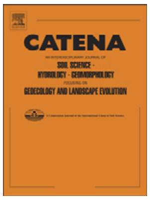
Catena 126 (2015) 53–58

Contents lists available at ScienceDirect

Catena

journal homepage: www.elsevier.com/locate/catena



A modified applicative criterion of the physical model concept for evaluating plot soil erosion predictions

V. Bagarello [a,*], V. Ferro [a], G. Giordano [a], F. Mannocchi [b], V. Pampalone [a], F. Todisco [b]

[a] *Dipartimento di Scienze Agrarie e Forestali, Università degli Studi di Palermo, Viale delle Scienze, 90128 Palermo, Italy*
[b] *Dipartimento di Scienze Agrarie, Alimentari ed Ambientali, Università degli Studi di Perugia, Borgo XX Giugno,74 06121 Perugia, Italy*

ARTICLE INFO

Article history:
Received 24 October 2013
Received in revised form 20 October 2014
Accepted 21 October 2014
Available online 20 November 2014

Keywords:
Soil erosion
Plot measurements
Soil loss data
Physical model

ABSTRACT

In this paper, the physical model concept by Nearing (1998. Catena 32: 15–22) was assessed. Soil loss data collected on plots of different widths (2–8 m), lengths (11–44 m) and steepnesses (14.9–26.0%), equipped in south and central Italy, were used. Differences in width between plots of given length and steepness determined a lower data correlation and more deviation of the fitted regression line from the identity one. A coefficient of determination between measured, M, and predicted, P, soil losses of 0.77 was representative of the best-case prediction scenario, according to Nearing (1998). The relative differences, $Rdiff = (P - M) / (P + M)$, decreased in absolute value as M increased only for erosion rates approximately >1 kg m^{-2}. An alternative applicative criterion of the physical model concept, based on the $|P - M|$ difference, was valid for the entire range of measured soil losses. In conclusion, the physical model should be defined in terms of perfect planimetrical equivalence. The best applicative criterion of the physical model concept may vary with the considered dataset, which practically implies the need to further test this concept with other datasets.

1. Introduction

Using a model for predicting soil loss due to water erosion is useful to predict both the aggressiveness of the phenomenon in an area of interest and the effects of different soil erosion control practices. These predictions have interest for many reasons, including safeguard of the people. For example, many valley towns in Italy are crossed by streams that are frequently covered by roads. In these cases, there is the need to reduce ordinary sediment yield to tolerable levels, so to minimize the risk of obstruction of the stream at its outlet. Obstruction phenomena can favor disastrous flooding during severe rainfall events (e.g., Bagarello et al., 2010b).

The performances of a soil erosion model have to be tested for establishing the expected reliability of the soil loss estimates (Foster and Lane, 1987; Quinton, 1994). The quality of the predictions can be established only if a criterion to discriminate between "acceptable" and "unacceptable" soil loss estimates is available.

Although an erosive event occurring on plots having identical characteristics in terms of soil, morphology, land use and crop management practices yields runoff and soil loss data varying from plot to plot (Bagarello and Ferro, 2004, 2006; Ruttimann et al., 1995; Wendt et al., 1986), a single or a few replicated plot soil loss data are generally collected for a given treatment. The circumstance that similar plots give different soil loss outputs affects the performance evaluation of a soil erosion model. In fact, for a particular condition, the departure between the measurement and the corresponding prediction has to take into account the prediction error, due to the model structure and the input data, and the deviation of the measured sample value from the representative mean value (Nearing, 2000).

According to Nearing (1998), the best possible model to predict soil loss from an area is a physical model of the area which is characterized by a similar soil type, land use, size, shape, slope and climatic inputs. In other words, the physical model obtained by a replicated plot is the best possible, unbiased, real world model. Using event soil loss data for approximately 3000 pairs of replicated plots, Nearing (1998) compared the measured, M, soil losses and the predicted, P, ones obtained by the physical model represented by the replicated plot. Nearing (1998) obtained a coefficient of determination, R^2, of the linear relationship between P and M equal to 0.77 and he concluded that an uncalibrated erosion model would not give a better overall result. Nearing (2000) also proposed that a soil erosion prediction has to be considered acceptable if the difference between the prediction and the measurement lies within the population of differences between pairs of measured values. Using a large set of plot soil loss data collected in different U.S. locations, Nearing (2000) also developed empirical relationships to predict the 90% and 95% occurrence intervals of the relative differences, $Rdiff$, between replicated plots as a function of the measured soil loss.

In a previous paper, the soil loss measurements carried out in Sicily, at the Sparacia station, supported the conclusion that a coefficient of determination between measured and predicted soil losses of 0.77 has to be considered as a benchmark or best-case prediction scenario

---

* Corresponding author. Tel.: +39 09123897053; fax: +39 091484035.
E-mail address: vincenzo.bagarello@unipa.it (V. Bagarello).

http://dx.doi.org/10.1016/j.catena.2014.10.021
0341-8162/© 2014 Elsevier B.V. All rights reserved.

