



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

PREDICTING THE PARTICLE SIZE DISTRIBUTION OF ERODED SEDIMENT USING ARTIFICIAL NEURAL NETWORKS

MARÍA PAZ LAGOS AVID

Thesis submitted to the Office of Research and Graduate Studies in partial fulfillment of the requirements for the Degree of Master of Science in Engineering

Advisor:

CARLOS BONILLA MELÉNDEZ

Santiago de Chile, December, 2016

© 2016, María Paz Lagos Avid



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

PREDICTING THE PARTICLE SIZE DISTRIBUTION OF ERODED SEDIMENT USING ARTIFICIAL NEURAL NETWORKS

MARÍA PAZ LAGOS AVID

Members of the Committee:

CARLOS BONILLA M.

RODRIGO CIENFUEGOS C.

FELIPE ABURTO G.

RICARDO GIESEN E.

Thesis submitted to the Office of Research and Graduate Studies in partial fulfillment of the requirements for the Degree of Master of Science in Engineering

Santiago de Chile, December, 2016

To my family, especially those who
are no longer here, and my friends
for their unconditional support.

ACKNOWLEDGEMENTS

First I want to thank my parents and my brother who supported me all this time. I also want to thank my high school friends who have accompanied me for many years, and my classmates and future colleagues over the years in the engineering school that made my college life something very important.

I express my deepest gratitude to my teacher guides Carlos Bonilla who believed in my abilities even before me. Without their guidance and support this study would not have been possible. I would also like to thank to the member of the research team of Professor Bonilla, because of their advice and collaboration help me to grow as a researcher.

This research was supported with funds from the National Commission for Scientific and Technological Research CONICYT/FONDECYT/Regular 1130928.

GENERAL INDEX

	Pág.
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	v
LIST OF FIGURES	vi
RESUMEN.....	vii
ABSTRACT	ix
1. INTRODUCTION	1
1.1. Overview	1
1.2. Hypothesis.....	6
1.3. Objective	6
1.4. Methodology	6
1.5. Results.....	7
1.6. Conclusions	7
2. LITERATURE REVIEW	8
3. MATERIALS AND METHODS	12
3.1. Soil database.....	12
3.2. Database analysis	17
3.3. Artificial neural network training and validation.....	17
3.4. Evaluation criteria	18
4. RESULTS AND DISCUSSION.....	19
4.1. Variables affecting the sediment particle size composition.....	19
4.2. Neural Network Model	20
4.3. Performance of the ANN model	24
4.4. Comparison with other sediment PSD models.....	27
5. CONCLUSIONS	28
REFERENCES.....	29

LIST OF TABLES

Table 1: Length, slope, rainfall amount and moisture content in the compiled soil database..	13
Table 2: Values used for normalization of input variables when building the artificial network model.	21
Table 3: Parameter values used for the artificial network model.....	22
Table 4: Maximum (y_{\max}) and minimum (y_{\min}) values used for each estimated soil fraction.....	23
Table 5: Evaluation of the artificial network model for predicting sediment particle size distribution.....	26

LIST OF FIGURES

Figure 1: Basic structure of an Artificial Neural Network.....	5
Figure 2: Soil texture class for the selected soil used in this study.....	14
Figure 3: Diagram of the artificial neural network model.	22
Figure 4: Comparison between the observed and estimated sediment PSD when using the artificial network model for training (a), validation (b), testing (c), and the entire database (d).....	25

RESUMEN

La erosión hídrica provoca la degradación del suelo y procesos de contaminación difusa. Los contaminantes son principalmente transportados sobre las superficies de partículas finas del suelo y del sedimento. Se han desarrollado varios modelos de pérdida de suelo y ecuaciones empíricas para la estimación de la distribución del tamaño del sedimento que abandona un terreno, incluyendo modelos en base física y ecuaciones empíricas. Por lo general los modelos de base física requieren una gran cantidad de datos, a veces superando la cantidad disponible en el área modelada. Por el contrario, las ecuaciones empíricas no siempre predicen la composición de los sedimentos asociados con eventos individuales y pueden requerir datos que no están disponibles. Por lo tanto, el objetivo de este estudio fue desarrollar un modelo para predecir la distribución del tamaño de partícula (DTP) del suelo erosionado. Se utilizaron 41 eventos de erosión provenientes de 21 suelos. Estos datos fueron compilados a partir de estudios previos. Se utilizó análisis de correlación y regresión múltiple para identificar las principales variables que controlan la DTP del sedimento. Estas variables fueron la distribución del tamaño de partícula en la matriz del suelo, la condición inicial de humedad del suelo, la erodibilidad del suelo y la geometría de la ladera. Con estas variables se calibró una red neural artificial utilizando datos de 29 eventos ($r^2 = 0,98, 0,97$ y $0,86$, para arena, limo y arcilla en el sedimento, respectivamente) y luego se validaron y probaron en 12 eventos ($r^2 = 0,74, 0,85$ y $0,75$, para arena, limo y arcilla en el sedimento, respectivamente). La red neural artificial se comparó con tres modelos empíricos. La red presentó un mejor desempeño en la predicción de DTP de sedimentos y la diferenciación de eventos de lluvia-escorrentía en el mismo suelo. Además de la calidad de las estimaciones de distribución de partículas, este modelo requiere un pequeño número de variables fácilmente obtenibles, proporcionando una rutina conveniente para predecir la DTP en sedimentos erosionados en otros modelos de transporte de contaminantes.

Palabras Claves: Distribución de tamaño de partículas, erosión hídrica, redes neuronales, sedimento.

ABSTRACT

Water erosion causes soil degradation and nonpoint source pollution. Pollutants are primarily transported on the surfaces of fine soil and sediment particles. Several soil loss models and empirical equations have been developed for the size distribution estimation of the sediment leaving the field, including physically-based models and empirical equations. Usually, physically-based models require a large amount of data, sometimes exceeding the amount of available data in the study area. Conversely, empirical equations do not always predict the sediment composition associated with individual events and may require data that are not available. Therefore, the objective of this study was to develop a model to predict the particle size distribution (PSD) of eroded soil. A total of 41 erosion events from 21 soils were used. This data was compiled from previous studies. Correlation and multiple regression analyses were used to identify the main variables controlling sediment PSD. These variables were the particle size distribution in the soil matrix, the antecedent soil moisture condition, soil erodibility, and hillslope geometry. With these variables, an artificial neural network was calibrated using data from 29 events ($r^2=0.98$, 0.97 , and 0.86 ; for sand, silt, and clay in the sediment, respectively) and then validated and tested on 12 events ($r^2=0.74$, 0.85 , and 0.75 ; for sand, silt, and clay in the sediment, respectively). The artificial neural network was compared with three empirical models. The network presented better performance in predicting sediment PSD and differentiating rain-runoff events in the same soil. In addition to the quality of the particle distribution estimates, this model requires a small number of easily obtainable variables, providing a convenient routine for predicting PSD in eroded sediment in other pollutant transport models.

Keywords: Neural networks; particle size distribution; sediment; water erosion.

1. INTRODUCTION

1.1. Overview

Erosion is the detachment and movement of soil or rock by water, wind, ice, or gravity. Water erosion is the most important type of soil degradation worldwide as it affects 56% of the total human-induced soil erosion (Oldeman, 1992). Soil erosion by water can be described as a three-step process: soil detachment, transport and deposition (Merritt, Letcher, and Jakeman, 2003). Detachment of the particles is caused by raindrop impact and/or runoff shear. Detached particles are transported by running water and deposited when the velocity of water decreases by the effect of slope or ground cover (Lal, 2001).

The consequence of water erosion are the soil loss (Morgan, 1995), decline in organic matter and nutrients (López et al., 2016; Morgan, 1995; Novotny and Chesters, 1989), and transport of contaminant such as pesticides, herbicides and heavy metals (Foster, Young, and Neibling, 1985; Gao, Magunhn, Spitzauer, and Kettrup, 1997; Selbig, Bannerman, and Corsi, 2013). The chemical transport capacity by sediment depends on its specific surface area (Deizman, Mostaghimi, Shanholtz, and Mitchell, 1987; Horowitz and Elrick, 1987; Young and Onstad, 1976), which in turn relates to the sediment particles size distribution. These particles are typically classified as sand (0.05-2.0 mm in diameter), silt (0.002-0.05 mm), clay (<0.002 mm), and aggregates (conglomerates of sand, silt and clay) (Soil Survey Division Staff, 1993). Particles with diameters <0.02 mm are particularly crucial for chemical transport because of their large surface area. Clay particles have the largest specific surface area, between $20 \text{ m}^2\text{g}^{-1}$ and $800 \text{ m}^2\text{g}^{-1}$ depending on the type of clay (Boonamnuyvitaya, Chaiya, Tanthapanichakoon, and Jarudilokkul, 2004; Slattery and Burt, 1997; Young and Onstad, 1976). In addition, fine particles are less prone to sedimentation, the contaminants adhered to there may be transported long distances, while sedimentation does not occur (Foster et al., 1985; Gabriels and Moldenhauer, 1978). Therefore, when

predicting the transport of soil-absorbed contaminants, it is necessary to use sediment particle size distribution (PSD) with an accurate assessment of the clay content (Foster et al., 1985; Meyer, Harmon, and McDowell, 1980).

