Spatial Interpolation of Rainfall Erosivity Using Artificial Neural Networks for Southern Brazil Conditions

Michel Castro Moreira(1)*, Thiago Emanuel Cunha de Oliveira(2), Roberto Avelino Cecílio(3), Francisco de Assis de Carvalho Pinto(4) and Fernando Falco Pruski(4)

(1) Universidade Federal do Oeste da Bahia, Centro das Ciências Exatas e das Tecnologias, Barreiras, Bahia, Brasil.
(2) Gapso - Supply Chain Management, Belo Horizonte, Minas Gerais, Brasil.
(3) Universidade Federal do Espírito Santo, Departamento de Engenharia Florestal, Alegre, Espírito Santo, Brasil.
(4) Universidade Federal de Viçosa, Departamento de Engenharia Agrícola, Viçosa, Minas Gerais, Brasil.

ABSTRACT: Water erosion is the process of disaggregation and transport of sediments, and rainfall erosivity is a numerical value that expresses the erosive capacity of rain. The scarcity of information on rainfall erosivity makes it difficult or impossible to use to estimate losses occasioned by the erosive process. The objective of this study was to develop Artificial Neural Networks (ANNs) for spatial interpolation of the monthly and annual values of rainfall erosivity at any location in the state of Rio Grande do Sul, and a software that enables the use of these networks in a simple and fast manner. This experiment used 103 rainfall stations in Rio Grande do Sul and their surrounding area to generate synthetic rainfall series on the software ClimaBR 2.0. Rainfall erosivity was determined by summing the values of the EI30 and KE >25 indexes, considering two methodologies for obtaining the kinetic energy of rainfall. With these values of rainfall erosivity and latitude, longitude, and altitude of the stations, the ANNs were trained and tested for spatializations of rainfall erosivity. To facilitate the use of the ANNs, a computer program was generated, called netErosividade RS, which makes feasible the use of ANNs to estimate the values of rainfall erosivity for any location in the state of Rio Grande do Sul.

Keywords: erosive potential of rainfall, soil conservation, universal soil loss equation.
INTRODUCTION

The use of physical and mathematical models to describe the erosion process has evolved from the 1950s with the advent of the Universal Soil Loss Equation (USLE) (Pruski, 1996). Developed by Wischmeier and Smith (1958), the USLE is an empirical model that allows estimation of soil losses and identification of the factors that have the greatest effect on them; among these factors is rainfall erosivity (R), which expresses the erosive capacity of rain.

Soil loss caused by rainfall in farmed areas is highly correlated with the product of two characteristics of rainfall: total kinetic energy and maximum intensity in 30 min (Wischmeier and Smith, 1958). This product was designated EI$_{30}$ and it has been able to explain 72 to 97 % of soil loss caused by rains. However, Gonçalves et al. (2006) point out that authors such as Hudson (1973) and Lal (1988) found that the EI$_{30}$ does not correlate well with soil losses in tropical regions, proposing an alternative method for these regions, termed KE $>$ 25, which is the sum of kinetic energy of rains with intensity of more than 25 mm h$^{-1}$.

However, taking Brazilian conditions into account, calculation of the EI$_{30}$ and KE $>$ 25 erosivity indexes becomes difficult, due to the small number of locations that have sufficient rain gauge records to enable rainfall estimation. Moreover, the data analysis process is quite slow and laborious (Bertoni and Lombardi Neto, 1999), which makes information related to rainfall erosivity available for only a small number of locations in the country.

To estimate the values of rainfall erosivity in places where not even rainfall records are available, authors such as Rufino et al. (1993), Bertoni and Lombardi Neto (1999), Silva (2004), and Mello et al. (2012) have used interpolation techniques to estimate the values of R.

