ELSEVIER

Catena 32 (1998) 15–22

CATENA

Why soil erosion models over-predict small soil losses and under-predict large soil losses

M.A. Nearing *

USDA-ARS National Soil Erosion Research Laboratory, West Lafayette, IN 47907-1196, USA

Received 13 June 1997; accepted 23 October 1997

Abstract

Evaluation of various soil erosion models with large data sets have consistently shown that these models tend to over-predict soil erosion for small measured values, and under-predict soil erosion for larger measured values. This trend appears to be consistent regardless of whether the soil erosion value of interest is for individual storms, annual totals, or average annual soil losses, and regardless of whether the model is empirical or physically based. The hypothesis presented herein is that this phenomenon is not necessarily associated with bias in model predictions as a function of treatment, but rather with limitations in representing the random component of the measured data within treatments (i.e., between replicates) with a deterministic model. A simple example is presented, showing how even a 'perfect' deterministic soil erosion model exhibits bias relative to small and large measured erosion rates. The concept is further tested and verified on a set of 3007 measured soil erosion data pairs from storms on natural rainfall and run-off plots using the best possible, unbiased, real-world model, i.e., the physical model represented by replicated plots. The results of this study indicate that the commonly observed bias, in erosion prediction models relative to over-prediction of small and under-prediction of large measured erosion rates on individual data points, is normal and expected if the model is accurately predicting erosion rates as a function of environmental conditions, i.e., treatments. © 1998 Elsevier Science B.V.

Keywords: Soil erosion; Sediment; Hydrologic modeling; Hydrology

1. Introduction

Soil erosion models tend to over-predict erosion for small measured values and under-predict erosion for large measured values. Risse et al. (1993) applied the empirically based USLE model (Wischmeier and Smith, 1978) to simulate erosion from

* Corresponding author. Fax: +1-765-494-5948; e-mail: nearing@ecn.purdue.edu

PII S0341-8162(97)00052-0

208 natural run-off and erosion plots with a total of > 1700 plot years of soil loss data. For both annual and average annual erosion data, the model tended to over-predict the values on the lower end of the scale, and under-predict those on the upper end, although the average erosion predicted for the entire data set was not greatly different from the average measured values. The linear regression parameters for predicted vs. measured erosion rates from Risse et al. (1993) were: (A) $m = 0.59$, $b = 1.16$ kg/m^2, and $r^2 = 0.58$ for the annual values, and (B) $m = 0.77$, $b = 0.42$ kg/m^2, and $r^2 = 0.75$ for the average annual values, where m is the regression slope, b is the y-intercept value, and r^2 is the coefficient of determination. Rapp (1994) performed a study of the Revised USLE (RUSLE) using the same data as that used by Risse et al. (1993) and found similar results. In the case of RUSLE, the linear regression parameters for predicted vs. measured erosion were: (A) $m = 0.49$, $b = 1.44$ kg/m^2, and $r^2 = 0.58$ for the annual values, and (B) $m = 0.64$, $b = 0.91$ kg/m^2, and $r^2 = 0.75$ for the average annual values (Rapp, 1994).

Results from testing of physically based, computer simulation models of soil erosion have produced similar results. Zhang et al. (1996) applied the Water Erosion Prediction Project (WEPP) model to data from 65 natural run-off plots from eight locations in the US. The data included 556 annual values and 4124 event values of erosion. Zhang et al. (1996) found that small erosion values tended to be over-predicted, and large erosion values tended to be under-predicted for event-by-event, annual totals, and annual average soil loss. This trend was consistent with that observed by Ghidey et al. (1995), as well as Kramer and Alberts (1995) in their applications of the WEPP model to other data sets. Zhang et al. (1996) also found a similar trend for the case of predicted vs. measured run off, with the exception of the calibrated annual average values of run off. It is interesting that even for the calibrated case, where the baseline soil infiltration parameter was adjusted to produce essentially matching total average annual run off, the annual (year-to-year) and storm run- off values also exhibited over-prediction at low measured values, and under-prediction at large measured run-off values.