Soil erosion can be estimated using: conceptual, empirical or physically based models. Conceptual models represent the watershed as storage systems. Empirical models are generally the simplest among the three types of model and are limited to those conditions for which they were developed. Physically based models are typically constructed by using mass conservation equations of sediment (Aksoy and Kavvas, 2005; Merritt et al., 2003).

Most current studies of water erosion have been made with models derived from the empirical Universal Soil Loss Equation (USLE) (Kinnell, 2010; Wischmeier and Smith, 1978) and its revised version RUSLE (Renard, Foster, Weesies, McCool, and Yoder, 1997). Models like the Erosion Productivity Impact Calculator (EPIC) (Sharpley and Williams, 1990) and the Soil Water Assessment Tool (SWAT) (Neitsch, Arnold, Kiniry, and Williams, 2011) use the USLE as a routine because it provides reasonable, long-term, annual soil loss estimates with little field information (Kinnell, 2010). However, USLE model does not consider the deposition of sediment in the modelled area (Merritt et al., 2003), and therefore does not calculate the composition of the sediment, which is necessary to estimate pollutants that bind preferentially to fine sediment particles (Aksoy and Kavvas, 2005).

Sediment particle size distribution can be estimated using multi-size erosion models, such as the Agricultural Non-Point Source Pollution (AGNPS) model (Young, Onstad, Bosch, and Anderson, 1989), the Areal Nonpoint Source Watershed Response Simulation (ANSWERS) model (Beasley, Huggins, and Monke, 1980), the Water Erosion Prediction Project (WEPP) model (Nearing, Foster, Lane, and Finkner, 1989), and the RUSLE2 model (Foster, 2008). AGNPS is a conceptual model (Merritt et al., 2003), whereas the others are physically based models, but all of them subdivide eroded soil into five particle size classes: clay, silt, sand, small aggregates and large aggregates. In RUSLE2 and WEPP, sediment particle composition at its point of detachment is

predicted with the equations developed by Foster et al. (1985) which use the soil matrix texture as input, an easily obtained parameter. Some of the previous models require a large amount of detailed information about climate, soil topography and land cover, and could exceed the available data in the modeled area.

An alternative to erosion models are empirical equations like Frere, Onstad, and Holtan (1975), Young and Onstad (1976), Young (1980) and Deizman et al. (1987). Frere et al. (1975) used texture information from 56 Midwest soils to develop a relationship between specific surface area and soil texture to estimate the particle size distribution of the eroded sediment. Young and Onstad (1976) used 45 Indiana soils and 30 Minnesota soils in addition to the Frere et al. (1975) data to develop a set of equations that require as input the particle size distribution of the soil matrix, organic matter, and water content at -15 bar pore pressure. Young (1980) built a database of 21 soils and developed three sets of empirical equations to approximate the undispersed particle size distribution of sediment from the dispersed matrix soil depending on the sediment size distribution of the matrix soil. Deizman et al. (1987) conducted 12 field experiments with a Groseclose silt loam soil to develop a set of equations that require as input the rainfall amount, slope, initial soil water content, and undispersed size distributions of the matrix soil. Some of the empirical equations listed above only consider soil properties, so they are unable to predict sediment particle size distribution based on rainfall, runoff or the size of erosion event.

Sediment particle size distribution is a function of soil properties, management, ground cover, slope, and detachment and transport processes (Alberts, Moldenhauer, and Foster, 1980; Alberts, Wendt, and Piest, 1983; Basic, Kisic, Nestroy, Mesic, and Butorac, 2002; Carkovic, Pastén, and Bonilla, 2015; Defersha and Melesse, 2012; Deizman et al., 1987; Foster et al., 1985; Gabriels and Moldenhauer, 1978; Kinnell, 2009; Martinez-Mena, Rogel, Albaladejo, and Castillo, 1999; Meyer, Harmon, and McDowell, 1980; Proffitt and Rose, 1991; Rienzi, Fox, Grove, and Matocha, 2013; Young, 1980; Zhang, Liu, Wang, and Wang, 2011a, 2011b; Ziadat and Taimeh, 2013). Many studies have been conducted to determine the sediment PSD and the factors

affecting distributions. Gabriels and Moldenhauer (1978) found that the sediment PSD had higher percentages of particles < 0.05 mm when the slope was less pronounced. Meyer et al. (1980) found that sediment PSD (1) did not vary significantly due to variations in the rainfall intensity and (2) was similar to the PSD of the matrix soil. Foster et al. (1985), based on the analysis of experimental data, concluded that the sediment sand content was directly related to sand in the matrix soil and inversely related to the clay content in the matrix soil. They also developed equations that describe the composition of sediment at its point of detachment. Later these equations have been tested and validated by other authors (Carkovic et al., 2015). Martinez-Mena et al. (1999) demonstrated that vegetal cover in natural plots reduces the energy available for water erosion. Similar results were obtained by Jin et al. (2009) and Zhang et al. (2011a). -They found that with the same ground cover condition, the fine fraction in the sediment decreased significantly when the rainfall intensity increased.

Defersha and Melesse (2012) conducted a series of laboratory experiments using simulated rainfall. They found that the effect of the slope and rainfall intensity on PSD varies with soil types and moisture contents. Similar results were obtained by Rienzi et al. (2013), indicating that sediment PSD depends on the antecedent moisture content.

Other tools such as artificial neural networks (ANNs) have been used to predict soil properties and soil related process (Baker and Ellison, 2008; Koekkoek and Booltink, 1999; Licznar and Nearing, 2003; Merdun, Çınar, Meral, and Apan, 2006; Tamari, Wösten, and Ruiz-Suárez, 1996; Wösten, Pachepsky, and Rawls, 2001). ANNs have a series of advantages, such as the ability to detect complex nonlinear relationships between dependent and independent variables, as its range of choices of structures of interconnections among components (Wösten et al., 2001). ANNs becomes complex formula in the relation between inputs with output values and, can be used like a regression formula (Wösten et al., 2001). An ANN consists of many interconnected simple computational elements called nodes or neurons. The outputs of neurons are used as input to other neurons in the network (Fig. 1). When the number of inputs is larger

than three, ANNs usually do better than regression techniques, so they are a good alternative for the development of empirical models.

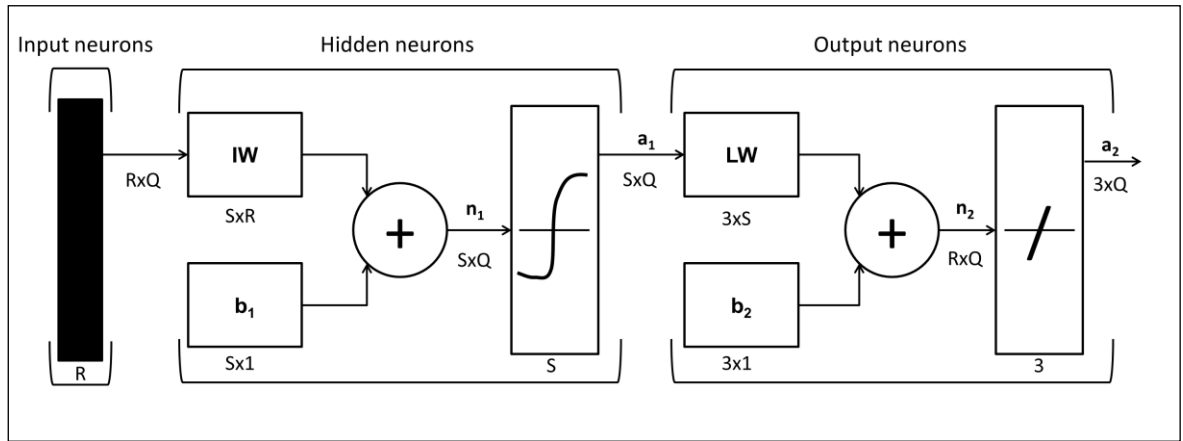


Figure 1: Basic structure of an Artificial Neural Network.

Many studies have been developed to identify and understand the factors controlling PSD. Most of the models developed in these studies use data from soil, slope, management, climate, cover, and irrigation/rainfall to estimate PSD (Defersha and Melesse, 2012; Gabriels and Moldenhauer, 1978; Martinez-Mena et al., 1999; Meyer et al., 1980; Rienzi et al., 2013; Zhang et al., 2011a). This information is not always available at the site of interest, precluding the use of these models in many applications. In contrast, the empirical equations assume very specific conditions in the matrix soil, and require that the PSD of the matrix soil is expressed as aggregate. Therefore, the objective of this study was to provide an empirical and more comprehensive equation for predicting the sediment PSD by using simple and typically measured soil properties.

