Earth Surface Processes and Landforms
Earth Surf. Process. Landforms **25**, 1035–1043 (2000)

EVALUATING SOIL EROSION MODELS USING MEASURED PLOT DATA: ACCOUNTING FOR VARIABILITY IN THE DATA

MARK A. NEARING*
National Soil Erosion Laboratory, USDA-ARS, West Lafayette, IN 47907-1196, USA

Received 10 September 1999; Revised 11 February 2000; Accepted 1 March 2000

ABSTRACT

One of the important methods used to evaluate the effectiveness of soil erosion models is to compare the predictions given by the model to measured data from soil loss collected on plots taken under natural rainfall conditions. While it is recognized that plot data contain natural variability, this factor is not quantitatively considered during such evaluations because our knowledge of natural variability between plots which have the same treatments is very limited. The goal of this study was to analyse sufficient replicated plot data and present methodology to allow the model evaluator to take natural, within-treatment variability of erosion plots into account when models are tested. A large amount of data from pairs of replicated erosion plots was evaluated and quantified. The basis for the evaluation method presented is that if the difference between the model prediction and a measured plot data value lies within the population of differences between pairs of measured values, then the prediction is considered 'acceptable'. A model 'effectiveness' coefficient was defined for studies undertaken on large numbers of prediction versus measured data comparisons. This method provides a quantitative criterion for taking into account natural variability and uncertainty in measured erosion plot data when those data are used to evaluate erosion models. Published in 2000 by John Wiley & Sons, Ltd.

KEY WORDS: soil erosion; soil conservation; soil erosion models; hydrology; surface water; sediment; spatial variability

INTRODUCTION

Many criteria may be and are used to evaluate soil erosion models. For the Water Erosion Prediction Project (WEPP) model, for example, a series of seven individual points were defined relative to 'Validation Criteria' (Foster and Lane, 1987). These included several subjective statements such as: 'The model is valid if it serves its intended purpose as defined by these specific User Requirements'; 'The model is based on scientific principles...'; 'The model gives results that are more useful for agency program objectives than those given by the USLE and applies to situations not appropriate for the USLE.' These types of criteria are important and necessary. Another type of evaluation criteria for models, and perhaps the most commonly considered one, is the comparison of model predictions to measured erosion data. The WEPP criteria (Foster and Lane, 1987) also addressed this type of evaluation: 'Judgements on the "goodness of fit" of the estimates from the procedure to observed data are to be based on the data sets as a whole and not on a few specific isolated data sets. Quantitative measures of the "goodness of fits" will be calculated and presented, but a specific quantitative level of accuracy figure is not being required because of the great variation in the experimental data that will be used in the validation.'

Quinton (1994) suggested a methodology for erosion model validation which uses a two-part process: corroboration and evaluation. Corroboration involves one-way and two-way sensitivity analysis, and the comparison of the model response to critical experimental data in order to examine the fundamental hypotheses imbedded in the model structure. Evaluation, in Quinton's scheme, involves the definition of the application and the selection of a success/failure criteria such as a certain percentage of error allowed, based on the intended application of the model. Then he proposes that the model be applied to the type of data to be

* Correspondence to: M. A. Nearing, National Soil Erosion Laboratory, USDA-ARS, West Lafayette, Indiana 47907-1196, USA. E-mail: nearing@ecn.purdue.edu
Published online 12 July 2000
Published in 2000 by John Wiley & Sons, Ltd.

used in the application, and that confidence limits be established using ranges of the input values. Finally, the model results are compared to the observed erosion data to determine the coincidence of the model output bands with the observed values. Quinton (1994) used this approach in his evaluation of the EUROSEM model.

Many studies have been made to compare model predictions to measured data. Risse *et al.* (1993) applied the Universal Soil Loss Equation (USLE) to 1700 plot × years of data from 208 natural runoff plots. Annual values of soil loss averaged 3·51 kg m^{-2} with an average magnitude (absolute value) of error of 2·13 kg m^{-2}, or approximately 60 per cent of the mean. Zhang *et al.* (1996) applied the WEPP computer simulation model to 290 annual values and obtained an average of 2·18 kg m^{-2} for the measured soil loss, with an average magnitude of error of 1·34 kg m^{-2}, or approximately 61 per cent of the mean. In both cases the relative errors tended to be greater for the lower soil loss values. Both the Risse *et al.* (1993) and the Zhang *et al.* (1996) studies were conducted without model calibration. Model input parameter values were not adjusted for the specific data used in the comparisons. Given these results and others from similar types of studies (Liu *et al.*, 1997; Rapp, 1994; Govers, 1991), the question remains: are the predictions acceptable? What is an acceptable and expected level of model prediction error?

