learning Technique

S. Akram[1,*], B. Yu[2], H. Ghadiri[1], and C. Rose[3]

[1]*Environmental Futures Research Institute, School of Environment, Griffith University, Nathan QLD 4111, Australia*
[2]*School of Engineering, Griffith University, Nathan QLD 4111, Australia*
[3]*Australian Rivers Institute, School of Environment, Griffith University, Nathan QLD 4111, Australia*

ABSTRACT. Grass strips are known as one of the most effective management practices in controlling sediment loss to rivers and other surface water bodies. Some physically-based models have been previously developed to predict the amount of sediment retention in grass strips. Although physically-based models can explain the effects and interactions of various factors, they tend to be sophisticated as they require a large amount of input data. A nonparametric supervised learning statistical model was developed to predict the efficiency of grass strips in trapping sediments. Grass type and density, inflow sediment particle size distribution, slope steepness, length of strip, and the antecedent soil moisture were the five major factors on which the statistical model was built. The model was assessed by comparing with an independent dataset. Estimated bias, coefficient of model efficiency, mean absolute percentage error, Pearson product-moment correlation coefficient of the model were 1.01, 0.54, 18.1and 76% respectively. Testing the model predictions, permuting the input data, showed that inflow sediment particle size distribution, length of the buffer strip, and the antecedent soil moisture are the most important factors upon the performance of grass strips in trapping sediments. From the model outputs for a range of likely scenarios it was concluded that very long strips are needed in extreme conditions such as steep slopes, wet soil and sparse grass strips in order to trap sediments effectively.

Keywords: grass strip, model, regression tree, sediment, supervised learning

1. Introduction

One of the commonly used BMPs (best management practices) to reduce the amount of sediment and other pollutants in aquatic systems is grass buffer strips. These are narrow strips of grass or other dense vegetation on sloping agricultural lands in estuaries, riparian zones or located around the lakes and freshwaters to protect these surface water bodies. The grass buffer strips bring about changes in hydrology and hydraulics of the water flow that makes these strips effective in reducing sediment delivery. Resistance to the flow and high infiltration are the two main hydrological effects occurring in the presence of grass buffer strips.

High hydraulic roughness of the vegetated strips effectively decreases the runoff velocity (Borin et al., 2005; Deletic and Fletcher, 2006; Le Bissonnais et al., 2004). Increased hydraulic roughness by the vegetative zone is a main process which makes grass strips effective in removing sediment and associate pollutants, through decreasing the flow velocity and allowing more time to sediments to settle. The hydraulic roughness is higher in stiff and erect species than more flexible and less dense types. Dense strips also produce higher resistance to surface flow than sparse ones.

Previous research shows that vegetative strips can significantly increase the infiltration rate of the soil. Shrestha et al. (2005) observed that the vertical hydraulic conductivity of soil increases dramatically with vegetation. Schmitt et al. (1999) found that doubling the width of a grass buffer strip doubles the infiltration rate. The infiltration rate is highly related to the vegetation type as the cumulative infiltration under switch grass is significantly higher than under pasture and is more effective in reducing the amount of fine particles in the outflow (Bharati et al., 2002; Blanco-Canqui et al., 2004).

If a grass strip is not submerged, it can efficiently remove inflow sediment. Whilst the efficiency of vegetative buffer strips in removing sediment can be considerable, they are much less effective in removing particulate or sediment-associated nutrients as they are mostly attached to fine particles (McKergow et al., 2004). Dillaha et al. (1989) found that the effectivness of vegetated buffer strips reduces with time, due to sediment deposition, while Hussein et al. (2007a) observerd that the outflow sediment concentration is reasonably constant during runoff events. This difference is likely a result of whether the main area of deposition is upstream the grass strips or

* Corresponding author. Tel.: +61 7 3735 6713; fax: +61 7 3735 7459.
 E-mail address: s.akram@griffith.edu.au (S. Akram).

within them.

There have been several fields as well as controlled studies to assess the effect of vegetation on reducing nutrients and sediment transport over the landscape and sediment and pollutants delivery to aquatic systems. Controlled condition or flume studies have been conducted mostly to provide an understanding of the effectiveness of different factors in vegetated buffer strips performance in controlled environments (Ghadiri et al., 2001; Hussein et al., 2007a, b; Meyer et al., 1995). Field experiments have been carried out to assess the effectiveness of vegetated buffer strips under natural conditions but with less control over interfering factors (Daniels and Gilliam, 1996; Ghadiri et al., 2011; Magette et al., 1989; Parsons et al., 1994; Robinson et al., 1996).

A few physically-based models have been developed to predict the efficiency of vegetated buffer strips in sediment removal (Deletic, 2001b; Hussein et al., 2007b; Munoz-Carpena et al., 1999; Newham et al., 2005). Some currently available models are not completely process-based and their sediment transport and deposition sub-models are not very well validated (Munoz-Carpena et al., 1999; Newham et al., 2007). Some of the existing models neglect one of the two major zones of "upstream" or "within the grass strips" in their assumptions (Deletic, 2001b; Hussein et al., 2007b). The existing models are limited to some specific conditions or cannot accurately predict the efficiency of grass strips in removing sediments.

This paper presents a new modelling approach to predict the efficiency of grass buffer strips in removing sediment under different conditions. The statistical modelling approach presented in this paper is easy to use, physically sensible and accurately predicts the efficiency of grass strips in removing sediment.

This model is based on bagged ensemble statistical approach which is a supervised learning technique that has received considerable attention and has been used successfully for modelling purposes in various scientific fields. The results can be significantly accurate, physically meaningful, and models are improvable over time by adding new experimental data to datasets the model is built on.

This paper will scrutinize the most effective factors in grass strips performance by considering physical processes as well as experimental observations. The objectives of this paper include (i) review previous lab and field studies, and compile a comprehensive dataset on sediment trap efficiency; (ii) develop and apply new statistical models to predict the trapping efficiency of grass buffer strips; (iii) test and validate the new model in terms of accuracy and whether the results make any physical sense; and (iv) use the model to describe the structure and functionality of grass strips in removing sediment in different conditions.

2. Literature Review of Performance of Grass Buffer Strips

The performance of vegetated strips is reported in this paper by using the concept of efficiency, defined as:

$$Eff(\%) = \frac{(MS_i - MS_o)}{MS_i} \times 100 \qquad (1)$$

where *Eff* is the vegetated strip efficiency in trapping sediment, MS_i is the total mass of sediment entering the strip, and MS_o is the total mass of sediment leaving the grass strip. In some studies (Ghadiri et al., 2001; Hussein et al., 2007a, b; Loch et al., 1999; Meyer et al., 1995; Munoz-Carpena et al., 1999), the effectiveness of grass strips in reducing the sediment concentration has been introduced as the index of efficiency of vegetative buffer strips. Most of these studies are carried out in controlled conditions on impermeable surfaces. As field experiments indicate that infiltration can be as important as vegetation resistance to surface flow and sediment trapping, comparing the inflow and outflow concentrations is not a reliable indicator of the buffer strips' performance.

Some mathematical models also consider concentration as the index of efficiency rather than mass. Experimental results, however, show that even in strips where the outflow sediment concentration is higher than that in the inflow, there is noticeable reduction in mass delivery of sediment and buffer strips can still be effective (Schoonover et al., 2006).

2.1. Area of Deposition

The place of sediment deposition is mostly related to slope, vegetation type and density, and inflow sediment particle size distribution.

In several studies considerable deposition has been occurred just in the backwater region upstream of the grass strips (Blanco-Canqui et al., 2004; Dillaha et al., 1989; Hussein et al., 2006; McGregor et al.,1999; Meyer et al.,1995). The backwater region is the area upstream of grass strips where flow depth is increased due to high resistance the vegetation culms and leaves produce to the flow. This makes the flow velocity decrease, so providing more chance for sediment particles to deposit if their settling velocity is high enough. In all studies for which most deposition occurred in the backwater region, the slope was less than 12% and vegetation was very dense and stiff, situations which will be referred to as "hedge strips". A high ratio of coarse particles in the inflow resulted in more deposition in the backwater region as resistance was the dominant factor rather than infiltration. The experimental data reviewed showed that resistive force offered by the hedge strips reduces the concentration of the sediment while infiltration reduces only the mass.