1.2. Hypothesis

This research is based on the hypothesis that it is possible to develop an equation for estimating the sediment particle size distribution in the sediment in water erosion events based on parameters easily measured in zones with different soil characteristics, precipitation, and topography.

1.3. Objective

The overall objective of this study was to develop an equation to estimate the sediment particle size distribution. For this purpose, the following specific objectives were established:

- a) Review and identify the equations and models available for estimating the sediment particle size distribution.
- b) Build a database of sediment and soil, by collecting actual data from previous studies.
- c) Identify key variables that control and explain the particle size distribution in the sediment.
- d) Formulate and validate an equation to estimate the particle size distribution in the sediment using the actual soil database.

1.4. Methodology

A sediment and soil database was built based on existing experiments. The resulting database consists of 41 erosion events from 21 soils collected in five different experiments conducted in the United States, Belgium, China and Ethiopia. Correlation and multiple regression analysis were used to identify the main variables controlling sediment PSD.

The main parameters influencing the sediment PSD were used for developing an artificial neural network to estimate the sand, silt, and clay in the eroded sediment, and the effectiveness of the model was evaluated using the database constructed and compared with existing empirical equations.

1.5. Results

An artificial neural network was built to estimate the sand, silt, and clay in the eroded sediment. The input variables used were particle size distribution in the soil matrix, the antecedent soil moisture condition, soil erodibility, and hillslope geometry. This model estimated the sand, silt, and clay with r^2 of 0.93, 0.95 and 0.85, respectively. The artificial neural network was compared with the empirical equation of Frere et al. (1975), Young and Onstad (1976), and Deizman et al. (1987). The network showed a better performance when predicting sediment PSD, and differentiating rain-runoff events in the same soil.

1.6. Conclusions

The main conclusions of this study were:

- 1) The interaction between matrix soil composition, the ratio between soil water content and total porosity, the erodibility factor, and the hillslope vertical length can be used to explain the sediment PSD.
- 2) These interactions can be represented by an artificial neural network for predicting the sand, silt, and clay in the eroded sediment.
- 3) The network provides an alternative for estimating sediment composition when technical and/or economic resources are scarce, contributing to pollutant transport modeling and control.

2. LITERATURE REVIEW

Water erosion is one of the major causes of soil degradation (Comino et al., 2016; Lal, 2001; Morgan, 1995; Oldeman, 1992). The consequences of water erosion include soil loss (Morgan, 1995), decline in organic matter and nutrients (López et al., 2016; Morgan, 1995; Novotny and Chesters, 1989), and transport of contaminants such as many pesticides (Foster et al., 1985; Gao et al., 1997; Selbig et al., 2013). The chemical transport capacity by sediment depends on its specific surface area (Deizman et al., 1987; Horowitz and Elrick, 1987; Young and Onstad, 1976), which in turn depends on the sediment particles size distribution. These particles are typically classified as sand, silt, clay, and aggregates (conglomerates of sand, silt and clay).

Using the USDA size classification system (Soil Survey Division Staff, 1993), diameters for sand are between 0.05 and 2.0 mm, between 0.002 and 0.05 mm for silt, <0.002 mm for clay, and between 0.002 and 2.0 mm for aggregates. Particles with diameters <0.02 mm are particularly crucial for chemical transport because of their large surface area. Clay particles have the largest specific surface area, between $20 \text{ m}^2 \text{ g}^{-1}$ and $800 \text{ m}^2 \text{ g}^{-1}$ depending on the type of clay (Boonamnuayvitaya et al., 2004; Slattery and Burt, 1997; Young and Onstad, 1976). Therefore, when predicting the transport of soil-absorbed contaminants, it is necessary to use sediment particle size distribution (PSD) with an accurate assessment of the clay content (Foster et al., 1985; Meyer et al., 1980).

Sediment particle size distribution can be estimated using multi-size erosion models, such as the Agricultural Non-Point Source Pollution (AGNPS) model (Young et al., 1989), the Areal Nonpoint Source Watershed Response Simulation (ANSWERS) model (Beasley et al., 1980), the Water Erosion Prediction Project (WEPP) model (Nearing et al., 1989), and the Revised Universal Soil Loss Equation version 2 (RUSLE2) model (Foster, 2008). AGNPS is a conceptual model (Merritt et al., 2003), whereas the others are physical models, but all subdivide eroded soil into five particle size classes: clay, silt, sand, small aggregates and large aggregates. ANSWERS, RUSLE2 and WEPP assume that the detached sediment particle size distribution is the

same as the matrix soil, and the deposition of these particles is selective for each. Some of these models require a large amount of input data, which can exceed available data in the modeled area.

Another way to estimate sediment particle size is by using empirical equations. Frere et al. (1975) used texture information from 56 Midwest soils to develop a relationship between specific surface area and soil texture to estimate the particle size distribution of the eroded sediment. In this study, the author assumed a specific surface area for each particle size. Young and Onstad (1976) used 45 Indiana soils and 30 Minnesota soils in addition to the Frere et al. (1975) data to develop a set of equations considering organic matter content and clay mineralogy. These equations require as input the particle size distribution of the soil matrix, organic matter, and water content at -15 bar pore pressure. Young (1980) built a database of 21 soils and developed three sets of empirical equations to approximate the undispersed particle size distribution of sediment from the dispersed matrix soil depending on the sediment size distribution of the matrix soil. Deizman et al. (1987) conducted 12 field experiments with a Groseclose silt loam soil using a rainfall simulator with an intensity of 50 mmh^{-1} in three runs. The plots were divided into conventional and no-tillage systems with slope from 8.5% to 9.7%. The results of the experiments showed that the rainfall amount, slope, initial soil water content, and undispersed size distributions of the matrix soil explain the behavior of the sediment PSD. Using these variables Deizman et al. (1987) developed empirical equations to describe the undispersed and dispersed size distributions of sediment from no-till and conventional tillage methods.

Some of the empirical equations listed above only consider soil properties, so they are unable to predict sediment particle size distribution based on rainfall, runoff or size of erosion event. The assumptions used in the empirical equations and data arrangement required may limit their applicability to other soils and soil conditions.

Sediment particle size distribution is a function of soil properties, management, cover, slope, and detachment and transport processes (Carkovic et al., 2015; Defersha and Melesse, 2012; Deizman et al., 1987; Foster et al., 1985; Gabriels and Moldenhauer,

1978; Kinnell, 2009; Martinez-Mena et al., 1999; Meyer et al., 1980; Young, 1980; Zhang et al., 2011b, 2011a). Many studies have been conducted to determine sediment PSD and factors affecting distributions. Gabriels and Moldenhauer (1978) conducted a series of experiments on four soils from Ames, Iowa, and two Belgian soils, A and B, using simulated rainfall of 63.5 mmh^{-1} intensity and a duration of 90 minutes to assess the effect of soil texture and rainfall intensity on sediment size distribution. They found that the sediment PSD had higher percentages of particles $< 0.05 \text{ mm}$ when the slope was less pronounced. Meyer et al. (1980) conducted a series of field experiments on 10 soils with slopes of 8% to less than 1% using a simulated rainfall of 67 mm h^{-1} intensity for one hour in order to compare sediment size distributions. They found that sediment PDS (1) did not vary significantly due to variations in the rainfall intensity and (2) was similar to the PSD of the matrix soil. Foster et al. (1985), based on the analysis of experimental data, concluded that the sediment sand content was directly related to sand in the matrix soil and inversely related to the clay content in the matrix soil. They also developed equations that describe the composition of sediment at its point of detachment.

Martinez-Mena et al. (1999) demonstrated that vegetal cover in natural plots reduces the energy available for water erosion. Similar results were obtained by Zhang et al. (2011a) conducting field experiments on a sandy loam soil under simulated rainfall with three intensities (60, 100, and 140 mm h^{-1}) for 60 minutes each and three cover percentages (0%, 30% and 80%) with a 15% slope to investigate the effect of rainfall intensity and vegetation cover on sediment PSD. Additionally, they found that with the same cover condition, the fine fraction in the sediment decreased significantly when the rainfall intensity increased.

Defersha and Melesse (2012) conducted laboratory experiments using simulated rainfall of 120, 70, and 55 mmh^{-1} intensity applied sequentially for 90 minutes with 9%, 25% and 45% slopes for three soil types that varied from clay to sandy clay loam to evaluate the effect of rainfall intensity, slope, soil types and antecedent moisture content on sediment PSD. They found that the effects of slope and rainfall intensity on PSD vary

with soil types and moisture contents. Similar results were obtained by Rienzi et al. (2013), indicating that sediment PSD depends on the antecedent moisture content.