Implicit in these and other studies of model predictions to measured data is the assumption that the measured data are representative of the erosion rates for the treatment under study. However, it is certainly true that data from soil erosion plots contain a great amount of unexplained variability, which is a critical consideration in using erosion data to evaluate the performance of soil erosion models. Consider the study conducted by Wendt *et al.* (1986), who measured soil erosion rates on 40 cultivated, fallow, experimental plots located in Kingdom City, Missouri, in 1981. They computed coefficients of variation for the 25 storms ranging from 18 to 91 per cent, with 15 of the storms falling in the range of less than 30 per cent. The more erosive storms tended to show the lower degree of variability. Perhaps the most important finding in that study was that 'only minor amounts of observed variability could be attributed to any of several measured plot properties, and plot differences expressed by the 25 events did not persist in prior or subsequent runoff and soil loss observations at the site.'

The results of the Wendt *et al.* (1986) are extremely important for the modeller, and they form the fundamental basis for this current research. The results of the Wendt *et al.* study strongly imply that, given our current ability to measure plot characteristics in the field, the act of adjusting input parameters for a model to represent the differences in measured plot characteristics will not improve model fit for the data from these plots.

Rüttimann *et al.* (1995) reported a statistical analysis of data from four sites, each with five to six reported treatments. Each treatment had three replications. Reported coefficients of variation of soil loss ranged from 3·4 to 173·2 per cent, with an average of 71 per cent. The authors concluded by suggesting 'as many replications as possible' for erosion experiments.

When comparing measured rates of erosion to predicted rates, it is to be expected that a portion of any difference between the two will be due to model error (including structural and input errors), but that a portion will also be due to unexplained variance of the measured sample value from the representative, mean value for a particular treatment. Consider the following example. Let us suppose that the model evaluator is working with a measured value of soil loss of 1·37 kg m^{-2}, and a model predicted value of 1·10 kg m^{-2}. Suppose further that the mean value of the population of measured replicates is 1·30 kg m^{-2}, though one cannot know this from the information given from a single plot. In this example, the prediction error is 1·10–1·37 = −0·27 kg m^{-2} (negative implies underprediction). In this case, of the −0·27 kg m^{-2}, −0·20 kg m^{-2} is due to the fact that the model missed the mean of the population by −0·20 kg m^{-2}, and −0·07 kg m^{-2} is due to the fact that the measured mean varied from the population mean by 1·30–1·37 = −0·07 kg m^{-2}. In essence, the issue for the model evaluator is to partition the 'prediction error' into the portion associated with its two components: the part associated with the difference between the prediction and the population mean, and the part associated with the difference between the individual sample value and the population mean. Sometimes the two factors produce 'prediction error' in the same direction, and sometimes they will partially compensate for one another. Unfortunately, knowledge of variability in soil erosion measurements has been limited in the past.

More recently, though Nearing *et al.* (1999) were able to use data from differences between replicated natural rainfall–erosion plots to quantify coefficients of variation for these plots. They found several important properties of erosion plot variance. Difference in variability between treatments was highly correlated to the magnitude of the measured soil loss. Thus, both the relative differences between replicated plot pair measurements and the coefficient of variation between replications tended to decrease as the magnitude of the measured soil loss increased. They also showed that this relationship between variance and the magnitude of measured soil loss was independent of whether the data were for a single storm, and yearly total, or a multi-year total. The information from that study (Nearing *et al.*, 1999) now provides the means to quantify variation and the means to take into consideration erosion variability when evaluating models using measured data.

A secondary aspect of the data which must be taken into account when evaluating a model relative to data is the observation that models will more often overpredict small, measured soil loss values and underpredict large, measured soil losses (Nearing, 1998). This is both an expected and normal phenomenon which results from applying a deterministic model to data which are naturally variable. However, while it is clear that many models exhibit this over- and underprediction phenomenon, it is not clear that data variability is the sole explanation. This tendency for over- and underprediction at small and large measured soil losses, respectively, must be taken into consideration during model evaluation.