Natural variation in soil loss data is large. Wendt et al. (1986) performed a study on 40 replicated, fallow, natural run-off plots near Kingdom City, MO and found that the coefficient of variation (CV) in measured soil loss on individual storms ranged from 18 to 83% and was dependent on the level of erosion observed (CV decreased with increasing mean soil loss). In that study, the authors could find no statistically significant relationships between soil loss and plot characteristics. The minor variations in soil type or slope were not correlated to soil loss, and run-off volume correlated to soil loss only for the storms that produced low run off. No spatial trends were found between plots, and no inherent plot differences relative to soil loss were observed. In other words, these were, for all practical purposes, replicated plots that would be modeled with identical model input parameters, and thus which would result in a single prediction value for each storm for all the plots. The implication of the study of Wendt et al. (1986) for erosion prediction is that, there is a limit to the accuracy of deterministic models because of the variation in soil erosion rates, which may be considered random from a practical standpoint. This is true irrespective of model type, whether empirical or physically based. The nature of the observed variation is discussed below.

The hypothesis of this study is that the consistently observed bias for soil erosion models to over-predict low measured rates and under-predict large measured rates is due, at least in part, to the fact that the models are deterministic in nature, and the measured data has a significant random component for which the models cannot account within the deterministic framework. A simplified, synthetic example is presented to illustrate in a general manner the basic nature of the concept. The objective of this study was to test this hypothesis by using 6014 measured soil loss data points from storms on replicated natural rainfall and run-off plots under both fallow and cropped conditions. In this study, the first plot is treated as the 'measured' soil loss value, and the replicated plot is treated as the unbiased physical model, i.e., predictor, of the measured plot data.

2. Conceptual considerations

Consider a 'perfect' deterministic soil erosion model. This model accurately predicts soil erosion as a function of treatment, i.e., for any particular combination of soil, plot size and shape, erosive inputs, plants, tillage, and other measurable factors, the model predicts the mean value of soil erosion for a population of replicates. As an example, we use the data from Wendt et al. (1986). They reported mean values and coefficients of variation for the data from individual storms during a 1-yr period for the 40 replicated, fallow plots in their study. They also reported that, except for the smaller events, the samples were normally distributed for individual storms. Thus, for demonstrative purposes, we consider here the 15 events with reported mean soil loss greater than 0.1 kg/m^2 (1 Mg/ha) and assume a normal distribution of replicates for each storm. We further assume that our hypothetical, perfect, deterministic model predicts the correct mean value of soil loss for each of the 15 storms as reported by Wendt et al. (1986), and we choose 40 data points, evenly spaced relative to probability of occurrence, from a normal distribution using the storm-by-storm means and variations as reported in the paper. Note that we do not assume a normal distribution for the data set as a whole, but only for within-storm replicates. The data for the series of storms as a whole was skewed. Fig. 1 shows a graph of the results of our model in terms of predicted vs. measured soil losses.

The average soil loss across plots for each storm is identical between our 'perfect' deterministic model and the measured data, and the total soil loss for the series of storms is the same for the model and the data. The model under-predicted half and over-predicted half of the total of 600 data points. Yet the regression line between predicted, P_{sl}, vs. measured, M_{sl}, soil loss was

$$P_{sl} = 0.903 \times M_{sl} + 0.0611 \quad (r^2 = 0.90). \tag{1}$$

For the range of measured values of less than 0.15 kg/m^2, 29% were under-predicted and 78% (102/131) were over-predicted, and for the upper range of M_{sl} greater than 0.8 kg/m^2, 67% (136/202) were under-predicted and 33% were over-predicted. In the mid-range, approximately half were under-predicted and half were over-predicted.