Many studies have been developed to identify and understand the factors controlling PSD. Most models developed in these studies use data from soil, slope, management, climate, cover, and irrigation/rainfall to estimate PSD. This information is not always available at the site of interest, precluding the use of these models in many applications. In contrast, the empirical equations assume very specific conditions in the matrix soil, and require that the PSD of the matrix soil is expressed as aggregate. Other tools such as artificial neural networks (ANNs) have been used to predict soil properties and soil related process (Baker and Ellison, 2008; Koekkoek and Booltink, 1999; Licznar and Nearing, 2003; Merdun et al., 2006; Tamari et al., 1996; Wösten et al., 2001). ANNs have several advantages, such as the ability to detect complex nonlinear relationships between dependent and independent variables, as its range of choices of structures of interconnections among components (Wösten et al., 2001). ANNs have a complex formula in the relationship between inputs and output values (Maren, Harston, and Pap, 1990) and can be used similar to a regression formula (Wösten et al., 2001).

The goal of this study is to provide an empirical and more comprehensive equation for predicting sediment PSD by using simple and typically measured soil properties. With this purpose, a sediment and soil database was compiled based on existing studies, and correlation and multiple regression analyses were used to identify the main variables controlling sediment PSD. With these variables, an artificial neural network was built to estimate the sand, silt, and clay in the eroded sediment. The effectiveness of the model was evaluated using the constructed database and compared with existing empirical equations.

3. MATERIALS AND METHODS

3.1. Soil database

A sediment and soil database was built using experiments conducted by Gabriels and Moldenhauer (1978) and data reported by Meyer et al. (1980), Deizman et al. (1987), Zhang et al. (2011a), and Defersha and Melesse (2012). These studies report the dispersed PSD of the matrix soil and sediment, organic matter content, hillslope length and slope, rainfall intensity and duration, and initial soil moisture conditions. Gabriels and Moldenhauer (1978) did not report the organic matter content of Belgian soils, and Meyer et al. (1980) did not report the initial soil moisture condition and the organic matter content for all soils. However, these missing data were estimated, as explained below, for utilization in this study.

The resulting database consists of 41 rainfall events from 21 soils collected in five different experiments conducted in the United States, Belgium, China, and Ethiopia. As shown in Table 1, the hillslope lengths ranged from 0.45 m to 18.3 m, with slopes from 0.5% to 45% and simulated rainfall from 25 mm to 245 mm applied in the experiments. The initial soil moisture conditions were oven dried, air dried, and pre-wetted. The studied soils had organic matter ranging from 0.8% to 14%, 1% to 88% in sand, 8% to 84% in silt, and 4% to 55% in clay (Fig. 1).

Table 1: Length, slope, rainfall amount and moisture content in the compiled soil database. Data for soils 1-6 are from Gabriels and Moldenhauer (1978), data for soils 7-16 are from Meyer et al. (1980), data for soil 17 are from Deizman et al. (1987), data for soil 18 are from Zhang et al. (2011a), and data for soils 19-21 are from Defersha and Melesse (2012). Soils 6 and 17 to 22 have more than one plot configuration.

	Soil	Hillslope length (m)	Hillslope slope (%)	Rainfall amount (mm)	Initial moisture
1	Haganer	0.45	9.0	95	Air dry
2	Ida	0.45	9.0	95	Air dry
3	Marshall	0.45	9.0	95	Air dry
4	Lutton	0.45	9.0	95	Air dry
5	Belgian silt loam A	0.55	31.0	81	Air dry
6	Belgian silt loam B	0.55	9.0, 31.0	81	Air dry
7	Grenada #1	0.90	1.0	67	NA
8	Cascilla	0.90	1.0	67	NA
9	Sharkey	0.90	0.5	67	NA
10	Bruin	0.90	0.5	67	NA
11	Vicksburg V	0.90	1.0	67	NA
12	Lexington	0.90	3.5	67	NA
13	Arkabutla	0.90	1.0	67	NA
14	Grenada #8	0.90	1.0	67	NA
15	Memphis	0.90	6.5	67	NA
16	Morganfield	0.90	0.5	67	NA
17	Groseclose silt loam soil	18.30	9.0	50, 25, 25	Dry-Wet-Very wet
18	Sandy loam soil	8.50	15.0	60, 100, 140	Dry
19	Alemaya black soil	0.45	9.0, 25.0, 45.0	245	Air dry- Pre wetted
20	Godie soil series	0.45	9.0, 25.0, 45.0	245	Air dry- Pre wetted
21	Alemaya series eroded	0.45	9.0, 25.0, 45.0	245	Air dry- Pre wetted

NA: Not available

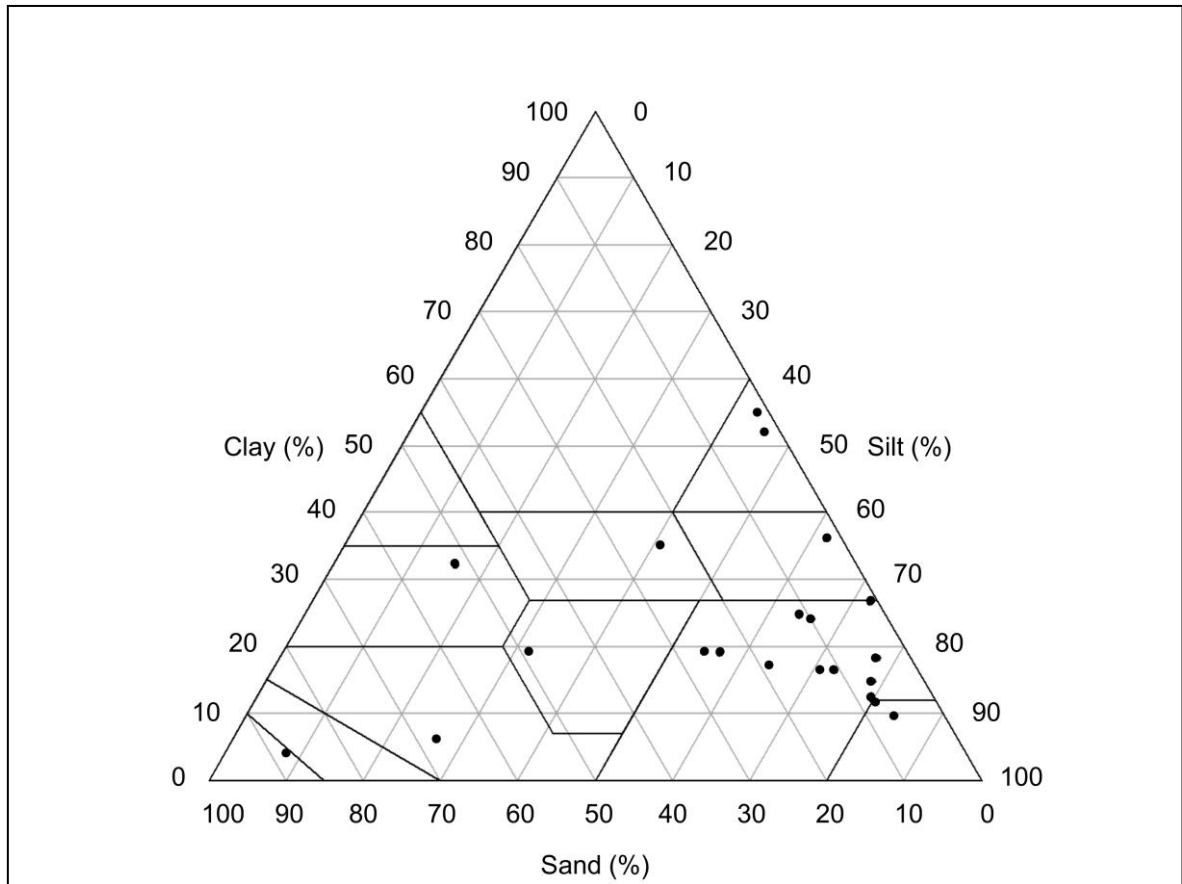


Figure 2: Soil texture class for the selected soil used in this study.

The literature review identified the main factors controlling sediment PSD. These factors were PSD of the matrix soil, initial soil water content, rainfall intensity and amount, and slope. Because not all these factors were provided by the authors, some estimation was necessary. In addition, hillslope vertical length, erodibility factor, ratio between soil water content and total porosity, and rainfall erosivity of the storm were estimated.

To analyze the effect of the slope and plot dimension, the correlation with sediment PSD and slope, hillslope length, and vertical length (θ_v) were computed. The hillslope vertical length was defined as follows:

$$\theta_v = \sin(\theta)\lambda \quad (1)$$

where θ and λ are the slope (rad) and length (m) of hillslope. The θ_v value provides a measurement of the hillslope steepness.