The objective of this study was to construct a method for incorporating erosional variability into a system of model evaluation whereby model predictions are compared to measured erosion plot data. The result is a quantitative, statistically based acceptance criterion for 'goodness of fit' for model predictions.

SOIL EROSION PLOT DATA

The soil erosion plot data used for this study were the same as those used by Nearing *et al.* (1999) in their quantification of replicate variability. Since those data are discussed elsewhere (Nearing *et al.*, 1999; Risse *et al.*, 1993) the list of site information is not repeated here. The data were taken from the repository of the USDA-ARS National Soil Erosion Research Laboratory located in West Lafayette, Indiana. Event values of soil loss were from seven sites in the United States, and there were a total of 2061 replicated storm events in the data set. Annual values of soil loss were used from 13 sites, with a total of 797 replicated pairs of plots. The plots ranged from 2 to 8 m in width, and most were 22 m in length. Slopes ranged from 3 to 16 per cent in steepness.

A DEFINITION FOR A MODEL EVALUATION CRITERIA

The thesis proposed here for defining an evaluation criterion for an erosion simulation model is: if the difference between the model prediction and the measured value lies within the population of differences between the measured data pairs, then the model reasonably reflects the erosion for that population. Another way of looking at this concept is that the replication of an individual plot may be considered as a 'real-world' physical model of that plot. The question, then, of whether or not a simulation model prediction is 'good' is made relative to how well that simulation model performs compared to the physical model as represented by the replicated plot.

As an example, suppose that we estimate for a particular set of conditions that 95 per cent of the values for differences in erosion between replicated plots within the population of replicates will fall within a certain range. To be more specific, suppose that the measured value of erosion, M, is, for example, 2 (units left undefined at this stage), and that we estimate that 95 per cent of the differences between replications for this set of conditions fall within a range of plus or minus 1 unit of the measured value. Let us further suppose that we apply an erosion simulation model to the data and predict an erosion value of P_s. Our null hypothesis is:

$$H_0: \quad P_s - M = 0 \tag{1a}$$

with the alternative:

$$H_1: \quad P_s - M \neq 0 \tag{1b}$$

If $P_s = 3{\cdot}1$ and we use the 95 per cent probability range as our criteria for rejection, then $(P_s - M) = 3{\cdot}1 - 2{\cdot}0 = 1{\cdot}1 > 1{\cdot}0$. We then must reject the null hypothesis, because the difference $(P_s - M)$ did not lie within the 95 per cent occurrence interval of the population of differences between measured values. In this case we conclude that the model is predicting inadequately. The probability of a type I error in this case (that we incorrectly rejected the null hypothesis) is less than 5 per cent ($\alpha = 0{\cdot}05$). Alternatively, suppose the value of the predicted erosion was 2·8, and so $(P_s - M) = 0{\cdot}8$. In this case we cannot reject the null hypothesis. In order to evaluate in this case the probability that we incorrectly failed to reject the null hypothesis, i.e. that we committed a type II error, we would have to suggest an alternative hypothesis, just as would be necessary in testing any statistical hypothesis.

QUANTITATIVE CRITERIA FOR EROSION MODEL EVALUATION

Relative difference, *Rdiff*, in erosion for this study was calculated by the function:

$$Rdiff = \frac{(P - M)}{(P + M)} \tag{2}$$

where P is the predicted erosion value given by either the results of a simulation model, P_s, or from the physical model represented by the replicated plot, P_R. Henceforth we refer to the relative difference from replicated plot data as $Rdiff_R$, and that from simulation model predictions as $Rdiff_S$. The data and analysis used here for computing $Rdiff_R$ are identical to there used in the previous study by Nearing *et al.* (1999) which focused on quantifying the variability. Since it was shown in the previous study that the variance of the replicated plots changed as a function of magnitude of measured soil loss, but not as a function of the measurement period, all of the event, annual and multi-year data herein were combined and analysed together. Each pair of replicates from the data was used to produce two data points. For the first data point one value (A) of the pair was chosen to serve as the 'measured' value, M, of erosion and the other (B) was considered to be the 'predicted' value, P_R, from the physical model. For the second data point, value B was used as the 'measured', M, and value A as the 'predicted', P_R. The resultant physical 'model' was unbiased in that total erosion was the same for both 'measured' and 'predicted' and the same number of data points were underpredicted as were overpredicted. The measured data with greater erosion rates showed, on average, lesser relative differences between replicates for both event and annual data (Figure 1). This result was consistent with the findings of Wendt *et al.* (1986).