(Bagarello and Ferro, 2012). The 95% occurrence interval for the data developed by Nearing (2000) included approximately 88–89% of the data collected at Sparacia. Taking into account that this discrepancy was moderate, i.e. a few percentage units, and considering that a large sample size and a wide variety of conditions were considered in the U.S. study, the conclusion by Nearing (2000) that the developed analysis should be usable for model validation studies in general was considered to be reasonable.

The influence of plot width, w, on the definition of the physical model should be considered. Bagarello and Ferro (2012) assumed that w did not affect the analysis and a single data set was considered for a given event for the 22 m long plots independently of w (2 or 8 m). The reasons of this choice were: i) the plots included in the investigation by Nearing (2000) ranged from 2 to 8 m in width, ii) a recent investigation (Bagarello et al., 2011) showed that soil loss differences between two plot widths were not statistically significant at Sparacia and plot width effects were negligible for the most erosive events, and iii) in plot soil loss models such as the Universal Soil Loss Equation (USLE) and its revised version (Renard et al., 1997; Wischmeier and Smith, 1978), soil loss per unit area is considered to depend on plot length but not on plot width. However, plot width affected measurement of soil loss for the less erosive events, suggesting a more appreciable dependence of the plot response on the local conditions in this last case. We did not find other investigations of the plot width effects on the measured soil loss in the literature. Therefore, establishing these effects with reference to the physical model concept is necessary to include data of appropriate quality in the (P, M) dataset.

Another point to be developed is the possibility to generalize the results by Nearing (2000), which was partially supported by Bagarello and Ferro (2012). Nearing (2000) used a huge dataset but his approach has a strong empirical connotation. For example, Nearing (2000) considered a minimum soil loss of 0.01 kg m^{-2} whereas smaller values were included in the investigation by Bagarello and Ferro (2012). Therefore, extending the investigation to other data and environments is desirable to be sure that the developed analysis is usable for model validation studies in general or to recognize the need or the opportunity to modify the procedure. In other terms, the methodology developed by Nearing (2000) to establish the effect of the severity of the erosive event on the expected differences between predicted and measured soil loss needs testing with data not included in the U.S. database. This test might suggest the opportunity to improve the methodology but such an improvement should be carried out by maintaining the centrality of the physical model concept. To our knowledge, however, no other studies are available which tested the physical model concept.

The general aim of this paper is to test the physical model concept by using soil loss data collected on plots of different lengths, widths and slopes at two experimental stations located in southern and central Italy. The three specific objectives are to: i) establish the plot width effects with reference to the physical model concept; ii) assess the applicability of the existing procedure to test plot scale soil erosion models; and iii) develop an alternative procedure to assess the suitability of an erosion model for soil loss prediction.

2. Materials and methods

Data for this investigation were collected at the "Sparacia" (south Italy) and "Masse" (central Italy) experimental stations for soil loss measurement (Table 1). The characteristics of the two stations were described in detail in other papers (e.g., Bagarello and Ferro, 2004; Bagarello et al., 2011; Todisco et al., 2012) and they were only summarized here for brevity reasons. In particular, the experimental station for soil erosion measurement "Sparacia" of the Department of Agricultural and Forestry Sciences of the Palermo University is located in western Sicily, southern Italy, approximately 100 km south of Palermo. It includes two plots of 8×44 m^2, two plots of 8×33 m^2, six plots of 8×22 m^2, two plots of 2×22 m^2, two plots of 4×11 m^2, and two plots of 2×11 m^2. The oldest plots (four plots of 8×22 m^2) were constructed in 1999, whereas the most recent plots (two plots of 2×22 m^2) were constructed in 2007. All these plots were installed on a 14.9% slope. Two plots of 6×22 m^2 were also realized on a 22.0% slope and other two plots (6×22 m^2) were constructed on a 26.0% slope. The area has a typical Mediterranean semi-arid climate with an average annual rainfall of approximately 700 mm. The soil has a clay texture (clay = 62%, silt = 33% and sand = 5%) and it shows a massive consistency in winter, when it is wet and fully swelled, but it develops a polygonal pattern of surface shrinkage cracks in late spring or early summer as the soil dries. The experimental station for soil erosion measurements "Masse" of the Department of Civil and Environmental Engineering of the Perugia University was established in 2007. It is located 20 km south of Perugia in the Umbria region (central Italy). The station includes ten plots: four plots of 8×22 m^2, two plots of 4×22 m^2, two plots of 4×11 m^2, and two plots of 2×11 m^2. All plots are oriented parallel to a 16% slope. The area has a characteristic Mediterranean climate with an average annual rainfall of 900 mm. The soil has a silty–clay–loam texture (clay = 34%, silt = 59% and sand = 7%). The structure is polyhedral angle and the gravel content is negligible. All considered plots were maintained in a cultivated fallow and rills were obliterated by hand implements at the end of each erosive event.