For high slopes where it is not possible to have an effective backwater length and depth in the upstream area, the area of deposition is shifted into the buffer (Hall et al., 1983; Patty et al., 1997; Robinson et al., 1996). The length and depth of backwater for the same flow rate and vegetation type and density changes substantially with slope. Therefore, it is almost impossible to have sediment trapping upstream of the grass buffer strips in slopes more than 12% (Ghadiri et al., 2001; Hussein et al., 2007b).

Data from the literature show that as the stiffness and

density of vegetation decreases the amount of sediment deposition within (rather than upslope) the buffer strips increases (Blanco-Canqui et al., 2004). This occurs because of reduced turbulence, shorter backwater, lower actual velocity and less friction slope and greater effectiveness of infiltration.

In less dense or stiff vegetation types, the backwater length is shorter and the depth lower in comparison to hedge strips. Therefore, lower amounts of sediment are trapped in the upstream region, and deposition within the grass strip is dominant.

Data from the literature show that the amount of sediment deposition in the backwater area is higher than that within the grass strip of the same length (Pan et al., 2010). Four major reasons can be enumerated for this: 1. The actual velocity is much higher within the strips comparing to the backwater as a large area of the flow cross section is barred by grass stems and leaves; 2. Most of the coarse particles have settled upstream, and the inflow through the strips is less concentrated; 3. Turbulence is high within the grass strips; 4. The friction slope is high and does not allow particles to settle.

Vegetation resistance to the flow can significantly reduce the flow velocity and subsequently provide coarser particles an opportunity to settle upstream the grass strips. As the settling velocity of fine particles is low, the high infiltration within the grass strips becomes the major factor in trapping fine and suspended sediments.

In low slopes, regardless of the different factors involved in different studies, the main area of deposition associated with hedgestrips is found to be in the backwater region upstream the strip (Blanco-Canqui et al., 2004; Hussein et al., 2007a; Meyer et al., 1995). Sediment trapping occurs more or less equally upstream and within dense, but less stiff grass species (Dillaha et al., 1989; Ghadiri et al., 2011; Loch et al., 1999). In sparse and limber grass strips, due to low resistance, no significant backwater forms upstream and deposition occurs within the buffer strips (Arora et al., 1996; McKergow et al., 2004; Young et al., 1980).

The shear stress within the grass buffer can be high. According to Hairsine and Rose (1992) the rate of entrainment and re-entrainment are related to the rate of energy expenditure of the flow known as stream power. According to Knighton (1999), stream power is equal to:

$$\Omega = \rho q S_f / g \tag{2}$$

where ρ is the density of water, q is the unit discharge, S_f is the friction slope, and g is acceleration due to gravity. As the effective flow width is reduced in grass strips, the actual unit discharge, hence the stream power, would be higher in dense grass strips.

As in many case studies the length of hedge strips is short the length of the decay curve, which is the place that water surface profile tends to decrease to the level of the exit point of the buffer, can be close to the strips length. Therefore the friction slope in narrow strips is higher than the bed slope,

which produces more pronounced stream power in this region (Ghadiri et al., 2001; Hussein et al., 2007a; Raffaelle et al., 1997; Rankins and Shaw, 2001). Erosion has been observed in narrow dense grass strips by Ghadiri et al. (2001), this being due to a high energy slope within the narrow strip.

2.2. Flow Rate

The effect of flow rate on erosion and deposition processes is twofold. For any given grass strip, higher flow rates produce longerbackwater lengths, which mean that there would be more room available for sediments to deposit in the upstream zone. However, according to the vertical velocity distribution of a sheet flow, velocity is a minimum in the bottom and maximum in the top of the water profile. This means that as flow rate increases the average velocity grows dramatically and there will be less chance for sediment particles to settle.

High flow rates provide less chance for water to infiltrate the soil and the rate of runoff reduction is less in higher flow rates than low flow rates in same areas. Therefore, there will be less sediment trapping due to infiltration within grass strips in high flow rates. The other fact is that within low flow rates, finer rather than coarser particles would be eroded in significantly higher rate due to preferential detachment.

2.3. Infiltration and Antecedent Soil Moisture

Infiltration is one of the two major components in trapping sediment in vegetative buffer strips. The overall experimental results show that the runoff reduction is higher in older, stiffer, and denser grass strips comparing to sparse and limber ones (Blanco-Canqui et al., 2004; Le Bissonnais et al., 2004; McKergow et al., 2004; Patty et al., 1997; Rankins and Shaw, 2001; Schmitt et al., 1999; VanDijk et al., 1996; Young et al., 1980).

Comparing the results of experimentswhere the only changing variable was the antecedent soil moisture shows significant reduction of the efficiency in wet soils comparing to dry soils. Results from Young et al. (1980), Dillaha et al. (1989), and Magette et al. (1989) show that the efficiency of grass strips in removing sediment is 15 to 40% higher in dry soils comparing to wet soils in the same areas.

In the experiments of Arora et al. (1996) the downstream edge of the source area plot was higher than the upstream edge of the vegetative buffer strips, therefore no backwater formed. As all sediment retention occurred within the grass strips, therefore infiltration was the key factor in removing sediment. Sediment removal was considerably higher while the antecedent soil moisture was low, and lower when the soil was wet.

2.4. Slope Steepness

Increasing the slope lowers the backwater length and experiments show that no effective backwater occurs when slope is more than 12% in most natural conditions. The negative influence of high slopeon sediment removal can be am-

ended partially by increasing the length of strip. The flow velocity is directly correlated with slope, and longer distance within steep strips is needed for sediments to deposit compared to the same conditions in planar areas. Also, as the water flows with higher velocity, the water and soil contact within the grass strips is of shorter duration, and therefore less water infiltrates soil comparing to steeper areas of same conditions.

The results of VanDijk et al. (1996) confirm the above arguments that unless the slope is high and no backwater forms upstream, the deposition takes place within the buffer strips, and the efficiency is considerably higher in longer strips. Parsons et al. (1994) results also confirm that in low slopes the effect of buffer length on the efficiency is low, but as the slope increases, the efficiency is much higher in longer buffers.

The concept of efficiency (Equation 1) cannot perfectly describe the performance of grass strips, as in many experiments such as Daniels and Gilliam (1996), Loch et al. (1999), and Ghadiri et al. (2001) the trapping efficiency was higher in steeper lands, due to higher erosion happening in the upstream area containing more coarse fractions due to preferential detachment. In the experiments where slope was the only variable and sediment delivery into buffers was constant, the efficiency of vegetated buffers decreased as slope increased (Deletic, 2005; Hussein et al., 2007a). Overall, sediment delivery is significantly less in lands of lower slope.

2.5. Plant Type and Density

The density of vegetation affects the performance of grass strips in removing sediment. Hedge strips produce longer and deeper backwater upstream of the hedges. This is a place that most of the deposition occurs in these types of grass strips (Blanco-Canqui et al., 2004; Dabney et al., 1995; Ghadiri et al., 2001; Hussein et al., 2007a, b; Meyer et al., 1995). Although the long and deep backwater profiles can trap a high ratio of incoming sediment, this only occur in low slopes. The results from Akram et al. (2014) shows although the backwater area upstream narrow hedge strips can trap a large fraction of incoming sediment in low slopes, the friction slope is high within narrow hedges and sediment deposition is less likely to occur within the hedges comparing to the upstream section. Filtering or deposition within the strips has been observed to be greater in still dense but less resistive grass than in previously mentioned hedge barrier strips.

Hedgetype strips are more effective in dispersing rill flows and decreasing rill erosion hazards. The combination of a narrow, old and dense hedge strip following by long dense strips can be very effective in lowering the fine particles loss as they need longer distance and less turbulent flow to settle (Blanco-Canqui et al., 2004).

2.6. Particle Size Distribution

Vegetated buffer strips of all types and lengths can trap inflows and fractions most effectively (Deletic, 2005; Hook,

2003; Hussein et al., 2006; Ma et al., 2013). In intense storms, the ratio of detachment of coarse particles increases and the efficiency of buffer strips soars (Daniels and Gilliam, 1996; Parsons et al., 1994; Robinson et al., 1996). On the other hand, due to the preferential detachment of fine particles in low intensity rainfall events the trapping efficiency is low. Therefore the efficiency of grass buffer strips in reducing sediment loss can be higher in intense events.

It is very unlikely to observe vegetated strips of any types and lengths reducing the concentration of particles less than 5 μm in the outflow (Deletic, 2005; Jin and Romkens, 2001). However, grass strips can be effective in significantly reducing the discharging mass of the fine particles if the infiltration within the strips is substantial. As particulate nutrients are associated with fine textured particles, the efficiency of buffer strips in reducing particulate nutrients loss is less than sediment as a whole. Increasingthe length can improve the performance of grass strips in reducing the sediment bound nutrients loss by increasing the runoff loss.