The next step in the analysis was to compute occurrence intervals for the data. The data were divided into half log-cycle intervals starting at 0·01 kg m^{-2}. For the data within each division, the 90 and 95 per cent frequency of occurrence of the data points were computed using the empirical distribution function with averaging (SAS Institute, 1988) (Figure 2). Intervals for probability of occurrence were considered for the purposes of this study to be equal to the frequency of occurrence of the data points. The limitation of this approach is the implied assumption that the samples represent the population of relative differences. Since the sample sizes were large, this assumption was considered to be reasonable for the proposed purposes. The advantage of using the assumption was that there was no need to assume a type, or types, or probability density function for the distribution.

The relative difference values, $Rdiff_{occ}$, representing the 90 and 95 per cent frequency of occurrence intervals were linear to the logarithm of the measured soil loss (Figure 2):

$$Rdiff_{occ} = m \log_{10}(M) + b \tag{3}$$

where:

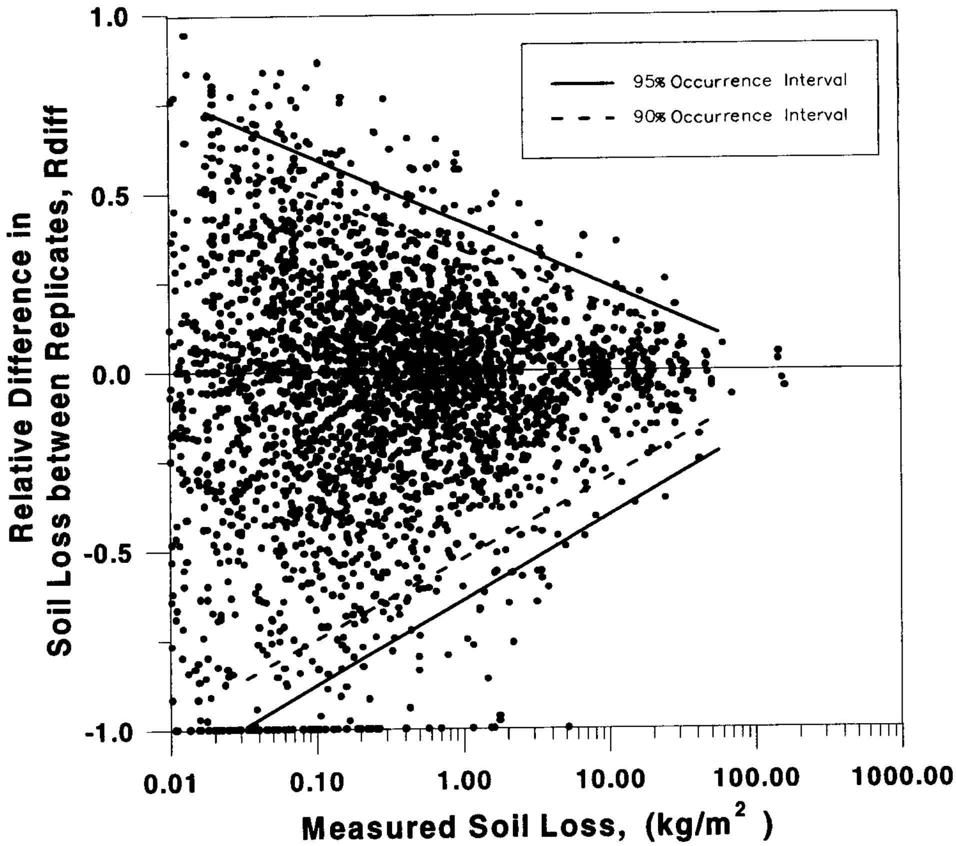

Figure 1. Relative differences in measurements of soil loss between replicated plots, $Rdiff_R$, as computed using Equation 2 vs. the measured soil loss value, M (kg m^{-2})

$m = +0{\cdot}236$, $b = -0{\cdot}641$ and $r^2 = 0{\cdot}97$ for the lower limit of the 95 per cent interval;
$m = -0{\cdot}179$, $b = +0{\cdot}416$ and $r^2 = 0{\cdot}98$ for the upper limit of the 95 per cent interval;
$m = +0{\cdot}225$, $b = -0{\cdot}524$ and $r^2 = 0{\cdot}96$ for the lower limit of the 90 per cent interval;
$m = -0{\cdot}155$, $b = +0{\cdot}338$ and $r^2 = 0{\cdot}98$ for the upper limit of the 90 per cent interval.