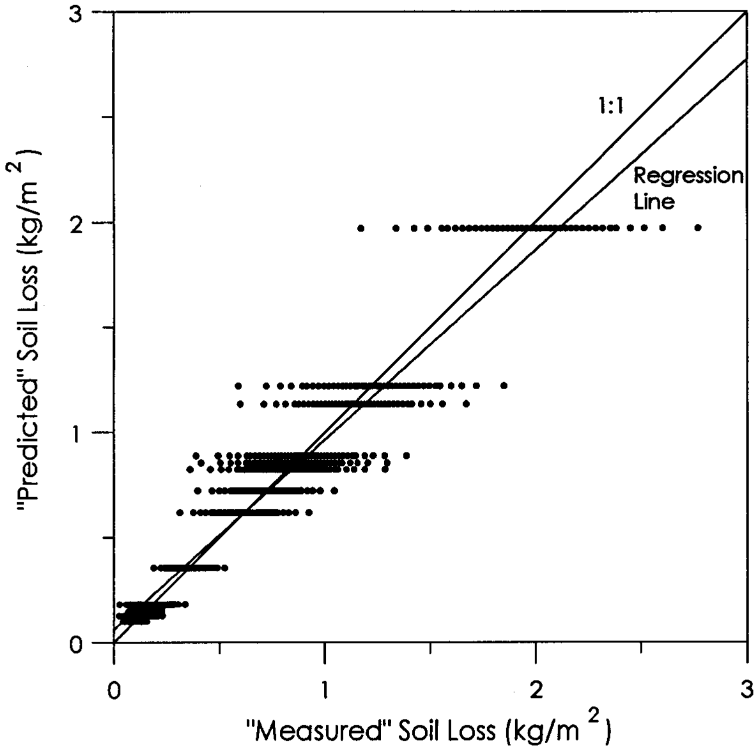

Fig. 1. 'Predicted' vs. 'measured' soil loss for the hypothetical, perfect, deterministic erosion model as applied to the data from Wendt et al. (1986). 'Predicted' values are the mean values of the 40 replicated erosion plots, and the 'measured' values are evenly spaced relative to probability of occurrence assuming normal distributions using the storm-by-storm means and variations as reported in Wendt et al. (1986).

The observed result of model bias is simply due to the fact that the 'model' used is deterministic, and the data contains random variation within treatments, rather than the fact that the model is biased relative to the means of the treatments. In other words, the phenomenon is mathematical in nature rather than a function of any bias inherent in the model itself. In this example, the assumption of normal distribution was used. The results would be slightly different quantitatively for different types of distributions, but qualitatively similar regardless of the assumed distribution type.

3. Data from replicated plots

Perhaps the best possible model for predicting the erosion from an area of land is a physical model of the area that has similar soil type, land use, size, shape, slope and erosive inputs, i.e., a replicated measurement. To test our hypothesis concerning the bias of deterministic erosion models for predicting erosion for individual plot measurements, we extracted information for pairs of replicated plots from the repository of soil loss data located at the USDA-ARS National Soil Erosion Research Laboratory. A total of 3007 pairs of data for individual storms were chosen for analysis (Table 1). Two data points were obtained from each pair of data. For the first data point, one value (A) of the pair was chosen to serve as the 'measured' value of erosion and the other (B) was considered to be the 'predicted' value from the physical model. For the second data point value (B) was used as the 'measured' and value (A) as the 'predicted'. The resultant 'model' was unbiased in the sense that total erosion was the same for both 'measured' and 'predicted', and the same number were under-predicted as were over-predicted.

The average soil loss rate per storm measurement was 0.222 kg/m^2. The regression line between predicted, P_{sl}, vs. measured, M_{sl}, soil loss was (Fig. 2)

$$P_{sl} = 0.876 \times M_{sl} + 0.0275 \ \ (r^2 = 0.77). \tag{2}$$

For the 2005 data points with the lowest 'measured' soil loss ($M_{sl} < 0.007$ kg/m^2), 50.4% were over-predicted, 34.2% were under-predicted, and 15.4% were reported as the same. For the 2005 values in the mid-range (0.007 $kg/m^2 < M_{sl} < 0.083$ kg/m^2), 40.5% were over-predicted, 42.8% were under-predicted, and 16.6% of the pairs had identical reported erosion rates. For the 2006 data points in the upper range of measured

Table 1
Site, copping and management, and data collection period for the replicated plot data used in this study

Site	Cropping and management	Years of record
Holly Springs, MS	Fallow	1961–1968
	Turn-plow corn	1961–1968
	Meadow, corn rotation	1961–1968
	Corn, soybeans rotation	1970–1980
	No-till corn and soybeans	1970–1980
	Conventional corn and soybeans	1970–1980
Madison, SD	Fallow	1962–1970
	Turn-plow corn	1962–1970
	Conservation tilled corn	1962–1970
	Conservation tilled oats	1962–1964
Morris, MN	Fallow	1962–1971
	Corn, oats, meadow rotation	1962–1971
Watkinsville, GA	Fallow	1961–1967
	Turn-plow corn	1961–1967
	Turn-plow cotton	1961–1967
	Corn, meadow rotation	1961–1967
Presque Isle, ME	Fallow	1961–1965
Pendelton, OR	Fallow	1980–1989
Tifton, GA	Fallow	1960–1966