Events with two or more replicated measurements for a given plot type (length, width and slope steepness) were included in this database and the physical model concept was tested according to Nearing (1998). Two data points were obtained from the soil loss data collected, for a given event, at the two available plots of given geometric characteristics. For the first data point, one value (A) of the pair was chosen to serve as the measured, M, value of erosion and the other (B) was considered to

Table 1

General characteristics of the sampled plots and erosive events.

Station	Plot width and length (m)	Number of plots	Slope steepness (%)	Sampling period	Erosive events	N	A_e (kg m^{-2})			$N_{<0.01}$ (%)	$N_{>1}$ (%)
							Min	Max	Mean		
Sparacia	2 × 11	2	14.9	09/2004–10/2011	21	42	0.0048	11.31	0.80	11.9	21.4
	4 × 11	2	14.9	09/2004–10/2011	22	44	0.0027	7.28	0.92	9.1	20.5
	2 × 22	2	14.9	09/2007–10/2011	11	22	0.0099	3.42	0.66	4.5	22.7
	8 × 22	6	14.9	11/1999–01/2012	52	235	0.00029	21.70	1.05	19.6	22.6
	6 × 22	2	22.0	09/2007–03/2012	19	38	0.011	8.35	1.48	0	44.7
	6 × 22	2	26.0	09/2007–03/2012	19	38	0.014	7.84	2.07	0	65.8
	8 × 33	2	14.9	01/2002–01/2012	39	78	0.00024	6.68	0.86	15.4	25.6
	8 × 44	2	14.9	09/2004–01/2012	23	46	0.00012	5.62	0.80	28.3	15.2
Masse	2 × 11	2	16.0	11/2008–12/2011	23	46	0.0065	3.48	0.59	2.2	19.6
	4 × 11	2	16.0	03/2008–05/2012	37	74	0.0024	2.33	0.32	12.2	8.1
	4 × 22	2	16.0	03/2008–05/2012	35	70	0.00075	1.17	0.13	35.7	1.4
	8 × 22	4	16.0	02/2008–05/2012	43	86	0.00040	0.96	0.06	58.1	0

A_e = event plot soil loss per unit area; N = sample size, i.e. number of individual plot soil loss data; $N_{<0.01}$ = percentage of A_e values smaller than 0.01 kg m^{-2}; $N_{>1}$ = percentage of A_e values greater than 1 kg m^{-2}.

be the predicted, P, value from the physical model. For the second data point, value (B) was used as the measured value and value (A) as the predicted one. The availability of a larger number of replicated soil loss measurements for a given event allowed calculation of more data points. For example, six (P, M) data pairs were obtained when data were collected from three simultaneously operating plots (Table 1).

At first, for a given slope steepness at a particular experimental area (Sparacia, $s = 14.9\%$; Masse, $s = 16.0\%$), a plot length was selected (11, 22 m) and a comparison between the P vs. M relationships determined by considering two alternative scenarios was carried out. In scenario 1, the physical model was defined in terms of plot length alone; in other words, differences in plot width were neglected and, for a given length, a plot having a width, for example, of 4 m was assumed to be the physical model of a plot of 8 m. Scenario 2 implied a physical model identical to the sampled plot also in terms of plot width. This preliminary analysis allowed to establish how the physical model should be defined.

Then, a single set of (P, M) data was developed by pooling together all available data from the two stations. The P vs. M relationship (Nearing, 1998) was determined, and a relative difference, *Rdiff*, was calculated for each (P, M) data pair according to the following relationship (Nearing, 2000):

$$Rdiff = \frac{P-M}{P+M}. \tag{1}$$

The (*Rdiff*, M) data pairs were used to verify if *Rdiff* decreases as M increases (Nearing, 2000), and also to compute the number of predictions for which *Rdiff* fell within the 95% occurrence interval that was calculated by the following relationships, derived by Nearing (2000):

$$R_{diff,INF} = 0.236 \ \log(M) - 0.641 \tag{2a}$$

$$R_{diff,SUP} = -0.179 \ \log(M) + 0.416 \tag{2b}$$

where $R_{diff,INF}$ and $R_{diff,SUP}$ are the lower and the upper limit of the interval and M is expressed in kg m^{-2}.

Finally, an alternative criterion for evaluating the reliability of a soil erosion predictive model was developed on the basis of the established relationship between $|P - M|$ and M.

3. Results and discussion

Measurements collected from November 1999 to May 2012 allowed to sample 11 to 52 erosive events, depending on the plot type established at a given experimental station (i.e., length, width and slope steepness, Table 1). A total of 819 individual soil loss data, A_e, varying from 0.00012 to 21.7 kg m^{-2} were considered. The wide ranges of A_e values suggested a good representativeness of the available dataset for the purposes of this investigation. The individual A_e values greater than 1 kg m^{-2} (=10 t ha^{-1}), which is an approximation of the tolerable soil loss in a year (Bagarello and Ferro, 2006), varied from 0 to 66% of the collected data, depending on the plot type (Table 1). The highest percentages of $A_e > 1$ kg m^{-2} values (45–66%) were detected on the most sloped plots of Sparacia ($s = 22$–26%), which was not surprising, being well known that slope steepness has a noticeable impact on soil erosion phenomena (Nearing, 1997; Renard et al., 1997; Wischmeier and Smith, 1978). The individual A_e values lower than 0.01 kg m^{-2}, which was the minimum value considered by Nearing (2000), varied with the plot type from 0 to 58% of the collected data (Table 1).