EXAMPLES OF MODEL EVALUATION FOR INDIVIDUAL DATA COMPARISONS

Data from the study by Risse *et al.* (1993) of the USLE were used to illustrate the model evaluation concept introduced here. Consider plot 3–28 at Watkinsville, Georgia, in 1954. Corn was grown on that plot that year, and the measured soil loss was 0·72 kg m^{-2}. The USLE prediction for that year, according to Risse *et al.* (1993), was 5·33 kg m^{-2}, and hence the computed $Rdiff_S$ (Equation 2) was +0·76. From Equation 3, we compute the 95 per cent probability of occurrence interval value for this particular measured value, which in this case gives us a range from −0·68 to +0·44. Thus we determine that this $Rdiff_S$ value (+0·76) falls outside the range of expected differences, as determined from values of $Rdiff_R$. Thus we reject the null hypothesis and conclude that Risse *et al.*'s USLE prediction is not equal to the measured value for that case. The probability that we have incorrectly rejected the null hypothesis is less than 5 per cent.

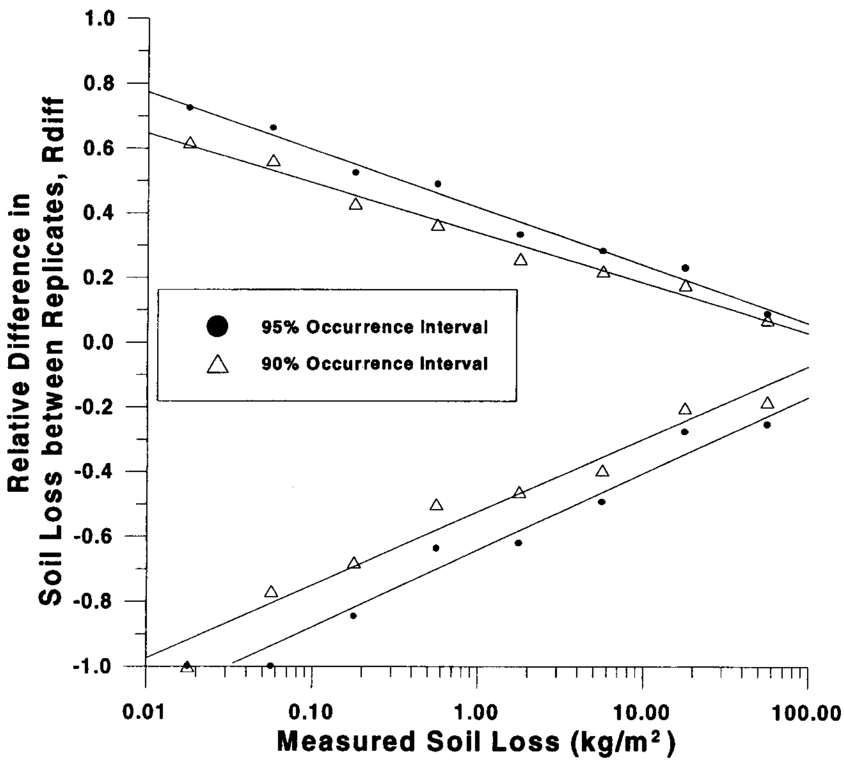

Figure 2. Occurrence intervals of 90 per cent and 95 per cent for the data shown in Figure 1

Now consider plot 1–2 in Tifton, Georgia, in 1954. Peanuts were grown that year, and the soil loss measured from the plot was 0·161 kg m^{-2}. The USLE prediction for that year, according to Risse *et al.* (1993), was 0·273 kg m^{-2}, and hence the computed $Rdiff_s$ (Equation 2) is +0·258. We test the null hypothesis that the model predicted value is equal to the measured value. From Equation 3 we find that this $Rdiff_s$ value falls within the 95 per cent probability of occurrence interval, and thus we do not reject the null hypothesis. For purposes of model evaluation we consider the error to be acceptable.