3. Data and Methods

Thirty four field as well as controlled experiments were compiled for this paper. The observations, results and experimental conditions of these studies were used to identify the effects and interactions of different factors in vegetative buffer strips performance in trapping sediments. Two non-parametric supervised learning models were developed using literature data of various conditions in order to predict the performance of vegetative buffer strips of different kinds in sediment trapping.

The factors affecting the performance of grass strips in removing sediment are numerous. Based on literature review, physical processes and sensitivity analysis of the previously developed models (Deletic, 2001a; Munoz-Carpena et al., 1999) show that the factors discussed in the previous section are the most important ones. In addition, several other factors can be important, such as time, flow type, planting slope, rill formation, cattle trampling, soil compaction (McDowell et al., 2003; Meyer et al., 1995; Tadesse and Morgan, 1996). It is not possible to consider all these factors in a model developed to simulate the physical processes involved. The statistical model developed for this paper should be able to simulate and predict most probable scenarios using limited data from major factors.

By collecting data from 34 different studies in various conditions, and considering grass type, sediment type, strip's length, slope steepness, and antecedent soil moisture as most effective factors in grass strips performance in removing sediment, two statistical models were developed. Nonparametric supervised learning (machine learning) methods were utilised to build these models.

Grass type, inflow sediment type, and the antecedent soil moisture were categorised as the first two were not continuous variables, and accurate measurements had not been carried out for the third factor in most studies. The categorisation of these

3 major factors is shown in Table 1. Responses were the efficiency of systems in removing sediment mass. These categorizations considered the major effective factors as spectrums and the observed data categories were entered as numerical vectors. The dataset collected from research papers is presented in the Appendix Table. The x1, x2, ..., x5 were factor identifiers which were used further in the regression tree model development. As examples, x1 = 2 means that the grass strip is dense, and x3 = 5 indicates that the slope steepness is 5%. Slope steepness and the length of the grass buffer strip were continuous factors, whilst the other factors were categorised. The observed efficiencies ranged from 7 to 100%.

Grass types and density is the first categorised factor which is split to three categories. Category 1 is for very dense and stiff strip types mostly consisted of dense hedges of vetiver and switch grass or similar species. Category 2 consisted of dense but less stiff species like fescue, sedge, bermuda grass, and similar species. It also includes stiff but less dense strips. Category 3 consisted of sparse and limber types like sparse meadow strips. Inflow sediment particle size distribution is categorised based on USDA soil texture triangular. The antecedent soil moisture was based on whether the soil was saturated, dry or the surface was impermeable at the time of the events.

Two statistical models were developed based on the data presented in the Appendix Table. The first one is a non-para metric supervised learning "regression tree", and the second one is a nonparametric supervised learning "bagging ensemble" (Dietterich, 2000). The regression tree was built to clarify the effects and importance of every factor in different conditions. The ensemble model accurately predicts the efficiency of grass strips in removing sediment. All data processing and model developing were performed using the Statistics toolbox in MATLAB software package.

3.1. Regression Tree

Regression trees are nonlinear predictive models which predict the response from various input variables by growing a binary tree (Razi and Athappilly, 2005). We choose left or right branches of a tree in each node based on the known conditions, and proceed to reach the outcome. The predicted response can be found by following the decisions from the root (starting) node down to a leaf node, which illustrates the outputs of the model. The prediction is indeed the aggregation of all the training data points that lead to that leaf.

A regression tree was developed in order to predict the

efficiency of grass strips in removing sediment resulted from the input data. The regression tree is created by using all input data and testing all conceivable binary splits on each factor. The procedure used for creating the regression tree is as following:

- Testing and trying all possible binary splits on every variable using all input data.
- Choosing a split with the best optimisation criterion. The mean-square error (MSE) was the criterion and the splits with the minimum MSE of predictions compared to the training data were chosen. Data used to build models is called "training data" and data utilised to test the models is called "test data" afterwards.
- Exert the split.
- Recursively iterating for the two child nodes.

Splitting is stopped when there are less than "Minleaf" observations in a node, where Minleaf is defined by the user, and refers to the minimum number of observations per tree leaf. Cross validation accuracy test was calculated for regression trees having different Minleaf numbers to find the optimum Minleaf value. Figure 1 shows the cross validation MSE for different Minleaf values.

To find the cross validation error, the training data were split into ten random parts. Ten new trees were trained, each by nine parts of the data. Then the accuracy of the new trees was examined by the data that are not part of that training tree. As this method tests new trees on new data it assesses the precisionof the tree accurately. As Figure 1 shows the Minleaf
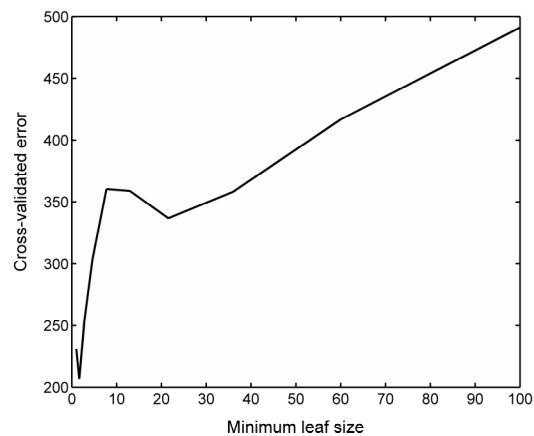

Figure 1. The cross-validated error over minimum leaf size of the regression tree.

Table 1. Categorising the Major Factors in Grass Buffer Strips Performance

Category	Grass (x1)	Sediment (x2)	Slope (x3)	Buffer length (x4)	Moisture (x5)
1	Dense hedges	Sandy loam	Continuous	Continuous	Dry
2	Dense grass strips	Silt loam	(from 1to 16%)	(from 0.14 to 27.43 m)	Saturated
3	Sparse strips	Clay loam			Impermeable
4		Silty clay loam			
5		Clay			

5

value of 2 yields the least error and this value was set as the Minleaf to build the tree. The cross validated MSE of the regression tree with Minleaf equal to two was 207.6.

Resubstitution error has also been calculated for the regression tree, where this error is the MSE between the observed efficiencies of the training data and the tree predictions based on the input training data. The predictions of the tree are not accurate if the resubstitution error is high, but low resubstitution error does not necessarily confirm the significant validity of the predictions. The measured resubstitution MSE for the regression tree with Minleaf equalto two was 136.5. These rates were acceptable values as the range of variation of the observed efficiencies was between 7 to 100%, and the squared root of the errors were 14.4 and 11.7% for cross-validated and resubstitution respectively.

Although Minleaf equal to two built the most accurate tree, the resulted regression tree was very leafy. A leafy tree is hard to interpret. Therefore, it was pruned to an appropriate level in order to be fitted in this paper. Leafy trees are usually accurate just for the training data, not for independent new conditions. They tend to overtrain, meaning they estimate the outcomes optimistically. The depth of the tree was controlled to give simple and accurate result, easy to interpret.

3.2. Ensemble Method

In supervised learning (Machine learning), "weak learner" refers to a model (learner) with high probability of error in predictions. Decision trees are classified as weak learners. Ensemble methods on the other hand are strong learners combining results from multitude weak learners (Dietterich, 2000).

An ensemble method was utilised to predict the performmance of grass buffer strips of different types and in different conditions in removing sediments. The results of this model as well as the regression tree statistical model were also used to interpret the physical processes taking place in different conditions.

An ensemble learning model is the combination of multiple weak learners such as regression trees to predict or solve complex problems. As the ensembles aggregate the decisions made by multiple regression trees, they improve the robustness of single regression tree models.

A matrix of input data has to be given to the model based on observations. This matrix is the same as the one used to build the regression tree and the predictor variable values for every event are also given to the model in the same way as for the regression tree (Appendix Table).

An applicable ensemble algorithm had to be chosen for creating the ensemble. The 'bootstrap aggregation' algorithm also called 'bagging' was chosen to build the ensemble as it is applicable for regression decision trees with more than two variables (Breiman, 1996). This algorithm was found to be significantly more accurate compared to the other tested algorithm 'boosting'. In bagging, boots trapped replicas of the training data are providing the diversity of classifiers. This

means that in each replica, unique training data subsets from the entire training dataset are randomly drawn. Each of these subsets is used to produce a different classifier which in this case is a regression tree. The final decision is made by combining the predictions made by single learners by taking a simple average from them.