For interpolation of rainfall erosivity, Silva (2004) and Gonçalves et al. (2006) used the interpolation method based on Inverse Distance Weighting. However, this method does not take into account the altitude of the site but rather the amount of rainfall erosivity of other locations considering an inversely proportional relationship to the distance between this site and each of the neighboring stations. An improvement in obtaining the R values could be achieved if interpolation methods took into account the altitude (Goovaerts, 1999; Gonçalves et al., 2006; Moreira et al., 2006), which is possible when using Artificial Neural Networks (ANN).

An ANN consists of a set of computing elements called artificial neurons, and its development aims to determine its architecture, that is, the number of layers and neurons in each layer, as well as to fit their free parameters $w$ and $b$, a phase known as training. The architecture of ANN varies according to the complexity of the problem and cannot be defined before training, constituting a search based on trial and error (Hagan et al., 1996).

Given the importance of knowing rainfall erosivity for designing the use of management practices for erosion control and the difficulty that technicians who work in soil and water conservation have in acquiring it, Moreira et al. (2008) developed a computer program using ANNs to estimate rainfall erosivity values in the state of Minas Gerais.

Due to the need for studies and subsidies for obtaining the rainfall erosivity index in Brazil, the objective of this study was to develop artificial neural networks for spatial interpolation of the monthly and annual values of rainfall erosivity at any location within the state of Rio Grande do Sul, as well as to develop a computer program that enables use of the network in a fast and simple manner.
MATERIALS AND METHODS

Obtaining rainfall erosivity

Considering the small number of rainfall stations in the region under study and the lack of sufficient rain gauge records to allow estimation of rainfall erosivity, synthetic series with 100-yr duration were generated (on a daily basis) for 94 rain gauge stations located in the state of Rio Grande do Sul from pluviograph information available on ClimaBR 2.0 software (Zanetti, 2003; GPRH, 2005), according to the method proposed by Cecílio et al. (2013). Moreover, a series of nine more stations located in the state of Santa Catarina were generated to increase the amount of data in the border region between those two states.

The synthetic series produced by ClimaBR 2.0 has the following information associated with rainfall: total daily rainfall, event duration, maximum instantaneous precipitation intensity, time of occurrence of the maximum intensity, and parameters that characterize the rainfall profile.

The monthly amount of rainfall erosivity for each station was obtained by the sum of the values of the EI30 or KE >25 erosivity indexes of erosive rainfalls in each month, which were obtained according to the procedures described in Wischmeier and Smith (1958) and Hudson (1973), respectively. Two methods for obtaining the kinetic energy of rainfall were used for calculating the indexes (Foster et al., 1981; Wagner and Massambani, 1988), and thus, four values for rainfall erosivity were obtained for each rainfall station.

Development of the artificial neural networks

For training the ANNs, 88 stations were randomly selected, and the remaining 15 were used for testing. Artificial neural networks were used with feedback and type 3-n1-n2-1 architecture, which consisted of an input vector with three variables, two intermediate layers with n1 and n2 artificial neurons, and one neuron in the output layer.

The input vector was made up of the latitude and longitude values of each station in decimal degrees, as well as its altitude value in meters. In the neuron of the output layer, a linear activation function was used to provide the rain erosivity value of the location represented by the input vector, in MJ mm h⁻¹ ha⁻¹ yr⁻¹ for the estimate considering the EI30 or MJ ha⁻¹ yr⁻¹ for the estimate considering the KE >25.

To ensure that each input parameter received equal attention during training, thus increasing parameter efficiency, both the input and the output data were standardized for the range between -1 and 1 (Maier and Dandy, 2000).

The ANNs were trained using the Error Backpropagation Learning Algorithm. In each interaction of this algorithm, the w and b parameters are updated by a training rule; that of Levenberg-Marquardt was adopted in this case. For development of the ANNs, different combinations of number of neurons, activation functions in the intermediate layers, and iteration numbers were tested. The number of neurons tested, 12, was limited by the number of samples used for training, as proposed by Hagan et al. (1996).