The erodibility factor (K) is a key factor for predicting the average annual soil loss in most erosion models, such as RUSLE (Renard et al., 1997), and it was included in our analysis as an input parameter. The value of K was computed as the erodibility factor developed in the Erosion Productivity Impact Calculator (EPIC) model (Sharpley and Williams, 1990), in which K is computed as follows:

$$K = (0.2 + 0.3 \cdot \exp[-0.0256 \cdot SAN \cdot (1 - SIL/100)]) \cdot \left(\frac{SIL}{CLA + SIL} \right)^{0.3} \cdot \left(1 - \frac{0.25 \cdot C}{C + \exp[3.72 - 2.95 \cdot C]} \right) \cdot \left(1 - \frac{0.7 \cdot SN_1}{SN_1 + \exp[22.9 \cdot SN_1 - 5.51]} \right) / 7.59 \quad (2)$$

where K is the erodibility factor ($t \ h \ MJ^{-1} \ mm^{-1}$); SAN , SIL , CLA , and C are the sand, silt, clay, and soil organic carbon contents (%), respectively; and $SN_I = 1 - SAN/100$. The C

value was computed from the organic matter content reported by the authors. The organic matter was assumed to be 2% for the 10 soils of Meyer et al. (1980), and for the 2 Belgian soils of Gabriels and Moldenhauer (1978). These 12 soils correspond to 13 events of the 41 events in the database.

To evaluate the effect of initial water content in the sediment PSD, the ratio (S) between the soil water content and total porosity was calculated as follows:

$$S = \frac{\theta}{\phi} \quad (3)$$

where θ is the soil water content at the beginning of the rainfall event ($\text{m}^3 \text{m}^{-3}$) and ϕ is the total porosity ($\text{m}^3 \text{m}^{-3}$). When the soil condition at the beginning of the experiment was described as pre-wetted, θ was assumed to be the content at -33 kPa. Conversely, when the soil was oven dried, θ was assumed to be the content at -1500 kPa, and when the soil condition was air dried, θ was assumed to be the average between -33 and -1500 kPa. When the initial soil water content was not reported, it was assumed to be the content at -33 kPa. All variables were obtained from Table 2 of Rawls et al. (1982) based on soil texture.

The effect of the rainfall was analyzed in three different ways: rainfall amount (mm), maximum 30-minute rainfall intensity I_{30} (mm h^{-1}), and storm rainfall erosivity R ($\text{MJ mm ha}^{-1} \text{h}^{-1}$). Erosivity is used in most erosion models and was included in the analysis because it combines the effects of duration, magnitude and intensity of each rainfall event (Foster, 2008; Lobo and Bonilla, 2015).

3.2. Database analysis

To determine both the dependence and redundant effects between the database variables, a correlation analysis with the Pearson correlation coefficient (r) was used. In addition, a multiple regression analysis was performed between the sediment PSD and soil and climate properties. The results of the analysis were used to determine input variables when building the neural network.

3.3. Artificial neural network training and validation

ANNs were tested with a two-layer feed forward network (with a hidden and an output layer) using the Levenberg-Marquardt back propagation algorithm as a training function (Hagan and Menhaj, 1994). The selected transfer function was hyperbolic tangent sigmoid, and the performance function was mean square error (MSE). The database was randomly divided in three sets: 70% of the data was used for training, 15% was used for validation during the training process, and 15% of the data set was used to test the ability of the trained neural network to predict new data. The number of neurons in the hidden layer was between 1 and 10. It is important to note that different neural networks trained on the same problem can give different outputs for the same input (Tamari et al., 1996). For this reason, a total of 100 iterations were performed for each number of tested neurons. The target of the ANNs was the sediment PSD as the dispersed fractions of sand, silt, and clay. All simulations were performed with the neural network toolbox of MATLABTM (The MathWorks Inc., USA).

Because of the random division of the database in the training, validation and testing, and the random selection of initial parameters values, using the network with minimum MSE between all iterations is not necessarily the best solution, as it could be removing appropriate information by rejecting networks with higher MSE (Perrone and Cooper, 1993; Tumer and Ghosh, 1996). Therefore, the criteria for choosing the best ANN were as follows: (1) predicted values are in the range of sediment PSD, (2) MSE

for both training and validation is minimal, and (3) as explained by Tamari et al. (1996), the maximum and minimum values of selected input variables are part of the training set.

As a result of the ANN structure, researchers have considered them “black boxes” (Olden and Jackson, 2002). The contribution of input variables in predicting output value is difficult to separate within the network, so the dependencies between variables and modeling mechanism are not explained by the network. Thus, to determine the relative importance of each input in the network, Garson’s algorithm was used (Garson, 1991). The procedure is described in detail by Olden and Jackson (2002); essentially, after selecting the best ANN, the relative importance (RI) for each input variable was calculated.

3.4. Evaluation criteria

The root mean square error (RMSE), coefficient of determination (r^2), and Nash-Sutcliffe model efficiency (NS) (Nash and Sutcliffe, 1970) were used to compare the observed and predicted sediment PSD. The Nash-Sutcliffe model efficiency was calculated as follows:

$$NS = 1 - \frac{\sum (O_i - P_i)^2}{\sum (O_i - \bar{O})^2} \quad (4)$$

where O_i and P_i are observed and predicted PSD, and \bar{O} is the observed mean of PSD.

4. RESULTS AND DISCUSSION

4.1. Variables affecting the sediment particle size composition

The correlation analysis identified four variables related to sediment PSD. Sand content in sediment showed a direct relationship with its content in the soil matrix ($r=0.91$). Sediment sand content was inversely related to silt content in the soil matrix and K factor value ($r=-0.74$ and -0.71 , respectively). In contrast to the results reported by Foster et al. (1985), no relationship between the sand in the sediment and clay in the soil matrix was found.

The silt content in sediment shows a direct relationship with its content in the soil matrix and with the K value ($r=0.90$ and $r=0.89$, respectively). The sediment silt content is inversely related to the sand content in the matrix, rainfall, and erosivity ($r=-0.77$, -0.71 , and -0.72 , respectively). The inverse relationship between the silt in the sediment and the rainfall is due to 18 events from Defersha and Melesse (2012), which have the same rainfall. If the analysis is performed without these events, no relationship is found between these variables ($r=-0.43$). Similar results were found for Deizman et al. (1987), where no trend was found when correlating rainfall and sediment PSD. Sediment clay content shows a direct correlation with its content in the soil matrix ($r=0.74$).

The analysis shows a correlation between sediment PSD and PSD in the matrix, rainfall, erosivity and K factor value. Therefore, they were used as input variables for developing the neural network model. Conversely, S and θ_v did not show a correlation with sediment PSD. However, S affects soil infiltration, and it is a good estimator of the total runoff volume, which is a key component of modeling soil loss and detachment of soil particles during a runoff event (Defersha and Melesse, 2012; Morgan, 1995). In addition, the slope and length of the plot affect soil loss in water erosion processes (Morgan, 1995). Deizman et al. (1987) found that as the slope decreased, the percentage of sediment PSD increased. Gabriels and Moldenhauer (1978) explained that the lower runoff velocity from lower slopes leaves the soil exposed to raindrop impacts for a

longer time, so particles are more available for erosion. The θ_v , which connect the two variables in a single parameter, could affect sediment PSD as S , because both variables were included as input values in network construction.

4.2. Neural Network Model

The ANN model developed for the prediction of sediment particle size distribution corresponds to a three stages' unified model: pre-processing, neural network, and post-processing. This model contains 8 neurons and 6 input variables, which are sand, silt, and clay in the matrix and S , K and θ_v .

The pre-processing step corresponds to a normalization of input variables contained in the range $[-1, 1]$. This was accomplished using the formula:

$$xp = \frac{(xp_{\max} - xp_{\min})}{(x_{\max} - x_{\min})}(x - x_{\min}) + xp_{\min} \quad (5)$$

where xp is the normalized input variable, xp_{\min} and xp_{\max} are the minimum and maximum value for xp equal to -1 and 1, respectively, x is the input variable (mass fraction), and x_{\min} and x_{\max} are the minimum and maximum value for the input variable (mass fraction). The value for x_{\min} and x_{\max} are given in Table 2.

Table 2: Values used for normalization of input variables when building the artificial network model.

	Sand	Silt	Clay	S	K (ton h MJ ⁻¹ mm ⁻¹)	θ_v
	(%)					
x_{min}	1	8	4	14	0.107	0.004
x_{max}	88	84	55	81	0.421	1.640

The normalized input xp_k ($k=1$ to 6 for sand, silt and clay in the matrix, S , K and θ_v , respectively) for the neuron j ($j=1$ to 8) are multiplied by weights IW_{jk} and summed together with the constant bias term b_j^1 . The resulting n_j^1 is the input for the hyperbolic tangent sigmoid function, which produces the output a_j^1 . At the same time, as shown in Fig. 3, a_j^1 are multiplied by layer weights LW_{ij} and summed together with the constant term b_i^2 to form the output a_i^2 ($i = 1$ to 3 for sand, silt and clay, respectively). The neural network equation is:

$$a_i^2 = \sum_{j=1}^8 LW_{ij} \text{tansig} \left(\sum_{k=1}^6 IW_{jk} xp_k + b_j^1 \right) + b_i^2 \quad (6)$$

where the values of LW , IW , b^1 and b^2 are shown in Table 3.