Bootstrap aggregation trains learners on re-sampled copies of the data. Re-sampling is done by bootstrapping observations, which is choosing n observations from the whole dataset with replacement for every new learner. For increasing the accuracy of the bagged trees, every tree within the ensemble can randomly choose variables for classification splits. The minimal leaf size of the bagged trees was set to five to have the high accuracy in the least running time.

The number of trees used to develop an ensemble has to be defined for building the ensemble. The appropriate size for an ensemble is one which balances speed and accuracy. It takes longer for large ensembles to train and produce predictions. The other problem that a large number of trees in an ensemble may cause is to make it overtrained and inaccurate. The MSE for different number of trees used for preparing the ensemble has to be calculated to find the optimum number of trees which the ensemble is built with.

Figure 2 shows the mean square error for different number of trees used in developing the bagging ensemble predictor. Two different tests have been applied to find the accuracy and quality of the ensemble over the number of trees that it is built on. The ensemble is evaluated on 'out of bag data' as well as 'cross validation' tests to find the optimum number of trees to develop the ensemble with.

For the cross validation test a five-fold cross-validated bagged ensemble was generated. It means that 80% of the data used for training and 20% for testing. The cross-validation loss as a function of the number of trees in the ensemble was tested and is depicted in Figure 2.

For finding the quality of the ensemble on out of bag data, the loss curve for out of bag estimates was generated and is illustrated in Figure 2. An average of 37% of the observations
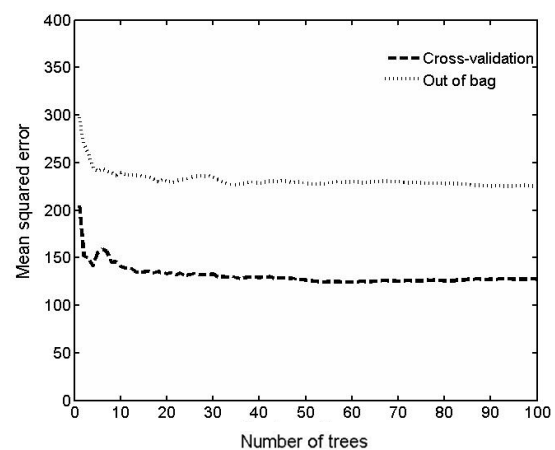

Figure 2. MSE of the ensemble method over different number of trees used for ensemble development.

is omitted for each regression tree. These are called 'out of bag' observations. The out of bag prediction is estimated by averaging predictions from all trees in the ensemble for every observation in the ensemble for which this observation is out of bag. The mean square error for a single observation is calculated by comparing the predictions and the actual observed efficiencies for this observation. The out of bag error is determined by comparing the out of bag predictions against the actual observations for all observations used for training.

As Figure 2 shows the errors were roughly constant in ensembles developing with more than 50 trees. Three hundred trees were used to prepare the ensembleas it has the lowest error and the model can be run quickly.

The ensemble predictions were calculated for a series of data sets and results were compared to actual experimental observations using different evaluation techniques. Therefore, the model was run 30 times, and in each run 80% of the data were used for training and the remaining 20% were utilised to test the model. The splitting was carried out randomly and the split was different for each of the 30 runs. The error estimations are the average of the values for every model evaluation techniques over the 30 runs.

The 'Bias' of themodel was calculated from the model predicted efficiencies and the actual observed grass strips efficiencies in trapping sediment as:

$$Bias = \frac{\sum M_i}{\sum O_i} \qquad (3)$$

where M_i and O_i are modelled and observed efficiencies respectively. The Bias criterion shows whether the model systematically overestimates or underestimates the trap efficiency depending on whether the bias is greater or less than unity.

The observed efficiencies of grass strips of different conditions were also compared to the model predictions. Coefficient of model efficiency Ec (Nash and Sutcliffe, 1970) was calculated from the observed and modelled data as:

$$Ec = 1 - \frac{\sum (O_i - M_i)^2}{\sum (O_i - \overline{O})^2} \qquad (4)$$

where $\overline{O}$ is the mean of the observed values. The model efficiency measures the level of accordance between the modelled and observed values. Ec values of 1 indicate a perfect fit, while negative Ec values indicate the model predictions are worse than those predicted simply by the average observed values.

In addition, the mean absolute percentage error (MAPE) was calculated as:

$$MAPE = \frac{\sum abs \frac{(O_i - M_i)}{O_i}}{n} \qquad (5)$$

where n is the number of observations. MAPE qualifies the magnitude of error in the modelled values and this measure of accuracy is expressed as percentage. MAPE equals to zero shows the perfect fit and low MAPE indicates better fit than high values.

The Pearson product-moment correlation coefficient (PWM) was the other criterion used for assessing the accuracy of model predictions as:

$$PWM = \frac{A}{B} \frac{n\sum O_i M_i - (\sum O_i)(\sum M_i)}{\sqrt{n\sum O_i^2 - (\sum O_i)^2}\sqrt{n\sum M_i^2 - (\sum M_i)^2}} \qquad (6)$$

$$A = 0.5 \frac{(\sum O_i^2 + \sum M_i^2)}{\sum O_i} \qquad (7)$$

$$B = 0.5 \frac{(\sum O_i^2 + \sum M_i^2)}{\sum M_i} \qquad (8)$$

PWM is a criterion, expressing the linear correlation between observed measurements and modelled outcomes. PWM equal to 1 indicates the perfect correlation between observed and modelled efficiencies.

4. Results

4.1. Nonparametric Supervised Learning Regression Tree

Figure 3 shows the regression tree built using data shown in the Appendix Table. The regression tree presented in Figure 3 is developed to clarify the effect of different factors on sediment removal with grass buffer strips. As regression trees are categorised as weak learners, therefore we did not use them for predictions.

The tree predictions can be derived by following the decisions from the rootnode down to the leaf nodes. The far left side of the tree shows that very high efficiencies in removing sediment is expected from low slope dry hedge strips with lengths more than 5.5 meters. The decision tree shows that the sediment trapping efficiency decreases as the inflow sediment contains higher portions of fine particles.

By comparing the leaf nodes which have close strips length and slope values, it is concluded that trapping efficiency is significantly less in wet soils comparing to dry soils.

Significant difference can be observed between the sediment trapping efficiency in different grass strip types (x1). Hedges can highly trap sediment even in high slopes and medium lengths, while sparse strips can trap sediment moderately. Grass strips of any kind cannot reduce the amount of sediment in the outflow significantly if the moisture content of the soil is high and the inflow sediment mostly consisted of very fine particles.
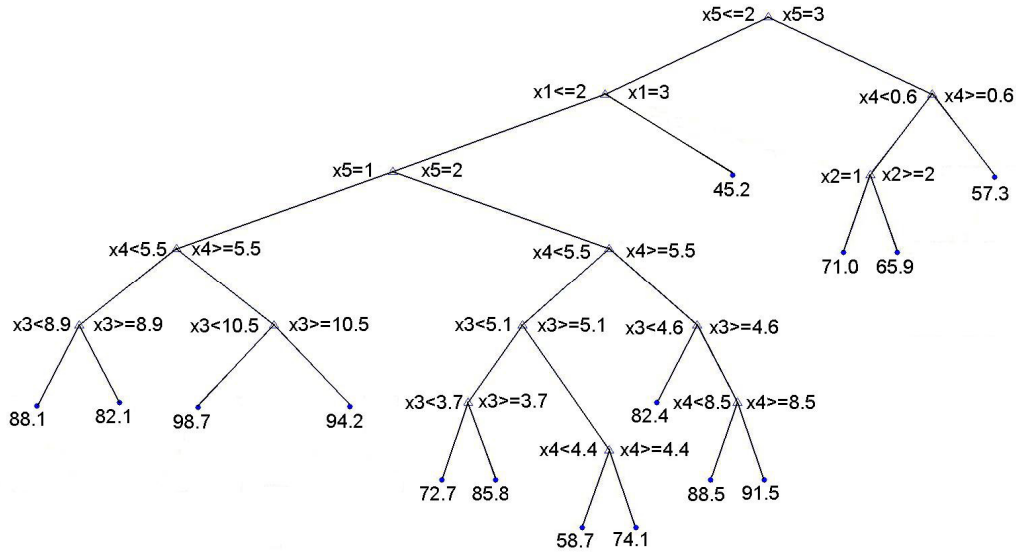
Figure 3. The regression tree of the effectiveness of grass strips in trapping sediment.