For a given experimental scheme, i.e., a given combination of experimental station, plot length and slope steepness, the linear regression analysis between the measured soil losses, M, and the predicted ones by the physical model, P, differed with the considered scenario in terms of physical model definition (Table 2). The first and most obvious difference was the sample size. Less data points were available for scenario 2 because a (P, M) data pair corresponding to two plots differing in width was included in scenario 1 but not in scenario 2. Sample sizes for scenario 2 were approximately 33 to 38% of those for the corresponding scenario 1. An exception was detected for the 22 m long plots of Sparacia (80%) because in this case soil loss data collected on the larger plots prevailed in comparison with the ones obtained on the narrower plots (Table 1). The four experimental schemes considered in this investigation yielded qualitatively similar results. In particular, an intercept, b_0, closer to 0, a slope, b_1, closer to 1 and a higher coefficient of determination, R^2, of the P vs. M linear relationship were detected when the physical model was defined in terms of planimetrical equivalence to the sampled plot (scenario 2, Table 2). Taking into account that a perfect model yields $b_0 = 0$, $b_1 = 1$, and $R^2 = 1$, this analysis induced to conclude that the physical model has to be defined in terms of perfect planimetrical equivalence. In other terms, the physical model of a plot is another plot having the same length and the same width of that plot.

A possible explanation of this result is that plot hydrological processes may vary with plot width. According to Bagarello et al. (2011), for example, the probability for runoff to become concentrated (higher runoff and soil loss) may be expected to increase in a narrow plot than a wide one since runoff has less possibilities to deviate from the flow direction of maximum slope in the former case. On the other hand, localized areas with a relatively low soil erodibility, high roughness or high infiltration rates, which reduce runoff and soil loss, can have a more appreciable effect in the narrow plots than the wide ones. To schematically illustrate the link between plot width and the spatial variability of the soil properties controlling the hydrologic behavior of a plot, Fig. 1 shows, for a 4 × 11 m^2 plot established at Sparacia, a map of the spatial distribution of the soil bulk density, ρ_b (Bagarello et al., 2010a), that influences both the soil water retention curve and the soil hydraulic conductivity function (Assouline, 2006a,b). Relatively high ρ_b values (i.e., >1.15 Mg m^{-3}) occupy 5.1% of the plot area whereas low ρ_b values (<0.9 Mg m^{-3}) occupy 5.5% of the plot. If a plot width of 2 m is considered, high and low ρ_b values occupy 8.3% and 11.0%, respectively, of the half plot on the left of the figure and 1.9% and 0%, respectively, of the half plot on the right. None of the two smaller plots reproduces the spatial

Table 2
Intercept, b_0, slope, b_1, and coefficient of determination, R^2, of the linear regression line between the measured soil loss, M, and the predicted one by the physical model, P, for two scenarios differing in terms of physical model definition.

Station	Plot length (m)	Slope steepness (%)	Scenario	Sample size	Intercept, b_0	Slope, b_1	Coefficient of determination, R^2
Sparacia	11	14.9	1 (λ)	262	0.6123	0.3793	0.144
			2 (λ, w)	86	0.4598	0.4651	0.216
	22	14.9	1 (λ)	1126	0.1658	0.8359	0.699
			2 (λ, w)	906	0.0961	0.905	0.819
Masse	11	16.0	1 (λ)	316	0.0978	0.7864	0.618
			2 (λ, w)	120	0.0522	0.8747	0.766
	22	16.0	1 (λ)	440	0.0338	0.6605	0.436
			2 (λ, w)	156	0.0129	0.8578	0.736

Scenario 1: the physical model was defined in terms of plot length, λ, alone. Scenario 2: the physical model was identical to the sampled plot in terms of both λ and width, *w*.