EXAMPLE OF MODEL EVALUATION FOR LARGER DATA SETS

The method for evaluation of a single data point may be extended to the larger data set, and from the analysis a model *effectiveness coefficient, e*, may be calculated. We define here the effectiveness coefficient as the fraction of simulation model predictions for which a model is effective in predicting the measured erosion, using the acceptance criteria discussed above. Using the 95 per cent occurrence intervals from the replicated erosion data would result in a value, $e_{(\alpha = 0\cdot 05)}$, which is the fraction of data points in the set for which the relative difference between the simulation model prediction and measured erosion value falls within the 95 per cent occurrence interval as given by Equation 3.

As an example, we consider here the studies of Risse *et al.* (1993) for the evaluation of the USLE and of Zhang *et al.* (1996) for the evaluation of WEPP. It is not necessary for our purposes, nor is it within the scope of this study, to justify these previously published studies. We simply use the results to illustrate the evaluation technique proposed here. The procedure was as follows.

1. List the measured and predicted data pairs.
2. Calculate the relative difference between measured and predicted soil loss, $Rdiff_s$, using Equation 2.
3. Compute the 95 per cent occurrence interval as given by Equation 3 for each data point.

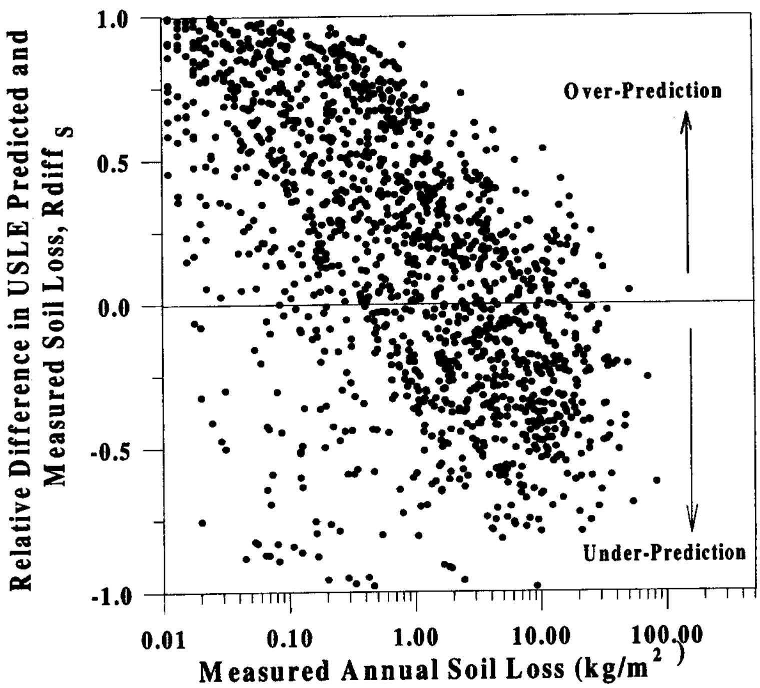

Figure 3. Relative differences between USLE predicted soil loss and measured soil loss, $Rdiff_S$, as computed using Equation 2 vs. the measured soil loss value, M (kg m^{-2})

4. Determine the number of predictions for which the $Rdiff_s$ value fell within the interval.
5. Calculate $e_{(\alpha = 0.05)}$ as the fraction of 'acceptable' predictions for the data set.

The results of the analysis for the annual predictions of the USLE data set made by Risse *et al.* (1993) produced an effectiveness coefficient of $e_{(\alpha = 0.05)} = 0.60$. In other words, of the soil loss predictions made in the study, 60 per cent of the differences between measured and predicted soil loss fell within the expected range of differences for two measured data points within the same population. For the multi-year plot totals of the data and USLE predictions used by Risse *et al.*, we computed an effectiveness coefficient of $e_{(\alpha = 0.05)} = 0.56$. For the Zhang *et al.* (1996) study of the WEPP model, we computed $e_{(\alpha = 0.05)} = 0.66$ and 0·65 for the annual and multi-year predictions, respectively. Since we used the 95 per cent occurrence interval criteria for evaluation, we would not expect the simulation model to have a value of $e_{(\alpha = 0.05)}$ greater than 0·95 unless it were predicting 'better' than the physical model represented by the replicate plots. It is worthwhile to note that the studies of Risse *et al.* (1993) and Zhang *et al.* (1996) were made without calibration to the data.