4.2. Nonparametric Supervised Learning Bootstrap Aggregation Ensemble Method

The average values of the Bias, Ec, MAPE, and PWM of the model for 30 runs were 1.01, 0.54, 18.1, and 0.76 respectively. As 20% of the data which was equal to 40 observations was reserved for testing the model and was not used for training purposes, it was expected that the model would significantly fit better by adding the remaining 40 split observations to the training data.

To obtain a better understanding of the physical processes, the importance of each variable was estimated in the final result. The values for every variable were transposed manually across the dataset to determine changes in the MSE values. The process was repeated for every variable. MATLAB functions have the capability of conducting this process and store the growth in MSE averaged over all tress in the ensemble and divided by the standard deviation taken over the trees, for every variable. Larger values represent more important variables. Figure 4 shows the importance of every variable in the ensemble.

Figure 4 shows that all five considered variables are important in grass strips performance in removing sediments. Sediment type, length of strips, and the antecedent soil moisture were the most effective variables. This result is similar to findings by Munoz-Carpena et al. (1999) in sensitivity analysis of VFSMOD model. Slope was the least effective factor and this is likely a result of the limited variation in slope among the compiled dataset, from 1 ~ 16% only.

Figure 5 shows the observed and modelled efficiencies for the whole 30 runs which were calculated by splitting non-repetitive 20% testing data.

Figure 5 illustrates the model over-predicts small sediment deliveries and under-predicts large sediment deliveries. The subject will be considered further in the discussion section.

New input data not included in model development was entered to the ensemble model to evaluate the effectiveness of grass buffer strips in various conditions in trapping sediments and the results are shown in Table 2. The model predictions are also mentioned in this table.

The results of the critically comprehensive literature review which considered the physics of the involving processes that affect hydraulics and hydrology of the flow, plus the results of the regression tree (Figure 3) and the scenarios which were run by the bagged ensemble model (Table 2) were used to better explain the physics of the occurring processes upstream and within grass strips.

Results of the ensemble model, which are in agreement with what have been observed in many studies, show that the backwater region is a very important zone for sediment deposition. The results show that considerable deposition occurs in very narrow dense hedges (type 1), where slope is low and sediment inflow contains considerable portion of coarse par-
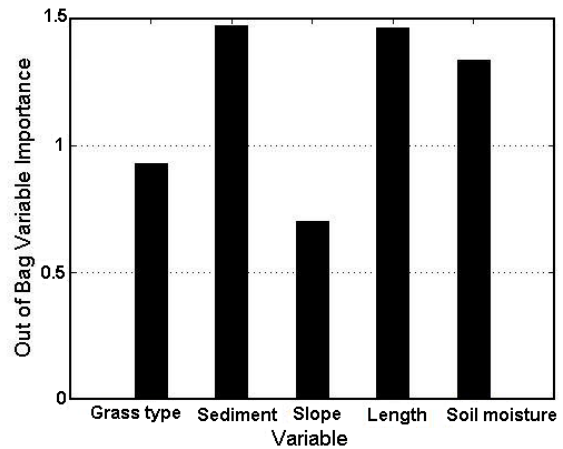

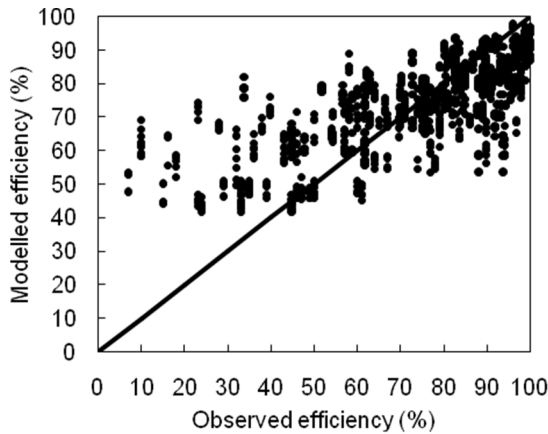**Figure 4.** The importance of every variable in the ensemble.

Figure 5. The observed and modelled efficiencies for 30 runs using non-repetitive 20% testing data.

ticles. This zone has to be considered separately, as the flow velocity decreases and particles may have enough opportunity to settle. As the infiltration rate is less in the backwater zone compared to the grass strip and the length of this zone is short, infiltration effects can be neglected in this zone.

Results of the statistical model show great difference in grass strips efficiency in trapping sediment between initially dry and wet soils. This shows the importance of infiltration in sediment deposition within the grass strips. The literaturere view confirms that even very fine particles of soil can be trapped in grass strips. A number of studies have shown that the mass of nutrients associated with sediment decreased in the outflow without significant change in the concentration. It can be concluded that particles are trapped due to infiltration re-

gardless of their size. Nevertheless, sediment deposition due to immersed weight is considerably less within the strips due to high turbulence, high friction slope, and high actual flow velocity.

Illustrations of the conclusions were prepared based on physics of the processes involved and the ensemble predictions for various possible scenarios (Table 2) to show the processes upstream and within grass strips in different conditions.

The illustrations (Figure 6) were prepared to show the effect and interactions of each of the five major factors, on which the model was built. They are based on some of the scenarios in Table 2 to compare the efficiencies in different conditions. For these figures it was assumed that the water flowing toward the grass strip consisted of 4 sand, 4 coarse silt, 4 fine silt, and 4 clay particles. It is similar to conditions defined in first 18 scenarios in Table 2. Deposited particles due to resistive force of the grass to the water flow are shown as hollow figures while particles settled due to infiltration are filled. The remaining particles in the outflow are shown over the water surface in the downstream zone.

As the infiltration rate is higher within the vegetated area compared to the upstream zone in the backwater region, it is assumed that the sediment deposition occurs in the backwater zone only due to the flow deceleration, and within the strips by deceleration as well as high infiltration. Particles settling due to flow deceleration are assumed to be less within the grass strips compared to the same water depths in the backwater zone for reasons given previously. It is also assumed that very fine particles only settle due to water infiltration into

Table 2. Different Scenarios for Grass Buffer Strips Conditions and the Ensemble Method Predictions

Scenario	Category					Predicted performance (%)
	Grass type	Sediment type	Slope (%)	Buffer length (m)	Antecedent soil moisture	
1	1	3	2	5	1	87.2
2	1	3	2	5	2	79.3
3	1	3	2	1	1	80.1
4	1	3	2	10	1	92.8
5	1	3	14	5	2	77.9
6	2	3	2	5	1	85.2
7	2	3	2	5	2	75.8
8	2	3	2	1	1	77.3
9	2	3	2	10	1	91.2
10	2	3	14	5	2	72.5
11	3	3	2	5	1	71.4
12	3	3	2	5	2	62.6
13	3	3	2	1	1	68.4
14	3	3	2	10	1	76.3
15	3	3	14	5	2	59.6
16	1	1	2	5	2	79.3
17	2	1	2	5	2	75.9
18	3	1	2	5	2	64.0
19	1	5	2	5	1	68.9
20	1	5	2	5	2	61.0
21	3	5	14	1	2	41.9

the soil. It is presumed that settling due to infiltration is not preferential and particles of different sizes have equal chance to be trapped as water infiltrates to soil.

Dry hedge strips of medium length and low slope (Figure 6a) can effectively trap almost all inflow sediment. The long and deep backwater region decelerates the water flow significantly and coarse and medium size particles trap in this region. As the root system is widespread in category 1 species and the soil is dry, the medium length strip can effectively trap almost all the remaining sediment in the flow within the grass strips. As the backwater region in low slope hedge strips traps the majority of the inflow sediment, the concentration of the sediment entering the grass strips is low. Therefore, the effectiveness of strips of these conditions in trapping sediment is high, even under wet soil conditions. In high slopes the length and depth of backwater decreases considerably. Thereupon, sediment removal in the backwater zone is considerably lower comparing to low slopes. Wet soils under these conditions cannot trap sediments effectively (Figure 6b). As the actual flow velocity is very high in hedge strips, particles have lower chance to settle due to gravity and infiltration compa- ring to sparser strips. The high performance of hedge strips is highly dependent to low slope steepness.