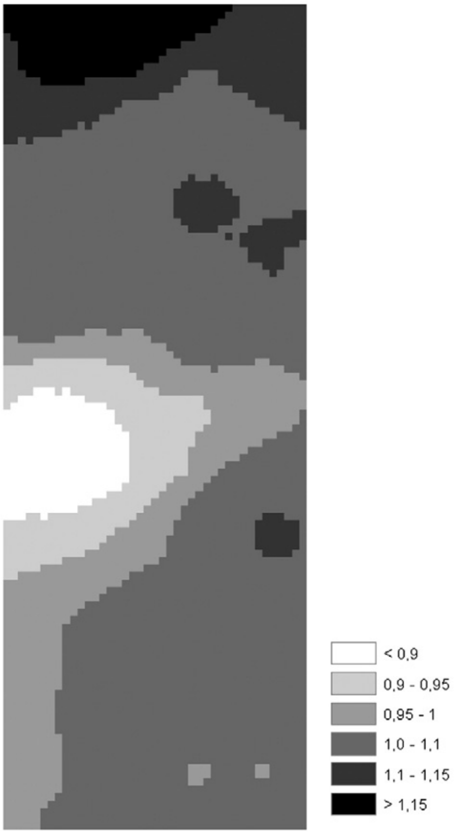


Fig. 1. Map of the spatial distribution of soil bulk density (Mg m^{-3}) for the N plot of 4 × 11 m^2 (from Bagarello et al., 2010a).

distribution of ρ_b detected in the wider plot. Therefore, differences in the hydrological response of narrow and wide plots can occur. Reasoning in terms of planimetrical equivalence, the longer plots yielded higher (Sparacia) or similar (Masse) coefficients of determinations as compared with the shorter plots (Table 2). A possible reason of this result is that local heterogeneities have a similar impact on the plot response for $\lambda = 11-22$ m or they are better averaged on the longer plots.

Therefore, an Italian (P, M) database was developed by considering the two stations (Sparacia and Masse), four plot lengths ($\lambda = 11$, 22, 33 and 44 m), four slope steepness ($s = 14.9$, 16.0, 22.0 and 26.0%), and defining a physical model as a plot planimetrically identical to the sampled one. The database included 1468 data pairs. This large sample size was possible because more than two plots of 8 × 22 m^2 were simultaneously operating at Sparacia. To be clearer, the number of (P, M) data pairs coincided with the number of individual plot soil loss data (Table 1) when two plots were simultaneously operating because two data pairs were obtained for each erosive event. The number of simultaneously operating 8 × 22 m^2 plots at Sparacia varied with the event, with a maximum of six. Therefore, 235 individual plot soil loss data yielded 884 (P, M) data pairs.

The mean of the 1468 values of M was 0.88 kg m^{-2} and the regression line between predicted, P, and measured, M, soil loss per unit area was (Fig. 2):

$$P = 0.115 + 0.870 \times M \tag{3}$$

having an R^2 value of 0.76 and a 95% confidence interval for the intercept and the slope equal to 0.061–0.169 and 0.844–0.895, respectively. Eq. (3) was very close to Eq. (2) by Bagarello and Ferro (2012), obtained by considering the Sparacia data collected on plots of $\lambda \geq 22$ m and defining the physical model in terms of plot length alone (intercept = 0.105, slope = 0.875, $R^2 = 0.77$), and it was also practically coincident with the P vs. M linear relationship obtained by Nearing (1998) (Fig. 2). Therefore, including additional data in the considered dataset and using a more restrictive criterion to define the physical model did not have an appreciable effect on the relationship between predicted and measured soil loss.

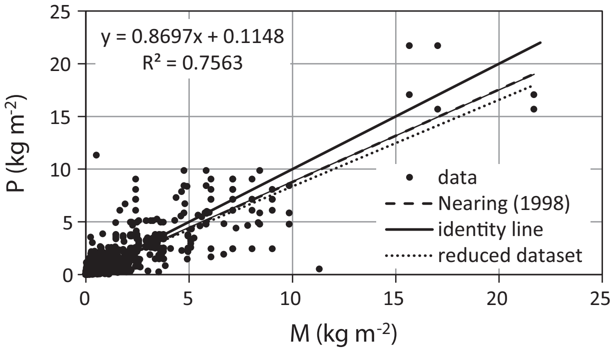


Fig. 2. Predicted, P, vs. measured, M, soil loss for the physical model of soil erosion as represented by pairs of replicated plots (sample size, $N = 1468$).

The dataset considered in this investigation differed appreciably from the one used by Nearing (1998) for several factors, including sample size, maximum and mean measured soil loss values, and probably applied experimental methodologies (Bagarello and Ferro, 2012). The Italian and the American regression lines, and also the two determination coefficients, were practically coincident notwithstanding the listed differences between the two datasets. Therefore, this investigation gave additional support to the conclusion by Nearing (1998, 2004) that an R^2 of 0.77 has to be considered as a benchmark or best-case prediction scenario. In other words, our results confirmed that it should not be expected that an erosion model would give better overall results than those obtained by Nearing (1998).

Fig. 2 also shows the P vs. M relationship obtained by neglecting the soil loss values >15 kg m^{-2}. There was a small deviation of this line (intercept = 0.145, slope = 0.821) from the one obtained with the complete dataset, and also a small decrease of R^2, equal to 0.67 with the reduced dataset. Although differences between the complete and the reduced datasets were not substantial, this comparison supported the importance to use the most representative possible datasets for determining the P vs. M relationship.