The graph of $Rdiff_s$ (Figures 3 and 4) for the annual values from the USLE and WEPP studies, respectively, tells a somewhat different story from the graph of $Rdiff_R$ values from the replicated plots (Figure 1). While there was a slight tendency for the measured plot data to be slightly overpredicted in the lower range and underpredicted in the upper range by the replicated plots (Nearing, 1998), the effect was not nearly as apparent as for the USLE and WEPP predictions.

DISCUSSION

Previous quantitative methods of evaluating soil erosion models have not explicitly taken consideration of variability in the measured data. One reason for this is that information on natural, or indeterminable, variability has been lacking. Perhaps another reason is that the extent of the natural variation between

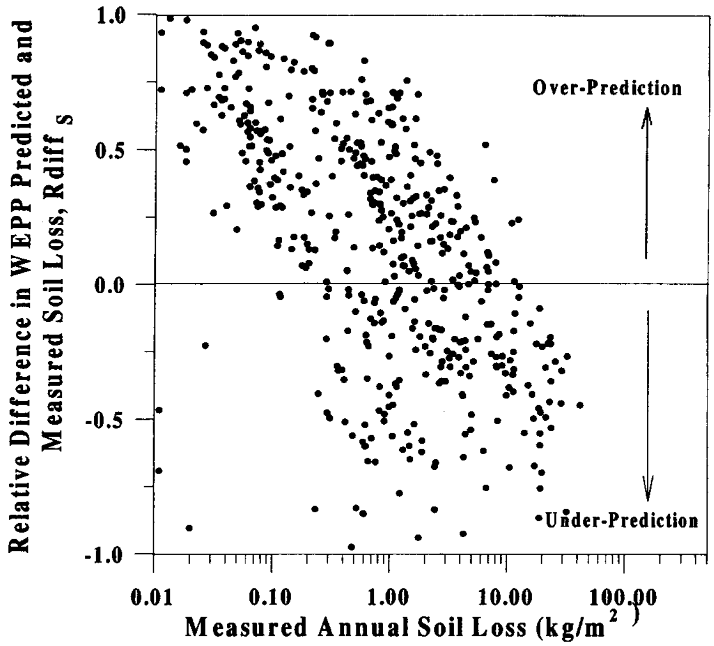

Figure 4. Relative differences between WEPP predicted soil loss and measured soil loss, $Rdiff_S$, as computed using Equation 2 vs. the measured soil loss value, M (kg m^{-2})

replicated erosion measurements has not been generally recognized. The implicit assumption, therefore, for such comparisons is that the data are correct and difference between predictions and measured data is due to model error. Such is not the case. This study, as well as the one by Nearing *et al.* (1999), provide extensive information on variation between replicated erosion plots under natural rainfall which allows us to take this variability into consideration. This study shows that once the natural variation is quantified, it is a relatively straightforward task to incorporate measured erosion variations into a quantitative scheme for model evaluation. In essence, this study proposes that the physical model represented by the replicated plot be set as the standard for the erosion simulation model. Given the state of our knowledge concerning erosion processes and environmental interactions, our lack of ability to adequately characterize plots in order to differentiate or explain differences in erosion from the replicated plots, and our inability to estimate model parameters, one would not expect that the simulation model (of any type) would perform up to the standard of the physical model. This is particularly true for the uncalibrated application of the model, where estimated model parameters were used. Such was the case of the two studies (USLE and WEPP) used as examples in this report.

We set as the specific objective of this study the goal of incorporating and quantitatively considering the issue of erosion data variability in the model evaluation process, and these proposed methods go a long way toward achieving that goal. On the other hand, the larger goal to which the introduction to this paper alludes is less clearly answered, and that is 'What is an acceptable and expected level of model prediction error?' We calculated effectiveness coefficients ranging from 0·56 to 0·66 for studies conducted previously on the USLE and WEPP, both uncalibrated to the data used. That is a relatively narrow range of difference in effectiveness for these two very different models. The USLE is an empirically based model and WEPP is a rather complex, physically based, continuous-simulation computer model. Perhaps an effectiveness on the order of 0·6 is what we can expect from uncalibrated erosion models. We await the results of further studies to see. We would expect that in a calibrated case a model might perform better than $e_{(\alpha = 0{\cdot}05)} = 0{\cdot}6$.