Since, at the beginning of the training, the free parameters are randomly generated, and these initial values may influence the final result of the training, the ANN representing each combination of parameters was trained 20 times. Of the 20 trained ANNs, the one with the highest value of the coefficient of determination (R²) was stored. The coefficient of determination (R²) was calculated from erosivity data from 15 stations of the test sample and those estimated by ANN.
In addition to the $R^2$, the confidence index (c) proposed by Camargo and Sentelhas (1997), which is calculated by multiplying the correlation coefficient (r) by the index agreement (d) proposed by Willmott (1981), was used in evaluation of the results obtained with the ANNs developed.

**Software development**

Considering that use of the ANNs developed requires a high-cost computer program and adequate training, a computer program that enables use of these networks in a user-friendly environment was developed.

Having acquired the ANNs, it was necessary to know their respective architectures, neuron activation functions, and the $w$ and $b$ free parameters to generate the mathematical functions that represent them. The $w$ and $b$ free parameters of ANNs were exported in text files, in which all the information relating to ANNs was stored.

Thus, with the data from architectures, neuron activation functions, and the $w$ and $b$ free parameters, it was possible to structure the mathematical equations representing the ANNs and then implement them computationally by using the Borland Delphi 7.0 programming environment.

The ANNs developed allow estimation of monthly values of rainfall erosivity, and a routine that calculates the annual rainfall erosivity from the sum of the monthly values was implemented.

Since ANNs require the altitude value to carry out calculation of rainfall erosivity for a given location, the database of the altimetry of the state of Rio Grande do Sul was incorporated into the computer program. The database was obtained from the GTOPO30 Project, which has spatial resolution of 1 km and has been developed worldwide by the United States Geological Survey (USGS).

To verify the adequacy of rainfall erosivity values estimated by the developed software, the erosivity values estimated by the software were compared with those found in the literature for different localities in the state of Rio Grande do Sul (Table 1). The comparative analysis was performed using the Relative Error (ER%), which was calculated by the equation

\[
ER\% = \frac{|R_o - R_i|}{R_o} \times 100
\]

where

$ER\% = \text{relative error, dimensionless}$; $R_o = \text{value obtained in the literature for rainfall erosivity, MJ mm h}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$; and $R_i = \text{value estimated by the software for rainfall erosivity, MJ mm h}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$.

**RESULTS AND DISCUSSION**

**Artificial neural networks**

The number of neurons in the intermediate layer and training cycles of the artificial neural networks developed for monthly estimate of rainfall erosivity in the state of Rio Grande do Sul are shown in Table 2, according to the $EI_{30}$ and $KE >25$ erosivity index and the methods for obtaining kinetic energy of rainfall proposed by Foster et al. (1981) and Wagner and Massambani (1988).

Considering that the networks developed could have a maximum number of 12 neurons, the architectures of the networks developed exhibited a median number of neurons, and the network for the month of June, considering the method of Foster et al. (1981),
had the smallest number of neurons (three in all). Although the number of neurons of the first and second layers ranged from one to six, none of ANNs developed had more than nine neurons in their layers.

Smaller architectures have greater generalization ability, while larger architectures have the capacity for faster learning, that is, they require fewer training cycles in their development (Lippmann, 1987). An architecture with minimal error is achieved by increasing the number of neurons or by increasing the number of training cycles up to a limit where there are no “memory” issues (Peng and Wen, 1999), which is characterized by having a low error in the training sample, but with a trend of increasing the error in the test sample.

Table 1. Rainfall erosivity estimated for different localities of the state of Rio Grande do Sul