As the resistance caused by the medium-density grass species is low comparing to hedge strips, the length of the backwater region is short and the water depth upstream the buffer strip is low. As shown in Figure 6c, the effectiveness of the antecedent soil moisture is more pronounced under these conditions, as more concentrated water flows through the strip. Dry, medium length grass strips of medium density in low slope fields, can efficiently trap sediment (Figure 6c). As the soil moisture increases, the effectiveness of these areas decreases considerably. As fine particles are assumed to be trapped only by infiltration within the grass strips, there will be considerable amount of fine sediments trapped under wet conditions. The significance of this issue is that particulate
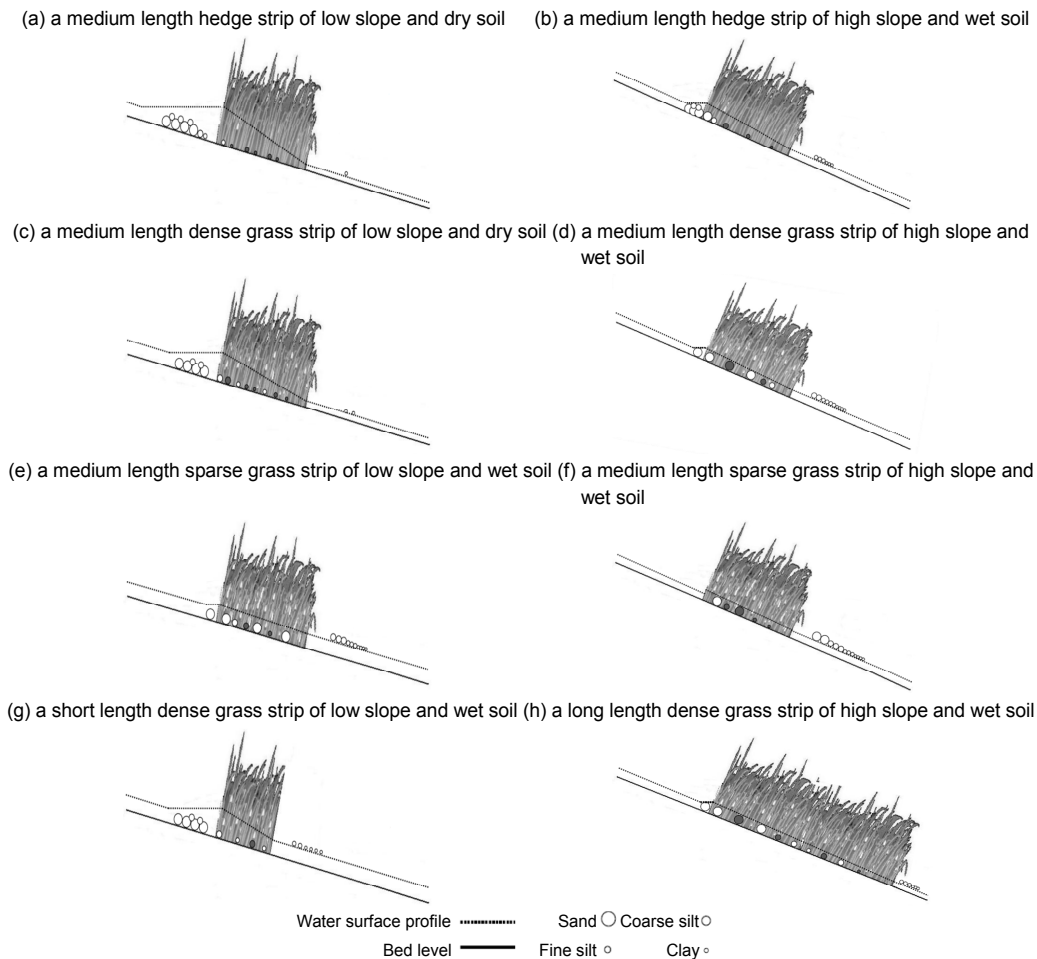
(a) a medium length hedge strip of low slope and dry soil (b) a medium length hedge strip of high slope and wet soil

(c) a medium length dense grass strip of low slope and dry soil (d) a medium length dense grass strip of high slope and wet soil

(e) a medium length sparse grass strip of low slope and wet soil (f) a medium length sparse grass strip of high slope and wet soil

(g) a short length dense grass strip of low slope and wet soil (h) a long length dense grass strip of high slope and wet soil

Water surface profile ········· Sand ○ Coarse silt ○

Bed level ▬▬▬ Fine silt ○ Clay ·

Figure 6. Particles transport and deposition in: (a) a medium length hedge strip of low slope and dry soil; (b) a medium length hedge strip of high slope and wet soil; (c) a medium length dense grass strip of low slope and dry soil; (d) a medium length dense grass strip of high slope and wet soil; (e) a medium length sparse grass strip of low slope and wet soil; (f) a medium length sparse grass strip of high slope and wet soil; (g) a short length dense grass strip of low slope and wet soil; and (h) a long length dense grass strip of high slope and wet soil.

pollutants are attached to fine particles and the delivery of pollution is expected to be very high under wet conditions. The efficiency will be much less in steep fields (Figure 6d).

Sparse grass strips can only be effective in low slopes and dry soils with lengths more than at least one metre. As no significant backwater forms even in low slopes, almost all sediment removal occurs within the strips. Sediment loss from wet soils or high slope fields with sparse grass strips is significantly high (Figures 6e and 6f).

Considering the physical processes, the prepared regression tree, and the ensemble model results presented in Table 2, it is concluded that increasing the length of grass strips can amend the adverse effects of a high ratio of fine particles in the inflow, sparse strips, high slopes, and high moisture content of the soil on sediment trapping efficiency to some extent. As conditions tend to extreme, like intense storms in steep fields, longer buffers can be remarkably more effective than short ones in removing sediment. Otherwise, short strips of hedges or dense species can trap coarse and medium size particles with very high efficiency if the slope is low and the soil is dry. Figures 6g and 6h show the effectiveness of grass strips length factor in performance of buffer strips under various conditions same as Table 2.

5. Discussion

Statistical modelling is a powerful tool for developing predictive models and decision support systems. Ensemble methods are precise predictors and user-friendly as there is no need for high proficiency tousethem.There have been several successful implementations of supervised learning for developing decision support systems in various science and technology fields (Bhattacharya et al., 2007; Hong, 2008; Shipp et al., 2002).

The advantage of statistical models in general, and the statistical model described in this paper in particular over physically-based models, is their high accuracy and simplicity in "application". The number of input variables needed for running physically-based models is usually high and some are hard for users to estimate. The statistical model described in this paper just needs five numbers/categories as input data which are all easy to measure or determine. For this reason, it can be expected that amongst two physically-based and statistical models with similar accuracy testing values,the statistical model have the advantages that the number of input data is less and easier to measure or estimate.

Figure 7 shows the bias of the developed ensemble model. It illustrates the difference between the observed and modelled efficiencies over the observed efficiencies. The scattered points should be randomly distributed above and below the zero value gridline for the y axis in high and low observed efficiencies in an unbiased model. Figure 7 shows that most of the y axis values are positive for observed efficiencies more than 75%, and negative for observed efficiencies less than 75%. The average observed efficiency for the whole dataset was also 75%. Therefore, the figure shows that the model

over-predicts small sediment deliveries and under-predicts large soil deliveries. According to the findings of Nearing (1998), this is a common problem for all soil erosion models, regardless of being physically-based or empirical. This study confirms the same trend for statistical modelling.

Nearing (1998) indicates that applying models to data which contain 'natural variations' that are not possible to catch by models, causes a bias in the model outcomes. The grass buffer strip models can basically be more biased as grassed areas are expected to have more variations in surface configurations than fallow fields around it. Notable variance can be observed in the performance of grass strips in same conditions in removing sediment (Daniels and Gilliam, 1996; Magette et al., 1989). The other important factor that causes the bias is that it is not possible to consider all effective factors involved. Every model regardless of its type neglects some factors. Also there are some factors that are not known to be effective or their effectiveness is not clear. However, in statistical modelling, important factors can be considered without knowing how exactly they influence the outcome.
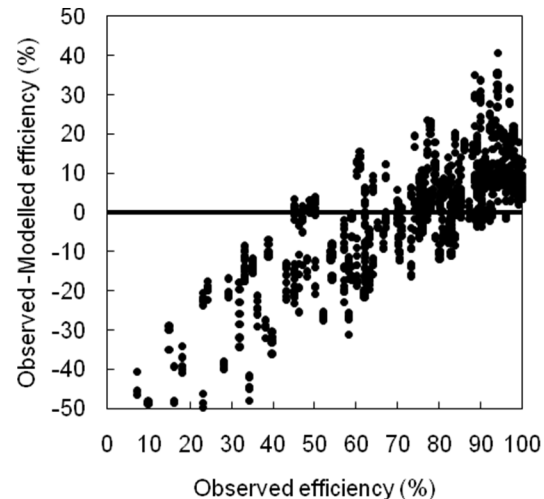

Figure 7. Bias of the model predictions.