For the Italian database, the percentage of *Rdiff* values falling within the 95% occurrence interval calculated by Eqs. (2a) and (2b) was equal to 88.9%. A similar result (i.e. a percentage of 87.8%) was obtained by only considering the M values greater than 0.01 kg m^{-2} ($N = 1194$ data pairs) that was the minimum soil loss in the U.S. investigation (Nearing, 2000). Bagarello and Ferro (2012) suggested that a 95% occurrence interval for the data including 88–89% of data, denoting moderate differences between the expected and the measured data (i.e. a few percentage units), was reasonably indicative of the usability of the analysis developed by Nearing (2000) for model validation studies in general. Therefore, this investigation was in line with the suggestion by Bagarello and Ferro (2012).

The premise of the analysis by Nearing (2000) was that the measured data with greater erosion rates showed, on average, less relative differences between replicates. This tendency was experimentally supported by Bagarello and Ferro (2012) (see their Figs. 4 and 5) but it was not so clear with the new Italian database (Fig. 3), which was more representative than the previously developed one since it included data from a larger variety of conditions (experimental stations, plot lengths and steepness) and it was based on a more restrictive definition of the physical model (planimetrical equivalence). The first point to be noted is that *Rdiff* showed a moderate but detectable tendency to decrease (i.e., higher percentage of negative values) as M increased. In

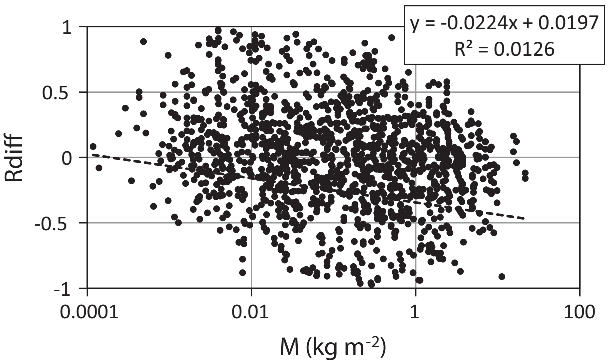


Fig. 3. Relative differences in measurement of soil loss between replicated plots, *Rdiff*, vs. the measured soil loss value, *M*, for the Italian database (sample size, $N = 1468$).

other terms, *Rdiff* was not perfectly symmetrical around the $Rdiff = 0$ line. This tendency was conceptually consistent with a regression line between P and M having a slope lower than 1 (Fig. 2), since both analyses suggested that the predicted value tends to be lower than the measured one for high erosion rates. Moreover, the regression line between *Rdiff* and *M*, characterized by a low but statistically significant ($p = 0.05$) R^2 value, suggested that the relative differences between P and M decreased with an increase of M. Three regions were distinguishable on the *Rdiff* vs. *M* plot, with approximate boundaries between adjacent regions at 0.01 and 1 kg m^{-2} (Fig. 3). In particular, relative differences decreasing with an increase in M were only detected for $M > 1$ kg m^{-2}. With reference to these data ($N = 316$), the percentage of *Rdiff* values falling within the 95% occurrence interval calculated by Eqs. (2a) and (2b) was equal to 85.1%. The relative differences also decreased with a decrease in M for $M < 0.01$ kg m^{-2}. This result was considered to be reasonable, since the differences between two replicated plots are expected to decrease when the event has a low erosive power. Two plots yield similar results because the rainfall–runoff event is able to detach and transport only a small amount of soil particles. For $0.01 < M < 1$ kg m^{-2}, a relationship between *Rdiff* and *M* was not detectable and the data points practically occupied all the space of the graph, suggesting that plot heterogeneities have a more noticeable impact on soil loss for intermediate levels of the erosion phenomenon. Therefore, the premise by Nearing (2000) was only confirmed with reference to the highest erosion rates (i.e., $M > 1$ kg m^{-2}) but also in this particular case the correspondence of the predicted occurrence interval with the data was not fully satisfactory. Developing specific confidence intervals with reference to the Italian database was not possible because the available sample size for highly erosive events is too small for carrying out an analysis of the data similar to the one carried out by Nearing (2000).

The availability of data from other locations could probably allow a recalibration of Eqs. (2a) and (2b) for predictive purposes. In any case, a point to be considered in the development of new confidence intervals for highly erosive events is the choice of the data to be included in the calculation of *Rdiff*. The *Rdiff* dataset used by Nearing (2000) included several values of $Rdiff = -1$, indicating a value of $P = 0$, but no points were plotted for $Rdiff = 1$. The reason was that, in this latter case, M is equal to zero and since the graph is logarithmic on the x-axis, no values of $M = 0$ can be plotted (Nearing et al., 1999). A value of *Rdiff* equal to -1 or 1 was not calculated in this investigation because soil loss values greater than zero for both replicated plots were considered (Bagarello and Ferro, 2012). This choice was made because the physical implication of the assumption by Nearing (2000), i.e. that only predictions, and not also measurements, can be equal to 0, was not considered to be fully realistic. Another reason was that $Rdff = -1$ or 1 can be obtained independently of the real difference between the measured and the predicted variable. In other terms, *Rdiff* ignores the fact that a prediction of 0 is more credible if the measured value is very close to zero (e.g., 0.0001 kg m^{-2}) and much less credible if the soil erosion rate is very high (e.g. 10 kg m^{-2}).