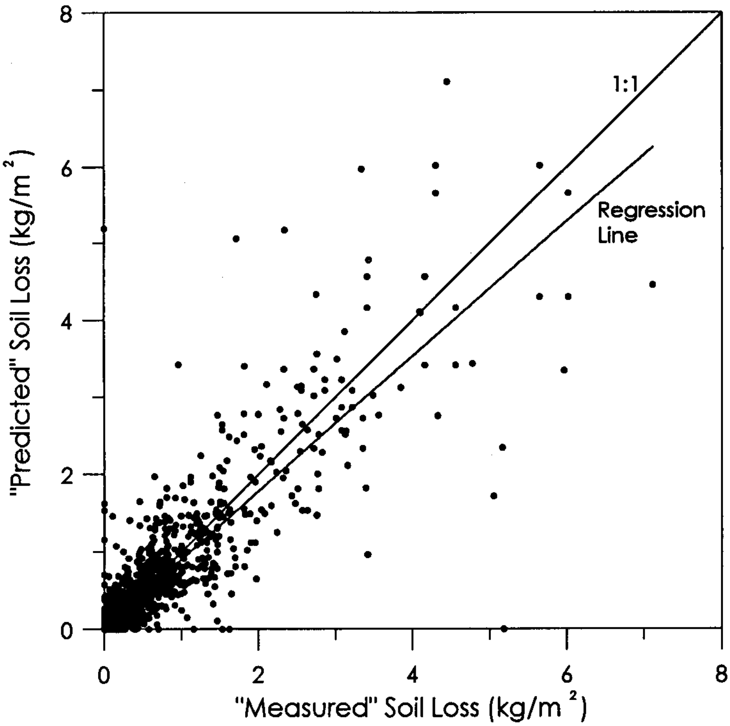

Fig. 2. 'Predicted' vs. 'measured' soil loss for the physical model of soil erosion as represented by pairs of replicated plots.

soil loss ($M_{sl} > 0.083$ kg/m^2), 41.6% were over-predicted, 55.5% were under-predicted, and 2.9% were reported as the same.

4. Discussion

The soil erosion model is a tool for identifying sources of soil erosional variances as a function of measurable quantities of the system of interest. This study is predicated on the premise that a practical limitation exists in defining an erosional system for the purposes of prediction. The study of Wendt et al. (1986) certainly supports this premise. There are sources of contributing variability that could theoretically be measured and which could potentially further explain the measured variance in soil loss between the replicates of the Wendt et al. study or other similar conditions. For example, local variations in rainfall intensities from plot to plot may have played a role, although the 15 non-recording gauges in the plot area showed only minor differences. There are, however, practical limitations to measurements. If one were to cover the plot with a rain gauge to measure the instantaneous, spatially distributed rainfall characteristics over the plot, the resultant run off and sediment generated would be affected. Another example is plot-to-plot variations in microtopography, and the way in which the microtopography evolves as a result of rainfall, run off, and erosion.

On the other hand, there may exist basic considerations relative to the physics of the erosion process that cause variations and which are essentially unaccountable. Lei et al. (1997) recently developed a finite element model that mimics the spatial and temporal evolution of rill development over time. The results of that numerical study indicate that sediment concentration in a rill will oscillate spatially in an apparent random fashion downstream in a rill even for conditions of constant flow discharge with time and downstream distance. Rills in the laboratory exhibit these alternating regions of deposition and detachment for the steady-state, uniform initial slope case that were effected by small-scale random variations in the initial conditions of the rill (Lei et al., 1997). Depending upon where the rill were to be sampled for sediment, a given rill may exhibit large variances in measured sediment discharge. Thus, the sediment discharge from a rill cannot be predicted without variance. The situation would certainly be much more complicated in the field with a variety of complex factors at work.

The results of this study indicate that the application of a deterministic model to data which, by their nature, contain 'natural variation', i.e., variation that the model is not capable of capturing, will effect a bias in the erosion predictions relative to values on the higher end vs. those on the lower end of the range of measured values. The results of the study do not suggest that this factor is necessarily the only one at work to cause this bias, but it certainly is significant. In our simple example presented above for the data from the Kingdom City plot data (Wendt et al., 1986), the variance values used may actually be lower than that which might be found in some other erosion data because of the fact that the plots were fallow. Cropped plots could be expected to have a greater level of variance between replicates because of the opportunity for more variation in the surface configuration associated with residue and canopy cover, although data to support this is not existent. An increase in variance would cause an increase in the apparent model bias.

An additional implication of this study is that there are practical limits on the level of fit that should be expected between measured and predicted values of erosion. The 'perfect' deterministic model discussed above for the data from Wendt et al. (1986) produced a r^2 value of 0.90, and the physical model represented by replicate plots for the case of individual storm data produced a r^2 value of 0.76. It should not be expected that an erosion model would give better overall results than those reported here.