The model developed in this research can be modified and improved over time. As studies to evaluate the effectiveness of grass strips in removing sediment and particulate nutrients are still ongoing in many research institutions, their results can be added to the currently gathered database to make the model more accurate as larger datasets can produce more accurate statistical models. In this study, it was observed that the *RMSE* of the model was significantly lower while 90% of the data was used for training and 10% for testing, than allocating 80% of the data for training and 20% for testing. The reason was that the first case had 20 more observations to build the ensemble with than the second case.

6. Conclusions

Using data in the literature to develop statistical models as well as considering the physical processes affecting the hy-

draulics and hydrology of flow upstream and through grass buffer strips showed that it is necessary to accurately consider grass resistance, slope steepness, antecedent soil moisture, sediment type, and strips length as important factors affecting the effectiveness of buffer strips in sediment removal.

Machine learning ensemble methods with the bootstrap aggregation algorithm can predict the performance of grass buffers of different types in removing sediment under certain conditions.

The regression tree built using the collected data from very different natural and controlled conditions, showed that unless the slope is low and the inflow sediment does not contain high portion of clay particles, moderate lengths of grass strips can trap large amounts of inflow sediment. The tree shows that in any grass strip type sediment removal would be low if the moisture content of the soil is high and the inflow sediment mostly consisted of very fine particles.

Evaluating the ensemble model with different model testing methods showed that the statistical non-parametric supervised learning ensemble developed in this research has prepared a strong learner for predicting the trapping efficiency of grass strips. Testing the ensemble learner predictions, permuting the inputdata, showed that sediment particle size distribution, length of strip, and the antecedent soil moisture were the most effective variables.

Comparing the results of the ensemble predictions for different prevalent scenarios showed that the backwater region upstream of dense grass strips (hedge strips) is the main region for sediment deposition on low slopes. The efficiency of grass strips in reducing the concentration of sediment is much higher for coarse than finer particles. Grass strips can substantially decrease the mass of fine particles if a significant reduction in runoff (ie. infiltration) occurs within the strip. As no backwater forms on high slopes and the flow velocity is high in steep lands, particles will not have enough time to deposit ahead of the strip. Having long grass strips can amend the low trapping efficiency associated with extreme conditions such as high slope, wet soil and sparse grass strips by providing more opportunity for particles to settle and more runoff reduction.

References

Akram, S., Yu, B., Ghadiri, H., Rose, C., and Hussein, J. (2014). The links between water profile, net deposition and erosion in the design and performance of stiff grass hedges. *J. Hydrol.,* 510(0), 472-479. http://dx.doi.org/10.1016/j.jhydrol.2014.01.001

Arora, K., Mickelson, S.K., Baker, J.L., Tierney, D.P., and Peters, C.J. (1996). Herbicide retention by vegetative buffer strips from runoff under natural rainfall. *Trans. ASAE,* 39(6), 2155-2162. http://dx.doi.org/10.13031/2013.27719

Bharati, L., Lee, K.H., Isenhart, T.M., and Schultz, R.C. (2002). Soil-water infiltration under crops, pasture, and established riparian buffer in Midwestern USA. *Agrofor. Syst.,* 56(3), 249-257. http://dx.doi.org/10.1023/A:1021344807285

Bhattacharya, B., Price, R., and Solomatine, D. (2007). Machine learning approach to modeling sediment transport. *J. Hydraul. Eng.,* 133(4), 440-450. http://dx.doi.org/10.1061/(ASCE)0733-942

9(2007)133:4(440)

Blanco-Canqui, H., Gantzer, C.J., Anderson, S.H., Alberts, E.E., and Thompson, A.L. (2004). Grass barrier and vegetative filter strip effectiveness in reducing runoff, sediment, nitrogen, and phosphorus loss. *Soil Sci. Soc. Am. J.,* 68(5), 1670-1678. http://dx.doi.org/10.2136/sssaj2004.1670

Borin, M., Vianello, M., Morari, F., and Zanin, G. (2005). Effectiveness of buffer strips in removing pollutants in runoff from a cultivated field in North-East Italy. *Agric., Ecosyst. Environ.,* 105(1-2), 101-114. http://dx.doi.org/10.1016/j.agee.2004.05.011

Breiman, L. (1996). Bagging predictors. *Mach. Learning,* 24(2), 123-140. http://dx.doi.org/10.1007/BF00058655

Dabney, S.M., Meyer, L.D., Harmon, W.C., Alonso, C.V., and Foster, G.R. (1995). Depositional patterns of sediment trapped by grass hedges. *Trans. ASAE,* 38(6), 1719-1729. http://dx.doi.org/10.13031/2013.27999

Daniels, R.B., and Gilliam, J.W. (1996). Sediment and chemical load reduction by grass and riparian filters. *Soil Sci. Soc. Am. J.,* 60(1), 246-251. http://dx.doi.org/10.2136/sssaj1996.03615995006000010037x

Deletic, A. (2001a). Modelling of water and sediment transport over grassed areas. *J Hydrol.,* 248(1-4), 168-182. http://dx.doi.org/10.1016/S0022-1694(01)00403-6

Deletic, A. (2001b). Modelling of water and sediment transport over grassed areas. *J. Hydrol.,* 248(1-4), 168-182. http://dx.doi.org/10.1016/S0022-1694(01)00403-6

Deletic, A. (2005). Sediment transport in urban runoff over grassed areas. *J. Hydrol.,* 301(1-4), 108-122. http://dx.doi.org/10.1016/j.jhydrol.2004.06.023

Deletic, A., and Fletcher, T.D. (2006). Performance of grass filters used for stormwater treatment - A field and modelling study. *J. Hydrol.,* 317(3-4), 261-275. http://dx.doi.org/10.1016/j.jhydrol.2005.05.021

Dietterich, T. (2000). *Ensemble Methods in Machine Learning, Multiple Classifier Systems,* Multiple Classifier Systems, Springer Berlin Heidelberg, 1-15. http://dx.doi.org/10.1007/3-540-45014-9_1

Dillaha, T.A., Reneau, R.B., Mostaghimi, S., and Lee, D. (1989). Vegetative filter strips for agricultural nonpoint source pollution-control. *Trans. ASAE,* 32(2), 513-519. http://dx.doi.org/10.13031/2013.31033

Ghadiri, H., Hussein, J., and Rose, C.W. (2011). Effect of pasture buffer length and pasture type on runoff water quality following prescribed burning in the Wivenhoe Catchment. *Soil Res.,* 49(6), 513-522. http://dx.doi.org/10.1071/SR11122

Ghadiri, H., Rose, C.W., and Hogarth, W.L. (2001). The influence of grass and porous barrier strips on runoff hydrology and sediment transport. *Trans. ASAE,* 44(2), 259-268. http://dx.doi.org/10.13031/2013.4687

Hairsine, P.B., and Rose, C.W. (1992). Modeling water erosion due to overland-flow using physical principles: 1. Sheet flow. *Water Resour. Res.,* 28(1), 237-243. http://dx.doi.org/10.1029/91WR02380

Hall, J.K., Hertwig, N.L., and Hoffman, L.D. (1983). Application mode and alternate cropping effects on atrazine losses from a hillside. *J. Environ. Qual.,* 12(3), 336-340. http://dx.doi.org/10.2134/jeq1983.00472425001200030008x

Hong, W.C. (2008). Rainfall forecasting by technological machine learning models. *Appl. Math. Comput.,* 200(1), 41-57. http://dx.doi.org/10.1016/j.amc.2007.10.046

Hook, P.B. (2003). Sediment retention in rangeland riparian buffers. *J. Environ. Qual.,* 32(3), 1130-1137. http://dx.doi.org/10.2134/jeq2003.1130

Hussein, J., Ghadiri, H., Yu, B., and Rose, C. (2007a). Sediment retention by a stiff grass hedge under subcritical flow conditions. *Soil Sci. Soc. Am. J.,* 71(5), 1516-1523. http://dx.doi.org/10.213 6/sssaj2006.0276

Hussein, J., Truong, P., Ghadiri, H., Yu, B., and Rose, C. (2006). Vetiver buffer strips: Modelling their effect on sediment and nutrient reduction from surface flow, *Proc. of the 4th International Vetiver Conference: Vetiver and People,* Caracas, Venezuela.