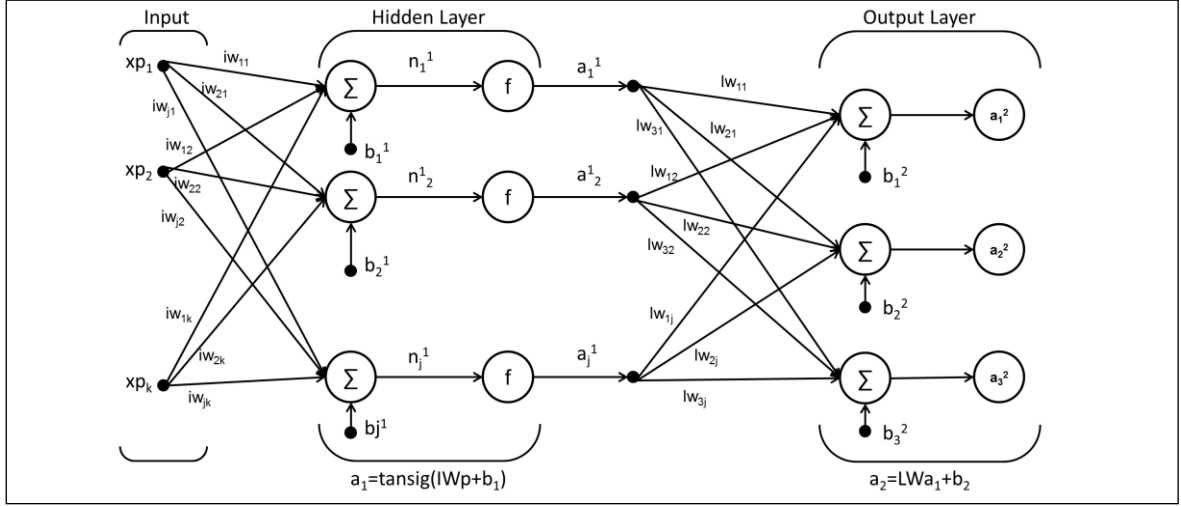


Figure 3: Diagram of the artificial neural network model.

Table 3: Parameter values used for the artificial network model.

IW								
$j \backslash k$	1	2	3	4	5	6	7	8
1	-0.711	-0.928	0.626	1.302	-1.480	-0.950	2.290	1.384
2	-0.746	-0.323	0.686	0.100	0.605	-0.681	0.015	-1.225
3	-1.432	1.601	0.042	-1.554	0.134	1.071	-0.341	1.266
4	-0.474	-0.512	0.542	-0.959	1.321	2.096	0.890	0.826
5	1.001	0.440	1.518	0.859	-0.399	-0.270	-1.137	0.604
6	1.221	0.879	1.693	1.752	0.412	2.016	0.656	-1.680
b^1								
j	1	2	3	4	5	6	7	8
	1.753	-1.253	-0.745	0.133	0.453	-0.044	1.236	1.472
LW								
$i \backslash j$	1	2	3	4	5	6	7	8
1	-0.208	0.530	-0.330	1.094	0.838	-0.602	0.646	0.911
2	-0.668	-0.609	1.487	-0.425	-0.387	-0.278	-0.680	-0.034
3	0.957	0.044	-1.314	-0.904	-0.598	1.059	-0.011	-1.141
b^2								
i	1			2			3	
	-0.873			0.931			-0.784	

The network output a^2 is normalized and must be reverse-processed to obtain the estimates values for each particle. This is accomplished using the formula:

$$y = \frac{(y_{\max} - y_{\min})}{(a_{\max}^2 - a_{\min}^2)} (a^2 - a_{\min}^2) + y_{\min} \quad (7)$$

where y is the estimated value (mass fraction), y_{\max} and y_{\min} are the maximum and minimum values (mass fraction) of y given in Table 4, a_{\max}^2 and a_{\min}^2 are 1 and -1, respectively, and a^2 is the network output of Eq. (6).

Table 4: Maximum (y_{\max}) and minimum (y_{\min}) values used for each estimated soil fraction.

	Sand	Silt (%)	Clay
y_{\min}	0.7	8.0	6.0
y_{\max}	86.0	86.6	70.7

Analysis of the relative importance of each input variable shows that sand and vertical length are the most important inputs to the model (RI=22%). As shown above, the sand content in the soil matrix was correlated with the contents of silt and sand in the sediment. However, the vertical length had no correlation with sediment PSD. This could be an effect of the complex interaction with the erosion factor; therefore, it is important to measure precisely the geometry of the hillslope to obtain a reliable estimate using the model. The least important variable was silt content in the matrix (RI=10%). Consequently, its value may be less accurate with respect to other variables. The other input variables have an RI=14%-17%. Based on results of the relative importance analysis of input variables, it is possible to conclude that inputs have a similar RI in the

ANN model (average RI=17%), and the process responsible for sediment particle distribution cannot be explained with a single variable.

4.3. Performance of the ANN model

The ANN model estimated the sediment PSD for the database with NS=0.96 and RMSE=0.04 for the training set, NS=0.91 and RMSE=0.05 for the validation set, NS=0.74 and RMSE=0.07 for the test set, NS=0.85 and RMSE=0.06 for the validation and test set together, and NS=0.94 and RMSE=0.05 for the total database (Fig. 4). Results for individual particles are presented in Table 5. The ANN model for an individual data set (test) overestimates sand in the sediment for one event and underestimates sand for one event. Similarly, clay in the sediment is underestimated in one event (Fig. 4.c). The event that overestimates the sand and underestimates the clay corresponds to the same event from the Alemaya series (Defersha and Melesse, 2012), with a 9% of slope and pre-wetting. This event has no particular characteristic making it different from other events associated with the Alemaya series, but it was the only event of this soil that was not used for network training. For this particular event, the silt fraction was well estimated. In contrast, the event that underestimated sand also overestimated the silt, and this event corresponds to an event from the Godie soil series (Defersha and Melesse, 2012), with a 25% of slope and air-dried soil. For this soil, 4 events were used for training, one for validation and one for testing. If the analysis was performed without these events, the r^2 and NS increased to 0.98 and 0.97 for sand, 0.97 and 0.96 for silt, and 0.99 and 0.98 for clay, and 0.98 and 0.98 for the test set, respectively. This result indicates that the ANN model predicts sediment PSD fairly well and provides a reliable tool for predicting sediment pollutant transport capacity, which depends on specific surface area (Horowitz and Elrick, 1987). Clay is predicted with lower precision with respect to other particles but is still a good estimate and is crucial for a pollutant transport model, as clay has the largest specific surface area (Boonamnuayvitaya et al., 2004).