There are advantages and disadvantages to various types of erosion models. Generally, empirical models such as the USLE tend to be easier to use because of the small number of input values required, and have less potential for the introduction of prediction errors because of uncertainty in the model input values relative to physically based models. Physically based erosion simulation models have a much more sophisticated model structure which, in theory, allows them to better describe the influence and interactions of many and various factors that influence erosion. Physically based models also often provide a different type of information than that given in empirical models. For example, the WEPP model (Flanagan and Nearing, 1995) estimates the spatial and temporal distributions of soil loss, sediment yield, sediment size characteristics, run-off volumes, soil water balance, and a myriad of other types of system information that the

USLE cannot provide. The USLE was designed only to predict long-term, average annual soil loss. However, because of inherent limitations in prediction capabilities as discussed in this study, as well as for other reasons associated with the trade-off between model complexity and the definition of model input values, one should probably not expect the physically based, deterministic models to predict more accurately the rates of erosion from specific land areas.

On the other hand, it is suggested here that a limitation of current erosion models is their deterministic nature, and the physically based models may have the greater potential as compared to the empirical models in moving from the deterministic framework to a probabilistic one (Singh et al., 1988; Wright and Webster, 1991). Monte Carlo simulations with the WEPP model, for example, have shown promise (Tiscarino-Lopez, 1994; Deer-Ascough, 1995).

References

Deer-Ascough, L.A., 1995. A framework for uncertainty analysis of complex process-based models. PhD Thesis, Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN, USA.

Flanagan, D.C., Nearing, M.A., 1995. USDA water erosion prediction project: Hillslope profile and watershed model documentation. NSERL Report No. 10, USDA-ARS National Soil Erosion Research Laboratory, West Lafayette, IN 47097-1196 USA.

Ghidey, F.E., Alberts, E.E., Kramer, L.A., 1995. Comparison of run-off and soil loss predictions from the WEPP hillslope model to measured values for eight cropping and management treatments. ASAE Paper No. 95-2385, Am. Soc. Agric. Eng. Publ., St. Joseph, MI.

Kramer, L.A., Alberts, E.E., 1995. Validation of WEPP 95.1 daily erosion simulation. ASAE paper No. 95-2385, Am. Soc. Agric. Eng. Publ., St. Joseph, MI.

Lei, T., Nearing, M.A., Haghighi, K., 1997. A model of soil erosion in a rill: Temporal evolution and spatial variability. Am. Soc. Agric. Eng. Paper 97-3068. ASAE Meeting, 10–14 Aug 1997, Minneapolis, MN.

Rapp, J.F., 1994. Error assessment of the revised universal soil loss equation using natural run-off plot data. MS Thesis, School of Renewable Natural Resources, University of Arizona, Tucson, AZ.

Risse, L.M., Nearing, M.A., Nicks, A.D., Laflen, J.M., 1993. Assessment of error in the universal soil loss equation. Soil Sci. Soc. Am. J. 57, 825–833.

Singh, V.P., Krstanovic, P.F., Lane, L.J., 1988. Stochastic models of sediment yield. In: Anderson, M.G. (Ed.), Modelling Geomorphological Systems. Wiley, Chichester, pp. 259–285.

Tiscarino-Lopez, M., 1994. A Bayesian–Monte Carlo approach to assess the uncertainties in process-based continuous simulation models. PhD Thesis, School of Renewable Natural Resources, University of Arizona, Tucson, AZ.

Wendt, R.C., Alberts, E.E., Hjelmfelt, A.T. Jr., 1986. Variability of run-off 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. Agricultural Handbook 537. USDA, Science and Education Administration, Washington, DC.

Wright, A.C., Webster, R., 1991. A stochastic distributed model of soil erosion by overland flow. Earth Surf. Processes Landforms 16, 207–226.

Zhang, X.C., Nearing, M.A., Risse, L.M., McGregor, K.C., 1996. Evaluation of run-off and soil loss predictions using natural run-off plot data. Trans. Am. Soc. Agric. Eng. 39 (3), 855–863.

ELSEVIER

Catena 38 (1999) 89–91

CATENA

www.elsevier.com/locate/catena

Correspondence

Response to 'Letter to the editor' by Mac Kirby and Richard Webster regarding my recent paper in Catena: 'Why soil erosion models over-predict small soil losses and under-predict large soil losses'

M.A. Nearing

USDA-ARS, National Soil Erosion Research Lab., 1196 Soil Building, West Lafayette, IN 7907-1196, USA

The author is happy to have received the attention of Dr. Kirby and Dr. Webster regarding my recent article in Catena. Dr. Kirby and Dr. Webster are, in my opinion, to be commended for bringing to fore the suggestion that when regression models are used to develop erosion prediction models, there are alternatives to the traditional least-square error approach to curve fitting. Unfortunately, they seem to be fundamentally confused about the paper in question.