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Data periods(1)</th>
<th>Latitude (south)</th>
<th>Longitude (west)</th>
<th>Altitude</th>
<th>Rainfall erosivity</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hulha Negra</td>
<td>1956-1984</td>
<td>31° 20'</td>
<td>54° 06'</td>
<td>214</td>
<td>6,209</td>
<td>Martins et al. (2009)</td>
</tr>
<tr>
<td>Ijuí</td>
<td>1963-1993</td>
<td>28° 23'</td>
<td>53° 54'</td>
<td>448</td>
<td>8,825</td>
<td>Cassol et al. (2008a)</td>
</tr>
<tr>
<td>Quaraí</td>
<td>1966-2003</td>
<td>30° 23'</td>
<td>56° 26'</td>
<td>100</td>
<td>9,292</td>
<td>Penalva et al. (2007)</td>
</tr>
<tr>
<td>Santa Maria</td>
<td>1963-2000</td>
<td>29° 41'</td>
<td>53° 48'</td>
<td>153</td>
<td>7,866</td>
<td>Cogol et al. (2006)</td>
</tr>
<tr>
<td>Santa Rosa</td>
<td>1975-2003</td>
<td>27° 51'</td>
<td>54° 29'</td>
<td>273</td>
<td>11,217</td>
<td>Mazurana et al. (2009)</td>
</tr>
<tr>
<td>São Borja</td>
<td>1956-2003</td>
<td>28° 39'</td>
<td>56° 00'</td>
<td>99</td>
<td>9,751</td>
<td>Cassol et al. (2008b)</td>
</tr>
<tr>
<td>Uruguaiana</td>
<td>1963-1991</td>
<td>29° 45'</td>
<td>57° 05'</td>
<td>74</td>
<td>8,875</td>
<td>Hickmann et al. (2008)</td>
</tr>
</tbody>
</table>

(1) Period of rain gauge data used to calculate the rainfall erosivity.

Table 2. Number of neurons in the intermediate layers and training cycles of the artificial neural networks developed for monthly estimate of rainfall erosivity in the state of Rio Grande do Sul according to the $E_{30}$ and KE $>$ 25 erosivity index and the methods for obtaining the kinetic energy of rainfall (F and WM)

<table>
<thead>
<tr>
<th>Month</th>
<th>$E_{30}$</th>
<th>$KE &gt;25$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>WM</td>
</tr>
<tr>
<td></td>
<td>$n_1$</td>
<td>$n_2$</td>
</tr>
<tr>
<td>Jan</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Feb</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Mar</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Apr</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>May</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Jun</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Jul</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Aug</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Sept</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Oct</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Nov</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Dec</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

F: Kinetic energy calculated by the method proposed by Foster et al. (1981). WM: Kinetic energy calculated by the method proposed by Wagner and Massambani (1988). $n_1$ and $n_2$: Number of neurons in the first and second intermediate layer, respectively.
Networks trained by the Levenberg-Marquardt training rule achieved better performances with fewer training cycles, and this behavior was attributed to the fact that networks with more than 200 training cycles had “memory” issues in most months. Because the training rule used in the study was the most efficient for convergence to the minimum error as a function of the number of training cycles (Hagan et al., 1996), low values were found for the number of cycles obtained, and the maximum number of training cycles was observed in only five ANNs.

The architectures of the ANNs showed median values for number of neurons. The network related to the month of September showed the greatest number of neurons (11 in all), considering the method proposed by Wagner and Massambani (1988).

For the number of training cycles, only six ANNs were trained with the maximum number of cycles, thus demonstrating behavior similar to that observed for the ANNs developed considering the EI\textsubscript{30} erosivity index.

The values of the coefficient of determination (R\textsuperscript{2}), confidence index (c), and performance rating (D) calculated from the 15 stations of the test sample using the artificial neural networks developed for the monthly estimate of rain erosivity in the state of Rio Grande do Sul according to the EI\textsubscript{30} and KE >25 erosivity indexes and the methods of obtaining kinetic energy proposed by Foster et al. (1981) and Wagner and Massambani (1988) are presented in table 3.

For EI\textsubscript{30}, of the 24 ANNs developed, corresponding to the 12 months of the year and to the two methods of estimating kinetic energy, nine were classified as “Excellent”, 10 as “Very good”, and only five as “Good”. The R\textsuperscript{2} ranged from 0.77 to 0.93, in which the ANNs classified as “Excellent” ranged from 0.86 to 0.92, those classified as “Very good” ranged from 0.77 to 0.93, and those classified as “Good” ranged from 0.77 to 0.86. Thus, it can be seen that the R\textsuperscript{2} values corroborated the performance rating proposed by Camargo and Sentelhas (1997).