There is usually a tendency for erosion models to underpredict large erosion values and overpredict small ones. Nearing (1998) explained part of the reason for this trend in terms of the natural variability of the measured within-treatment variability. The concept presented in that study was based on the concept that even the 'perfect' model, as defined by a model which accurately predicts mean values of measured treatment replicates, will show the tendency to overpredict the small erosion values and underpredict the larger values of measured erosion rates. The explanation for this is that 'smallness' and 'largeness' of measured values are composed of two components: treatment mean and variance around the treatment mean. The 'smallness' and 'largeness' associated with the variance within treatments are not addressed by the deterministic model (which attempts to predict only the mean for the treatment), and thus the bias in the model predictions relative to the measured values. The concept was illustrated (Nearing, 1998) using the data shown in Figure 1. It is apparent, however, in comparing Figure 1 to Figures 3 and 4 that this factor only explains a portion of the issue of underprediction and overprediction bias. In other words, there is a great deal more underprediction of the large values and overprediction of the small values of measured erosion for the simulation models (Figures 3 and 4) as compared to the physical model represented by the replicate plots (Figure 1). No attempt is made here to explain this phenomenon.

The method presented here is not intended to be, and should not be considered, the complete answer to model evaluation and validation studies. This method proposes an objective measure for the proportion of predictions where the error between predicted and measured erosion values can be justified in terms of natural variation of the measured data. This is an important objective measure, but does not negate the criteria, both subjective and objective in nature, cited by Foster and Lane (1987) for the WEPP model, by Quinton (1994) who worked with the EUROSEM model, or others. Hopefully, however, this study will call to attention the issue of natural variability in measured data and the implications for model evaluation. Measured data are not sacred or infallible. They contain an element of natural variability which the models cannot explain.

The proposed method is limited to application for measured erosion plot data. Plot data are often important in evaluating hillslope erosion model predictions. The wide range of geographic conditions, rainfall regimes, erosion rates, soil types, etc. represented in this data set gives plausible cause to believe that the variations used in this study should be applicable for erosion plot, model validation studies in general. The situation becomes more complicated with regard to watershed or field-scale application. Replications at the watershed scale are inherently more difficult to define at the larger scale, and upscaling variability is not straightforward.

REFERENCES

Foster GR, Lane LJ. (compilers). 1987. User Requirements: USDA-Water Erosion Prediction Project (WEPP). NSERL Report No. 1, USDA-ARS National Soil Erosion Research Laboratory: West Lafayette, IN.

Govers G. 1991. Rill erosion on arable land in central Belgium: rates, controls, and predictability. *Catena* **18**: 133–155.

Liu BY, Nearing MA, Baffaut C, Ascough JC. II. 1997. The WEPP watershed model: III. Comparisons to measured data from small watersheds. *Transactions of the America Society Agriculture Engineers* **40**(4): 945–951.

Nearing MA. 1998. Why soil erosion models over-predict small soil losses and under-predict large soil losses. *Catena* **32**: 15–22.

Nearing MA, Govers G, Norton LD. 1999. Variability in soil erosion data from replicated plots. *Soil Science Society of America Journal* **63**(6): 1829–1835.

Quinton JN. 1994. The validation of physically-based erosion models - with particular reference to EUROSEM. PhD thesis, Cranfield University.

Rapp JF. 1994. Error assessment of the Revised Universal Soil Loss Equation using natural runoff plot data. MS thesis. School of Renewable Natural Resources, University of Arizona.

Risse LM, Nearing MA, Nicks AD, Laflen JM. 1993. Assessment of Error in the Universal Soil Loss Equation. *Soil Science Society of America Journal* **57**: 825–833.

Rüttimann 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.

SAS Institute. 1988. SAS/STAT User's Guide. Release 6·03 Edition. Cary, NC.

Wendt RC, Alberts EE, Hjelmfelt AT Jr. 1986. Variability of runoff and soil loss from fallow experimental plots. *Soil Science Society of America Journal* **50**: 730–736.

Zhang XC, Nearing MA, Risse LM, McGregor KC. 1996. Evaluation of runoff and soil loss predictions using natural runoff plot data. *Transactions of the American Society of Agricultural Engineers* **39**(3): 855–863.