The first two sentences of the comment state that the models that I used are empirical and 'essentially regression' based. This is simply not true. Neither of the two models that I use in the paper (Figs. 1 and 2) is derived from regression. The first is simply a hypothetical model that predicts the means of treatments. The second is the 'model' of the replicate plot. The regression lines in the figures have nothing whatsoever to do with the development of any model. I used regression only to show that there exists a tendency for these models to under-predict at the high end and over-predict at the low end. In fact, regression is irrelevant to the argument. The same point is made simply by counting the number of over- and under- predictions within specified ranges, which I also did in the paper.

Secondly, and more importantly, the comments in the letter by Dr. Kirby and Dr. Webster are focussed primarily on the development of models, rather than the evaluation of models. There is a fundamental difference between the two goals, particularly concerning the issue of regression. They refer to the regression of a measured Y value on a measured independent variable X, and the desire 'to predict future erosion'. In other words, they refer to model development. In that case, certainly the use of Y on X regression may not be preferable, as is well described in the letter by Dr. Kirby and Dr.

PII: S0341-8162(99)00057-0

Webster. In the Catena paper, however, I discuss model evaluation, and in particular the comparison of model predictions to measured data. In this case, it is important that the Y value be evaluated relative to the X value, because the model evaluator is interested in how the model prediction compares to a measured data point. Thus, when one uses a model to predict a measured data point, the terminology and fundamental concept of over- or under-prediction is made in reference to the measured value, not the predicted one. If the model is 'over-predicting', it means that the value of the prediction is greater than the value of the measured. The evaluation of the model only works in one direction, i.e.; the measured value is not to be 'evaluated' relative to the predicted one. From that perspective, regression of, and comparison of, the predicted value to the measured is the appropriate perspective. One will find it important to know, for example, that for smaller measured values (of soil loss in this case) the model will over-predict more often than not.

We often conduct model evaluation studies using only a few data points. Recently, Liu et al. (1997) made a comparison between WEPP (A physically- based model) and data from 15 watersheds. For a watershed model evaluation, 15 is a large number. There was the normal scatter in the data, but the paper drew a significant amount of criticism (from the journal reviewers) because two of the watersheds in meadow, which had extremely low measured erosion rates, were over-predicted by a large amount. Actually, the authors (myself included) were quite happy with the results, because regardless of the fact that the predicted values were 'too high', the predicted values were still low in an overall sense with the other data. The other aspect of this, which is addressed only in passing in the Catena paper discussed here, is that readers and reviewers often have an inflated view of the levels of correlation they expect from a model. It is probably not realistic to expect (for the uncalibrated case) that the r^2 can or should be better than that of the replicated plot data ($r^2 = 0.77$) in the Catena paper, i.e., Fig. 2.

The letter to the editor also makes the claim that 'From Nearing's discussion it seems that he and other erosion scientists want the 1:1 line for their predictions.' It is not clear to me that my discussion in the Catena paper indicates a 'want' for a 1:1 line for predictions. Quite the contrary. The basic concept of the paper was to illustrate that 'over-prediction of small and under-prediction of large measured erosion rates * (are) normal and expected if the model is accurately predicting erosion rates as a function of environmental conditions.'

My point in the Catena paper was that when one conducts an evaluation of a model using measured erosion data, one should expect that there will be a tendency for the model to produce more over-predictions on the low end of the scale and more under-predictions on the upper end. This is true when a model is deterministic - in the sense that the model attempts to predict a mean value for a treatment - rather than probabilisitic - in the sense that the model attempts to predict a population of probable or possible responses for a given treatment. It is really a very simple idea, and one which is clearly expected, but it seems to have thrown many erosion modelers off when they find this phenomenon to occur. It is clear that regression is not needed to make this point in my paper.

Dr. Kirby and Dr. Webster suggest something quite different. They suggest using an alternative method of regression to develop an erosion model if a regression-based approach to model development is used. The idea proposed is interesting and promising for the case of model development, but irrelevant to the thesis of my Catena paper.

References

Liu, B.Y., M.A. Nearing, Baffaut, C., J.C. Ascough II. 1997. The WEPP watershed model: III. Comparisons to measured data from small watersheds. Trans. of ASAE 40(4): 945–951.