Overall, for the ANNs developed there was no distinction, in terms of performance, for the erosivity estimate considering the two methods for obtaining the kinetic energy value proposed by Foster et al. (1981) and Wagner and Massambani (1988).

For KE >25, the results of the performance rating showed that out of the 24 ANNs developed, five were classified as “Excellent”, nine as “Very good”, and 10 as “Good” (Table 3). Ten ANNs developed for estimation of erosivity were rated as “Good”, compared with only five ANNs according to the EI\textsubscript{30} index. This result is associated with greater complexity in fitting ANNs to estimate erosivity considering the KE >25 erosivity index in relation to the EI\textsubscript{30} index.

Considering the R\textsuperscript{2} values and performance ratings, the ANNs are efficient in estimating rainfall erosivity since among the 48 ANNs developed, 14 were rated as “Excellent”, and 19 obtained a rating of “Very good” and 15 were rated as “Good”. There were no ratings below “Good”, considering the Camargo and Sentelhas (1997) rating.

**Software**

The software developed to estimate monthly and annual values of rainfall erosivity in any location within the state of Rio Grande do Sul was named netErosividade RS, and it can be obtained free of charge at http://www.gprh.ufv.br. The main screen of netErosividade RS, in which the user identifies the target location is presented in figure 1. This procedure can be performed in three ways: the first is by clicking on the location on the map of the state of Rio Grande do Sul (field 1); the second is by choosing the name of a municipality in the state (field 2); and the third is by supplying values of latitude, longitude, and altitude of the location of interest (field 3).
Upon choosing a location, two lines (horizontal and vertical) are laid out on the map, indicating the selected location at their intersection. The estimated erosivity values are made available in field 4 of figure 1, where the tabs for the months of the year can be seen; the last one (annual) represents the erosivity of annual rainfall. It should be noted that ANNs related to erosivity of rainfall on an annual basis were not developed, so the annual values presented are the result of the sum of the monthly erosivities.

Table 3. Coefficient of determination ($R^2$), confidence index (c), and performance rating (D) of ANNs developed for the monthly estimate of rainfall erosivity according to the EI$_{30}$ and KE >25 erosivity index and the methods for obtaining kinetic energy of rainfall (F and WM)

<table>
<thead>
<tr>
<th>Month</th>
<th>EI$_{30}$</th>
<th>KE &gt;25</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>WM</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>c</td>
</tr>
<tr>
<td>Jan</td>
<td>0.78</td>
<td>0.80</td>
</tr>
<tr>
<td>Feb</td>
<td>0.79</td>
<td>0.80</td>
</tr>
<tr>
<td>Mar</td>
<td>0.93</td>
<td>0.83</td>
</tr>
<tr>
<td>Apr</td>
<td>0.90</td>
<td>0.89</td>
</tr>
<tr>
<td>May</td>
<td>0.90</td>
<td>0.86</td>
</tr>
<tr>
<td>Jun</td>
<td>0.77</td>
<td>0.82</td>
</tr>
<tr>
<td>Jul</td>
<td>0.80</td>
<td>0.67</td>
</tr>
<tr>
<td>Aug</td>
<td>0.86</td>
<td>0.72</td>
</tr>
<tr>
<td>Sept</td>
<td>0.78</td>
<td>0.76</td>
</tr>
<tr>
<td>Oct</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>Nov</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>Dec</td>
<td>0.77</td>
<td>0.81</td>
</tr>
</tbody>
</table>

F: Kinetic energy calculated by the method proposed by Foster et al. (1981). WM: Kinetic energy calculated by the method proposed by Wagner and Massambani (1988).