From a scientific point of view, the sign of the difference between P and M has to be determined to establish how to improve a soil erosion predictive tool. From a practical point of view, however, the absolute $P - M$ difference is enough to establish the accuracy level of the predictions. The $|P - M|$ was found to increase with M according to the following relationship (Fig. 4):

$$|P{-}M| = 0.356 \times M^{0.91} \tag{4}$$

having a coefficient of determination $R^2 = 0.72$ and a 95% confidence interval of the exponent of 0.88–0.94, denoting a nonlinearity of the relationship. In relative terms, Eq. (4) predicts a departure of the prediction from the measured value that decreases from 80% to 27% as M increases from the minimum (0.00012 kg m^{-2}) to the maximum (21.7 kg m^{-2}) measured value.

Eq. (4) can be viewed as an alternative approach for applying the physical model concept by Nearing (2000) since it predicts, for a given soil loss value (M), what is the mean absolute difference associated with the sampling of another, identical plot. A soil loss prediction by a model is accurate enough if the absolute difference with the measured value does not exceed the $|P - M|$ value calculated by Eq. (4). The least restrictive criterion, using a relationship enveloping all data points, could alternatively be proposed. An intermediate criterion between the suggested regression line and a data enveloping line could also be developed by carrying out a frequency analysis of the data divided into half log-cycle intervals, similar to the one carried out by Nearing (2000) to estimate 95% occurrence intervals for the data.

The approach presented in this investigation is promising, because it was found to be valid for the entire range of the measured soil loss values, but it cannot still be suggested for a general use, because the analysis had an empirical character and data were only collected in two experimental stations. Developing a larger database including other experimental areas is therefore advisable to confirm the general validity of the fitted relationship. For example, it could be interesting from both a scientific and a practical point of view to use the dataset by Nearing (2000) to compare his approach with the one developed here. The availability of more data might also allow considering alternative criteria to establish the accuracy of a predicted soil loss value. In this investigation, we considered that, for a given measured value, the $|P - M|$ difference discriminating between acceptable and unacceptable soil loss predictions was definable by an average of the experimental differences between replicated plots.

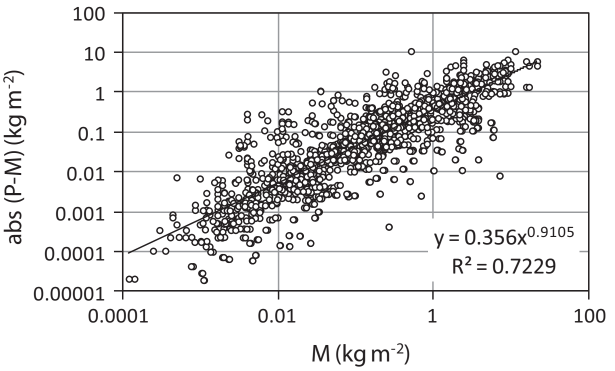


Fig. 4. Absolute differences in measurement of soil loss between replicated plots, abs ($P - M$), plotted against the measured value, M, for the Italian database (sample size, $N = 1468$).

4. Conclusions

The performances of a soil erosion model have to be tested for determining the expected reliability of the soil loss predictions. An acceptable evaluation criterion of the performances of a deterministic model has to take into account that, for a particular treatment, a portion of the difference between the measured and the predicted erosion rate is due to unexplained variance of the measured sample value from the representative mean. The concept of a physical model represented by a replicated soil loss measurement is theoretically based and also attractive to evaluate plot soil loss predictive models.

At first, this investigation, using the plot soil loss data collected at the Sparacia and Masse stations, in Sicily (south Italy) and Umbria (central Italy), respectively, showed that the physical model has to be defined in terms of perfect planimetrical equivalence. In other terms, the physical model of a plot is another plot that has the same length and the same width of that plot. Differences in width between the plot and its physical model determine a lower correlation of the data and a more appreciable departure of the regression line between the predicted and the measured soil loss from the identity line.

Then, the Italian database developed in this investigation supported the conclusion that a coefficient of determination between measured and predicted soil loss values of 0.77 has to be considered as a benchmark or best-case prediction scenario. Therefore, it should not be expected that an uncalibrated, deterministic erosion model would give more accurate results than those obtained by a replicated plot measurement.

The available data also showed that the relative differences between the predicted and the measured soil loss, *Rdiff*, decreased with an increase in the measured value, *M*, only for high *M* values, i.e. greater than approximately 1 kg m^{-2}. With reference to these high values, the 95% occurrence interval for the data, calculated by the single available procedure developed in the U.S., included approximately 85% of the experimentally determined relative differences. Therefore, the procedure developed for the U.S. was not usable for the entire range of measured soil losses in Italy. When the *Rdiff* vs. *M* trend was similar for the U.S. and the Italian data (i.e., high measured values), a need to develop larger confidence intervals for Italy was detected.