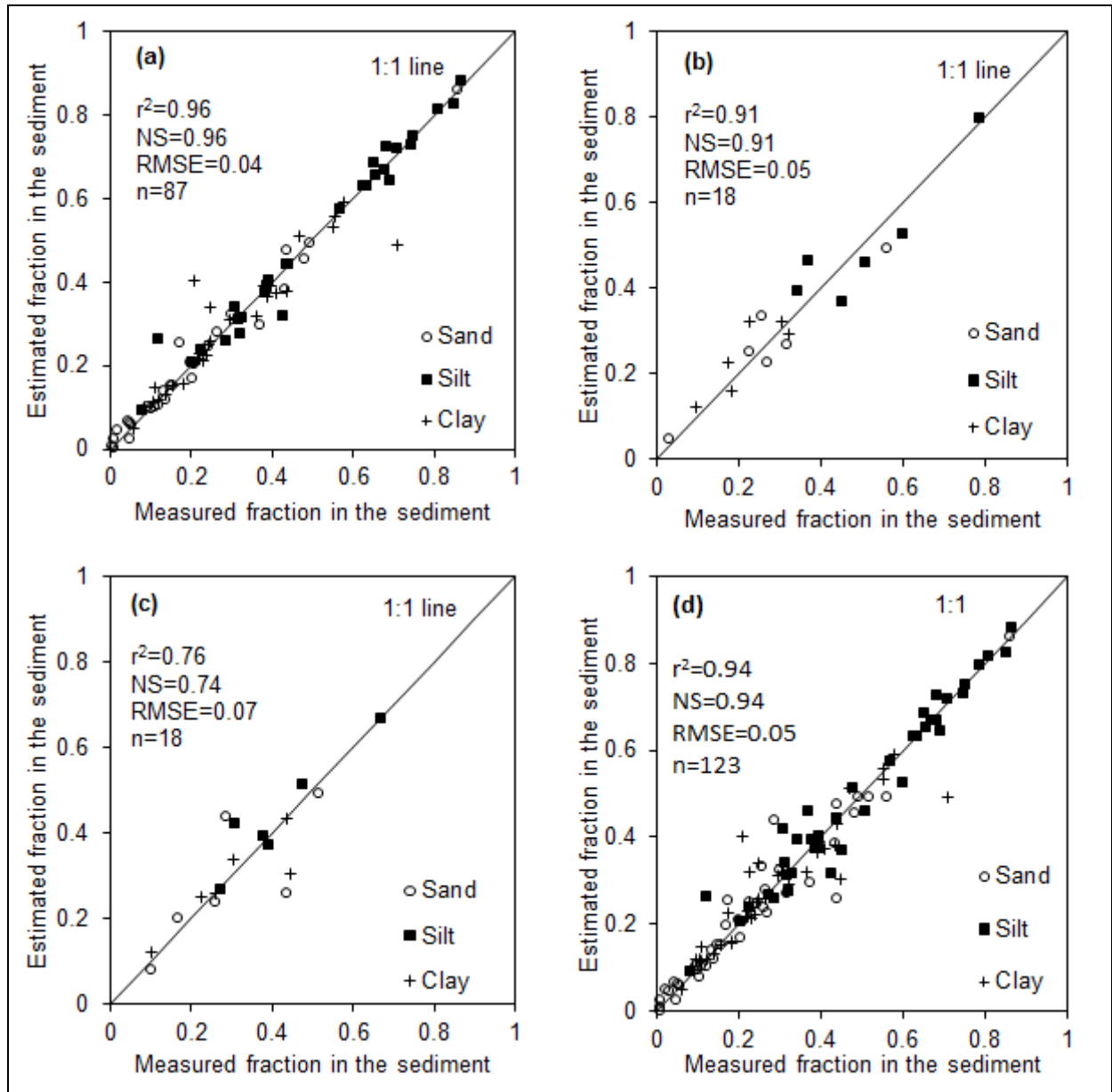


Figure 4: Comparison between the observed and estimated sediment PSD when using the artificial network model for training (a), validation (b), testing (c), and the entire database (d).

Table 5: Evaluation of the artificial network model for predicting sediment particle size distribution.

		Sand	Silt	Clay
Train	r^2	0.98	0.97	0.86
	NS	0.98	0.97	0.86
	RMSE	0.03	0.04	0.06
	n	29	29	29
Validation	r^2	0.91	0.82	0.74
	NS	0.89	0.82	0.63
	RMSE	0.05	0.06	0.05
	n	6	6	6
Test	r^2	0.59	0.89	0.76
	NS	0.54	0.86	0.74
	RMSE	0.10	0.05	0.06
	n	6	6	6
All	r^2	0.93	0.95	0.85
	NS	0.93	0.95	0.85
	RMSE	0.05	0.04	0.06
	n	41	41	41

To develop the ANN model, it was necessary to assume an organic matter content equal to 2% for Meyer et al. (1980) soils and the Gabriels and Moldenhauer (1978) soils. A value of 60% of the Meyer et al. (1980) soils and 33 % of the Gabriels and Moldenhauer (1978) soils were used to train the ANN, which represents 31% of the training data set. A change in the organic matter of 1% causes the ANN model to underestimate the average of the sand in 4% and overestimate the average of the silt and clay in 18% and 14%, respectively. Conversely, a 4% change in the organic matter content causes the ANN to underestimate sand average in 1% and overestimate the average of silt and clay in 10% and 10%, respectively. Despite these results for both 1% and 4% of organic matter content, the model predicted the trends in sediment PSD. Finally, it is important to mention that the ANN model was developed using as a domain

the values presented in Table 2; consequently, the neural network does not work a priori when the input values are out of the range of the table.

4.4. Comparison with other sediment PSD models

The empirical equations of Frere et al. (1975), Young and Onstad (1976), and Deizman et al. (1987) were evaluated using the database and compared with the results of the ANN. The Frere et al. (1975) and Young and Onstad (1976) equations were evaluated using the entire database. They estimated the sediment PSD with $r^2=0.64$, and 0.76 , $NS=0.63$, and 0.75 , and $RMSE=0.13$ and 0.11 , respectively. The equation of Young and Onstad (1976) predicted the PSD more accurately than that of Frere et al. (1975) due to the incorporation of the organic matter content and water content as equation input variables. However, both equations were unable to estimate the PSD for soil with more than one event. For the Deizman et al. (1987) equations, the numbers of events available in the database were not enough for a correct analysis. However, for the 3 events available, the r^2 was 0.98 , $NS=0.98$, and $RMSE=0.03$. Among the three studies, the empirical equations developed by Deizman et al. (1987) predicted the sediment PSD more accurately. These results are explained by the three selected events used to test Deizman's equation corresponding to events used by Deizman et al. (1987) for equation development.

The ANN model more accurately estimates sediment particle size distribution compared to empirical equations developed by Frere et al. (1975) and Young and Onstad (1976) ($r^2=0.94$, $NS=0.94$, and $RMSE=0.05$), and compared to Deizman et al. (1987) ($r^2=0.99$, $NS=0.99$, and $RMSE=0.004$), it responds to different events in the same soil and does not require aggregate particle information.

5. CONCLUSIONS

A direct relationship between sediment particle size distribution (PSD) and the particle size distribution in the matrix soil was found. PSD in the matrix soil can be used as a first approach for predicting sediment PSD from a water erosion event. Moreover, a reliable relationship between sediment PSD and rainfall, rainfall erosivity and erodibility factor was found. This study demonstrates that there is an interaction between matrix soil composition, the ratio between soil water content and total porosity, the erodibility factor, and the hillslope vertical length that can be used to explain sediment PSD.

The use of artificial neural networks proved to be suitable for building a suitable model for predicting sediment PSD in soil loss events across a wide range of soils and hillslope geometries. The developed ANN is a two-layer feed forward network with 8 neurons and 6 inputs, which is capable of simultaneously predicting dispersed particle size distributions in sediments.

The ANN model developed in this study is meant to be used as a simple predictor for sand, silt and clay fractions in the sediment. Although physical models can provide more accurate results, this model requires fewer input values, which are obtainable from field and soil surveys, and is easily coupled with other sediment/contaminant models. Finally, this ANN model provides an alternative for estimating sediment composition when technical and/or economic resources are scarce, contributing to pollutant transport modeling and control.

REFERENCES

- Aksoy, H., and Kavvas, M. L. (2005). A review of hillslope and watershed scale erosion and sediment transport models. *Catena*, 64(2), 247–271.
- Alberts, E. E., Moldenhauer, W. C., and Foster, G. R. (1980). Soil aggregates and primary particles transported in rill and interrill Flow. *Soil Science Society of America Journal*, 44, 590–595.
- Alberts, E. E., Wendt, R. C., and Piest, R. F. (1983). Physical and chemical properties of eroded soil aggregates. *Transactions of the ASAE*, 26(2), 465–471.
- Baker, L., and Ellison, D. (2008). Optimisation of pedotransfer functions using an artificial neural network ensemble method. *Geoderma*, 144(1), 212–224.
- Basic, F., Kistic, I., Nestroy, O., Mesic, M., and Butorac, A. (2002). Particle size distribution (texture) of eroded soil material. *Journal of Agronomy and Crop Science*, 188(5), 311–322.
- Beasley, D. B., Huggins, L. F., and Monke, E. J. (1980). ANSWERS: A Model for Watershed Planning. *Transactions of the ASAE*, 23(4), 938–944.
- Boonamnuyvitaya, V., Chaiya, C., Tanthapanichakoon, W., and Jarudilokkul, S. (2004). Removal of heavy metals by adsorbent prepared from pyrolyzed coffee residues and clay. *Separation and Purification Technology*, 35(1), 11–22.
- Carkovic, A. B., Pastén, P. A., and Bonilla, C. A. (2015). Sediment composition for the assessment of water erosion and nonpoint source pollution in natural and fire-affected landscapes. *Science of the Total Environment*, 512–513, 26–35.
- Comino, J. R., Iserloh, T., Lassu, T., Cerdà, A., Keesstra, S. D., Prosdocimi, M., ... Ries, J. B. (2016). Quantitative comparison of initial soil erosion processes and runoff generation in Spanish and German vineyards. *Science of the Total Environment*, 565, 1165–1174.
- Defersha, M. B., and Melesse, A. M. (2012). Effect of rainfall intensity, slope and antecedent moisture content on sediment concentration and sediment enrichment ratio. *Catena*, 90, 47–52.