Figure 1. NetErosividade RS main screen.
The netErosividade RS allows one to print reports containing relevant information on the location (Name, Latitude, Longitude, and Altitude), the map of the state, illustrating the location of interest and the monthly and annual values of rainfall erosivity, considering the EI₃₀ and KE >25 erosivity indexes, and the two methods for obtaining the rainfall kinetic energy (Foster et al., 1981; Wagner and Massambani, 1988). To do so, the user only has to press the “Report” (“Relatório”) button (field 5). Figure 2 presents a report given by netErosividade RS for the location of Caxias do Sul, RS, Brazil.

The relative error ER% (Table 4) ranged from 3.3 to 38.7 %, respectively in the municipalities of São Borja and Encruzilhada do Sul. The municipalities of Rio Grande, Santa Maria, São Borja and Uruguaiana presented relative errors lower than 10 %, demonstrating the capacity of netErosividade RS to estimate the rainfall erosivity for these locations. The municipalities of Encruzilhada do Sul and Hulha Negra, which showed the highest relative errors, had overestimated erosivity values in relation to the study of Eltz et al. (2011) and Martins et al. (2009). Eltz et al. (2011) considered a series of pluviograms from 1958 to 1988, while Martins et al. (2009) used data recorded from 1956 to 1984. Due to climate change scenarios, the rainfall erosivity values calculated with data until the 1980s may not reflect the climatic behavior in the region in recent years. Marengo (2006) found the largest positive anomalies of rainfall in the south of Brazil in 1997, with rainfall up to 300 % above normal in northwestern Rio Grande do Sul. Pace (2011) states that an increase in rainfall was observed in southern Brazil in the last fifty years.

The increase in rainfall in the region indicates that the erosivity values calculated with more recent periods of rainfall data may be higher. This demonstrates the importance of using current values of erosivity, in order to represent the erosive capacity of rainfalls over the past few years.

Figure 2. Report generated by netErosividade RS for Caxias do Sul, RS, Brazil.
Overall, for the ten locations used in the comparative analysis, the mean relative error of the estimates carried out by netErosividade RS was 15.6 %. It should be noted that the periods of rainfall data used for the development of neural networks were different from data periods used by the works presented in table 1. The difference in netErosividade RS databases and in the erosivity databases from the literature can explain the variation of the values of erosivity by both methods. For instance, Eltz et al. (2013) pointed to the difference between the periods of databases to justify the difference between the erosivity value estimated by them and Cemetrs (2011), since they used the period from 1963 to 1993 and Cemetrs (2011) used the period from 1976 to 2005. Thus, the difference in the estimates performed by netErosividade RS (15.6 %) may be associated with the fact of databases to calculate the rainfall erosivity were not the same.

By considering that rainfall erosivity is one of the factors of the Universal Soil Loss Equation (USLE) to estimate soil loss, the overestimation or underestimation of the rainfall erosivity value may result in higher or lower estimated value of soil losses. Thus, the estimation of erosivity factor of rainfall may undermine the planning and design of soil and water conservation practices and water, plus the recommendation that the values estimated by netErosividade RS should be evaluated by a professional who has the knowledge of the characteristics of the region.

### CONCLUSIONS

Artificial neural networks can be an alternative for spatial interpolation of rainfall erosivity in the state of Rio Grande do Sul, Brazil.

The comparison between the values estimated by the software and those found in the literature demonstrated the capacity of netErosividade RS to estimate the rainfall erosivity for these locations.

The netErosividade RS allows the monthly and annual values of rainfall erosivity for any location within the State of Rio Grande do Sul to be obtained in a fast and easy manner.

### ACKNOWLEDGMENTS

The authors thank the Research Support Foundation of the State of Minas Gerais (Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG) and the Brazilian Council of Technological and Scientific Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq) for financial support.
REFERENCES


Marengo JA. Mudanças climáticas globais e seus efeitos sobre a biodiversidade: caracterização do clima atual e definição das alterações climáticas para o território brasileiro ao longo do século XXI. Brasília, DF: Ministério do Meio Ambiente; 2006.