Hussein, J., Yu, B., Ghadiri, H., and Rose, C. (2007b). Prediction of surface flow hydrology and sediment retention upslope of a vetiver buffer strip. *J. Hydrol.,* 338(3-4), 261-272. http://dx.doi.org/10. 1016/j.jhydrol.2007.02.038

Jin, C.X., and Romkens, M.J.M. (2001). Experimental studies of factors in determining sediment trapping in vegetative filter strips. *Trans. ASAE,* 44(2), 277-288.

Knighton, A.D. (1999). Downstream variation in stream power. *Geomorphology,* 29(3-4), 293-306. http://dx.doi.org/10.1016/S016 9-555X(99)00015-X

Le Bissonnais, Y., Lecomte, V., and Cerdan, O. (2004). Grass strip effects on runoff and soil loss. *Agronomie,* 24(3) 129-136. http://dx. doi.org/10.1051/agro:2004010

Loch, R.J., Espigares, T., Costantini, A., Garthe, R., and Bubb, K. (1999). Vegetative filter strips to control sediment movement in forest plantations: Validation of a simple model using field data. *Soil Res.,* 37(5), 929-946. http://dx.doi.org/10.1071/SR98089

Ma, L., Pan, C., Teng, Y., and Shangguan, Z. (2013). The performance of grass filter strips in controlling high-concentration suspended sediment from overland flow under rainfall/non-rainfall conditions. *Earth Surf. Process. Landforms,* 38(13), 1523-1534.

Magette, W.L., Brinsfield, R.B., Palmer, R.E., and Wood, J.D. (1989). Nutrient and sediment removal by vegetated filter strips. *Trans. ASAE,* 32(2), 663-667. http://dx.doi.org/10.13031/2013.31054

McDowell, R.W., Drewry, J.J., Muirhead, R.W., and Paton, R.J. (2003). Cattle treading and phosphorus and sediment loss in overland flow from grazed cropland. *Soil Res.,* 41(8), 1521-1532. http://dx.doi.org/10.1071/SR03042

McGregor, K.C., Dabney, S.M., and Johnson, J.R. (1999). Runoff and soil loss from cotton plots with and without stiff-grass hedges. *Trans. ASAE,* 42(2), 361-368. http://dx.doi.org/10.13031/2013.133 67

McKergow, L.A., Prosser, I.P., Grayson, R.B., and Heiner, D. (2004). Performance of grass and rainforest riparian buffers in the wet tropics, Far north Queensland: 2. Water quality. *Aust. J. Soil Res.,* 42(4), 485-498. http://dx.doi.org/10.1071/SR02156

Meyer, L.D., Dabney, S.M., and Harmon, W.C. (1995). Sediment-trapping effectiveness of stiff-grass hedges. *Trans. ASAE,* 38(3), 809-815. http://dx.doi.org/10.13031/2013.27895

Munoz-Carpena, R., Parsons, J.E., and Gilliam, J.W. (1999). Modeling hydrology and sediment transport in vegetative filter strips. *J. Hydrol.,* 214(1-4), 111-129. http://dx.doi.org/10.1016/S0022-1694(98)00272-8

Nash, J.E., and Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I - A discussion of principles. *J. Hydrol.,* 10(3), 282-290. http://dx.doi.org/10.1016/0022-1694(70)90255-6

Nearing, M.A. (1998). Why soil erosion models over-predict small soil losses and under-predict large soil losses. *Catena,* 32(1), 15-22. http://dx.doi.org/10.1016/S0341-8162(97)00052-0

Newham, L., Rutherford, J., and Croke, B. (2005). *A Conceptual Model of Particulate Trapping in Riparian Buffers,* CSIRO Land and Water Technical Report 21/05.

Newham, L.T.H., Weber, T.R., Baker-Finch, S.C., Post, D.A., and Rutherford, J.C. (2007). *Testing and Application of the Riparian Particulate Model,* Modsim 2007: International Congress on Modelling and Simulation 2109-2115.

Pan, C.Z., Ma, L., and Shangguan, Z.P. (2010). Effectiveness of grass strips in trapping suspended sediments from runoff. *Earth Surf. Process. Landforms,* 35(9), 1006-1013. http://dx.doi.org/10.100 2/esp.1997

Parsons, J.E., Gilliam, J.W., Munoz-Carpena, R., Daniels, R.B., and Dillaha, T.A. (1994). Nutrient and sediment removal by grass and riparian buffers, in: K.L.C.e. (Ed.), *Proc. of 2nd Environmentally Sound Agriculture Conference,* ASAE: St. Joseph, MI, pp. 147-154.

Patty, L., Réal, B., and Joël Gril, J. (1997). The use of grassed buffer strips to remove pesticides, nitrate and soluble phosphorus compounds from runoff water. *Pestic. Sci.,* 49(3), 243-251. http://dx.doi.org/10.1002/(SICI)1096-9063(199703)49:3<243::AI D-PS510>3.0.CO;2-8

Raffaelle, J.B., McGregor, K.C., Foster, G.R., and Cullum, R.F. (1997). Effect of narrow grass strips on conservation reserve land converted to cropland. *Trans. ASAE,* 40(6), 1581-1587. http://dx.doi.org/10.13031/2013.21422

Rankins, A., and Shaw, D.R. (2001). Perennial grass filter strips for reducing herbicide losses in runoff. *Weed Sci.,* 49(5), 647-651. http://dx.doi.org/10.1614/0043-1745(2001)049[0647:PGFSFR]2.0. CO;2

Razi, M.A., and Athappilly, K. (2005). A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert Syst. Appl.,* 29(1), 65-74. http://dx.doi.org/10.1016/j.eswa.2005.01.006

Robinson, C.A., Ghaffarzadeh, M., and Cruse, R.M. (1996). Vegetative filter strip effects on sediment concentration in cropland runoff. *J. Soil Water Conserv.,* 51(3), 227-230.

Schmitt, T.J., Dosskey, M.G., and Hoagland, K.D. (1999). Filter strip performance and processes for different vegetation, widths, and contaminants. *J. Environ. Qual.,* 28(5), 1479-1489. http://dx.doi. org/10.2134/jeq1999.00472425002800050013x

Schoonover, J.E., Williard, K.W.J., Zaczek, J.J., Mangun, J.C., and Carver, A.D. (2006). Agricultural sediment reduction by giant cane and forest riparian buffers. *Water Air Soil Pollut.,* 169(1-4), 303-315. http://dx.doi.org/10.1007/s11270-006-3111-2

Shipp, M.A., Ross, K.N., Tamayo, P., Weng, A.P., Kutok, J.L., Aguiar, R., Gaasenbeek, M., Angelo, M., Reich, M., Pinkus, G.S., Ray, T.S., Koval, M.A., Last, K.W., Norton, A., Lister, T.A., Mesirov, J., Neuberg, D.S., Lander, E.S., Aster, J.C., and Golu, T.R. (2002). Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat. Med.,* 8(1), 68-74. http://dx.doi.org/10.1038/nm0102-68

Shrestha, G., Stahl, P.D., and Ingram, L. (2005). Influence of Reclamation Management Practices on Soil Bulk Density and Infiltration Rates on Surface Coal Mine Lands in WYOMING. *National Meeting of the American Society of Mining and Reclamation,* Breckenridge, CO, USA.

Tadesse, L.D., and Morgan, R.P.C. (1996). Contour grass strips: A laboratory simulation of their role in erosion control using live grasses. *Soil Technol.,* 9(1-2), 83-89. http://dx.doi.org/10.1016/09 33-3630(95)00037-2

VanDijk, P.M., Kwaad, F.J.P.M., and Klapwijk, M. (1996). Retention of water and sediment by grass strips. *Hydrol. Process.,* 10(8), 1069-1080. http://dx.doi.org/10.1002/(SICI)1099-1085(199608)10: 8<1069::AID-HYP412>3.0.CO;2-4

Young, R.A., Huntrods, T., and Anderson, W. (1980). Effectiveness of vegetated buffer strips in controlling pollution from feedlot runoff. *J. Environ. Qual.,* 9(3), 483-487. http://dx.doi.org/10.213 4/jeq1980.00472425000900030032x