Finally, an alternative criterion was developed, since the absolute difference between the predicted and the measured soil loss was found to increase monotonically with the soil erosion rate. This criterion seems promising since it was usable for the entire range of sampled soil erosion values.

Developing a single data set by a contribution of several authors working in different parts of the world could be a desirable step to develop the most possible robust criterion for plot soil erosion model validation studies.

Acknowledgments

All authors set up the research and contributed to both analyze the data and write the manuscript. V. Bagarello, V. Ferro, G. Giordano and V. Pampalone provided data for the Sparacia station. F. Mannocchi and F. Todisco provided data for the Masse station. This research was financially supported by the PRIN 2010–2011 Project (B71J12000610001).

References

Assouline, S., 2006a. Modeling the relationship between soil bulk density and the water retention curve. Vadose Zone J. 5, 554–563.

Assouline, S., 2006b. Modeling the relationship between soil bulk density and the hydraulic conductivity function. Vadose Zone J. 5, 697–705.

Bagarello, V., Ferro, V., 2004. Plot-scale measurement of soil erosion at the experimental area of Sparacia (southern Italy). Hydrol. Process. 18, 141–157.

Bagarello, V., Ferro, V., 2006. Erosione e Conservazione del Suolo. McGraw-Hill, Milano 88-386-6311-4 (539 pp. (in Italian)).

Bagarello, V., Ferro, V., 2012. Testing the “physical model concept” by soil loss data measured in Sicily. Catena 95, 1–5. http://dx.doi.org/10.1016/j.catena.2012.02.017.

Bagarello, V., Di Stefano, C., Ferro, V., Iovino, M., Sgroi, A., 2010a. Physical and hydraulic characterization of a clay soil at the plot scale. J. Hydrol. 387, 54–64. http://dx.doi.org/10.1016/j.hydrol.1010.03.029.

Bagarello, V., Di Stefano, C., Ferro, V., 2010b. L'evento alluvionale dell'Ottobre 2009 nel Messinese. Quad. Idronom. Mont. 29 (2), 25–45 (in Italian).

Bagarello, V., Ferro, V., Giordano, G., Mannocchi, F., Pampalone, V., Todisco, F., Vergni, L., 2011. Effect of plot size on measured soil loss for two Italian experimental sites. Biosyst. Eng. 108, 18–27. http://dx.doi.org/10.1016/j.biosystemseng.2010.10.006.

Foster, G.R., Lane, L.J. (compilers), 1987. User requirements: USDA-Water Erosion Prediction Project (WEPP). NSERL Report no.1, USDA-ARS National Soil Erosion Research Laboratory, West Lafayette, IN.

Nearing, M.A., 1997. A single continuous function for slope steepness influence on soil loss. Soil Sci. Soc. Am. J. 61, 917–919.

Nearing, M.A., 1998. Why soil erosion models over-predict small soil losses and under-predict large soil losses. Catena 32, 15–22.

Nearing, M.A., 2000. Evaluating soil erosion models using measured plot data: accounting for variability in the data. Earth Surf. Process. Landf. 25, 1035–1043.

Nearing, M.A., 2004. Capabilities and limitations of erosion models and data. ISCO 2004 — 13th International Soil Organization Conference “Conserving soil and water for society — sharing solutions”, Brisbane, July.

Nearing, M.A., Govers, G., Norton, L.D., 1999. Variability in soil erosion data from replicated plots. Soil Sci. Soc. Am. J. 63, 1829–1835.

Quinton, J.N., 1994. The Validation of Physically Based Erosion Models — With Particular Reference to EUROSEM(PhD thesis) Cranfield University.

Renard, K.G., Foster, G.R., Weesies, G.A., McCool, D.K., Yoder, D.C., 1997. Predicting Soil Erosion by Water — A Guide to Conservation Planning With the Revised Universal Soil Loss Equation (RUSLE). Agr. Handbook no.703. U.S. Dept. of Agric.

Ruttimann, M., Schaub, D., Prasuhn, V., Ruegg, W., 1995. Measurement of runoff and soil erosion on regularly cultivated fields in Switzerland — some critical considerations. Catena 25, 127–139.

Todisco, F., Vergni, L., Mannocchi, F., Bomba, C., 2012. Calibration of the soil loss measurement method at the Masse experimental station. Catena 91, 4–9.

Wendt, R.C., Alberts, E.E., Hjelmfelt Jr., A.T., 1986. Variability of runoff and soil loss from fallow experimental plots. Soil Sci. Soc. Am. J. 50, 730–736.

Wischmeier, W.H., Smith, D.D., 1978. Predicting Rainfall Erosion Losses. A Guide to Conservation Planning. USDA Agriculture Handbook no. 537. U.S. Gov. Print. Office, Washington DC.