- Deizman, M. M., Mostaghimi, S., Shanholtz, V. O., and Mitchell, J. K. (1987). Size distribution of eroded sediment from two tillage systems. *Transactions of the ASAE*, 30(6), 1642–1647.
- Foster, G. R. (2008). Revised Universal Soil Loss Equation Version 2 (RUSLE2). *Science Documentation, USDA-Agricultural Research Service*, 2(Versión 2), 1–355.
- Foster, G. R., Young, R. A., and Neibling, W. H. (1985). Sediment composition for nonpoint source pollution analyses. *Transactions of the ASAE*, 28, 133–139, 146.
- Frere, M. H., Onstad, C. A., and Holtan, H. N. (1975). ACTMO, an agricultural chemical transport model. *U.S. Agricultural Research Service. ARS-H-3*, 1–11.
- Gabriels, D., and Moldenhauer, W. C. (1978). Size distribution of eroded material from simulated rainfall: effect over a range of texture. *Soil Science Society of America Journal*, 42, 954–958.
- Gao, J. P., Magunhn, J., Spitzauer, P., and Kettrup, A. (1997). Distribution of pesticides in the sediment of the small Teufelsweher pond (southern Germany). *Water Research*, 31(11), 2811–2819.
- Garson, G. D. (1991). Interpreting neural-network connection weights. *AI Expert*, 6(4), 46–51.
- Hagan, M. T., and Menhaj, M. B. (1994). Training feedforward networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, 5(6), 989–993.
- Horowitz, A. J., and Elrick, K. A. (1987). The relation of stream sediment surface area, grain size and composition to trace element chemistry. *Applied Geochemistry*, 2(4), 437–452.
- Jin, K., Cornelis, W. M., Gabriels, D., Baert, M., Wu, H. J., Schiettecatte, W., ... Hofman, G. (2009). Residue cover and rainfall intensity effects on runoff soil organic carbon losses. *Catena*, 78(1), 81–86.
- Kinnell, P. I. A. (2009). The influence of raindrop induced saltation on particle size distributions in sediment discharged by rain-impacted flow on planar surfaces. *Catena*, 78(1), 2–11.
- Kinnell, P. I. A. (2010). Event soil loss, runoff and the Universal Soil Loss Equation family of models: A review. *Journal of Hydrology*, 385(1–4), 384–397.

- Koekkoek, E. J. W., and Booltink, H. (1999). Neural network models to predict soil water retention. *European Journal of Soil Science*, 50, 489–495.
- Lal, R. (2001). Soil degradation by erosion. *Land Degradation and Development*, 12(6), 519–539.
- Licznar, P., and Nearing, M. A. (2003). Artificial neural networks of soil erosion and runoff prediction at the plot scale. *Catena*, 51(2), 89–114.
- Lobo, G. P., and Bonilla, C. A. (2015). Sensitivity analysis of kinetic energy-intensity relationships and maximum rainfall intensities on rainfall erosivity using a long-term precipitation dataset. *Journal of Hydrology*, 527, 788–793.
- López, P., López-Tarazón, J. A., Casas-Ruiz, J. P., Pompeo, M., Ordoñez, J., and Muñoz, I. (2016). Sediment size distribution and composition in a reservoir affected by severe water level fluctuations. *Science of the Total Environment*, 540, 158–167.
- Maren, A. J., Harston, C. T., and Pap, R. M. (1990). *Handbook of neural computing applications*. San Diego: Academic Press.
- Martinez-Mena, M., Rogel, J. A., Albaladejo, J., and Castillo, V. M. (1999). Influence of vegetal cover on sediment particle size distribution in natural rainfall conditions in a semiarid environment. *Catena*, 38(3), 175–190.
- Merdun, H., Çınar, Ö., Meral, R., and Apan, M. (2006). Comparison of artificial neural network and regression pedotransfer functions for prediction of soil water retention and saturated hydraulic conductivity. *Soil and Tillage Research*, 90(1–2), 108–116.
- Merritt, W. S., Letcher, R. A., and Jakeman, A. J. (2003). A review of erosion and sediment transport models. *Environmental Modelling & Software*, 18(8–9), 761–799.
- Meyer, L. D., Harmon, W. C., and McDowell, L. L. (1980). Sediment sizes eroded from crop row sideslopes. *Transactions of the ASAE*, 23(4), 891–898.
- Morgan, R. P. C. (1995). *Soil Erosion and Conservation* (Second Edi). Essex, England: Longman Group Limited.
- Nash, J. E., and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models. *Journal of Hydrology*, 10(3), 282–290.
- Nearing, M. A., Foster, G. R., Lane, L. J., and Finkner, S. C. (1989). A process-based

soil erosion model for USDA-Water Erosion Prediction Project Technology. *Transactions of the ASAE*, 32(5), 1587–1593.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., and Williams, J. R. (2011). Soil and Water Assessment Tool: Theoretical documentation version 2009. *Texas Water Resources Institute*, 1–647.

Novotny, V., and Chesters, G. (1989). Delivery of sediments and pollutants from nonpoint sources: a water quality perspective. *Journal of Soil and Water Conservation*, 44(28), 568–576.

Oldeman, L. R. (1992). The global extent of soil degradation. In *Soil Resilience and Sustainable Land Use* (pp. 99–118). Wallingford: CAB International.

Olden, J. D., and Jackson, D. A. (2002). Illuminating the “black box”: A randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, 154(1–2), 135–150.

Perrone, M. P., and Cooper, L. N. (1993). When networks disagree: ensemble methods for hybrid neural network. In R. J. Mammone (Ed.), *Neural Networks for Speech and Image Processing* (pp. 126–142). London: Chapman and Hall.

Proffitt, A. P. B., and Rose, C. W. (1991). Soil erosion processes. II. Settling velocity characteristics of eroded sediment. *Australian Journal of Soil Research*, 29(5), 685–695.

Rawls, W. J., Brakensiek, D. L., and Saxton, K. E. (1982). Estimation of Soil Water Properties. *Transactions of the ASAE*, 25(5), 1316–1320.

Renard, K. G., Foster, G. R., Weesies, G. A., McCool, D. K., and Yoder, D. C. (1997). *Predicting soil erosion by water: A guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE)*. USDA Agriculture Handbook (Vol. 703). Washington, D.C.

Rienzi, E. A., Fox, J. F., Grove, J. H., and Matocha, C. J. (2013). Interrill erosion in soils with different land uses: The kinetic energy wetting effect on temporal particle size distribution. *Catena*, 107, 130–138.

Selbig, W. R., Bannerman, R., and Corsi, S. R. (2013). From streets to streams: Assessing the toxicity potential of urban sediment by particle size. *Science of the Total Environment*, 444, 381–391.

- Sharpley, A. N., and Williams, J. R. (1990). Erosion Productivity Impact Calculator: 1. Model Documentation (EPIC). *Technical Bulletin-United State Department of Agriculture*.
- Slattery, M. C., and Burt, T. P. (1997). Particle size characteristics of suspended sediment in hillslope runoff and stream flow. *Earth Surface Processes and Landforms*, 22(8), 705–719.
- Soil Survey Division Staff. (1993). *Soil Survey Manual*. Soil Conservation Service. U.S. Department of Agriculture Handbook 18.
- Tamari, S., Wösten, J. H. M., and Ruiz-Suárez, J. C. (1996). Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Science Society of America Journal*, 60(6), 1732–1741.
- Tumer, K., and Ghosh, J. (1996). Error correlation and error reduction in ensemble classifiers. *Connection Science*, 8(3–4), 385–404.
- Wischmeier, W. H., and Smith, D. D. (1978). Predicting rainfall erosion losses: a guide to conservation planning. *Agriculture Handbook No. 537*, (537), 285–291.
- Wösten, J. H. M., Pachepsky, Y. A., and Rawls, W. J. (2001). Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251, 123 – 150.
- Young, R. A. (1980). Characteristics of eroded sediment. *Transactions of the ASAE*, 23(3), 1139–1142, 1146.
- Young, R. A., and Onstad, C. A. (1976). Predicting particle-size composition of eroded soil. *Transactions of the ASAE*, 19(6), 1071–1075.
- Young, R. A., Onstad, C. A., Bosch, D. D., and Anderson, W. P. (1989). AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. *Journal of Soil and Water Conservation*, 44(2), 168–173.
- Zhang, G. H., Liu, G. B., Wang, G. L., and Wang, Y. X. (2011a). Effects of vegetation cover and rainfall intensity on sediment-associated nitrogen and phosphorus losses and particle size composition on the Loess Plateau. *Journal of Soil and Water Conservation*, 66(3), 192–200.
- Zhang, G. H., Liu, G. B., Wang, G. L., and Wang, Y. X. (2011b). Effects of vegetation

cover and rainfall intensity on sediment-bound nutrient loss, size composition and volume fractal dimension of sediment particles. *Pedosphere*, 21(5), 676–684.

Ziadat, F. M., and Taimeh, A. Y. (2013). Effect of rainfall intensity, slope, land use and antecedent soil moisture on soil erosion in an arid environment. *Land Degradation and Development*, 24(6), 582–590.