CAPABILITIES AND LIMITATIONS OF EROSION MODELS AND DATA

M.A. Nearing[A]
[A] Southwest Watershed Research Center, USDA-ARS, Tucson, AZ, USA.

Abstract
Variability in soil erosion data from replicated plots is large. One might think of the replicated plot as the best "real-world, physical model" of soil erosion, and that the physical model represented by the replicate plot represents a best-case scenario in terms of erosion prediction. In this study, replicated plot pairs for 2061 storms, 797 annual erosion measurements, and 53 multi-year erosion totals were used to estimate the natural variance of erosion data. Coefficients of variation ranged on the order of 14% for a measured soil loss of 20 kg/m^2 to greater than 150% for a measured soil loss of less than 0.01 kg/m^2. The r^2 for the fit for the replicate plot model was 0.76. This fit sets a benchmark for what one can expect for soil erosion models in general. This paper also discusses the critical nature of continuous simulation modeling in predicting erosion reliably. Results of simulation testing with the WEPP model indicate that 60 to 200 years of continuous simulation are required in order to quantify erosional response to plus or minus 10%. Single storm models do not have the capacity to accurately characterize erosional response of the complex and dynamic erosional system.

Additional Keywords: soil erosion, soil conservation planning, uncertainty, data variability

Introduction
The purpose of this paper is to present a realistic overview of soil erosion modeling capabilities and limitations. The data and model applications will focus on hillslope scale processes, but have obvious implications for sediment generation and sediment yield across larger scales as well. There are three major points upon which the paper will focus. The first has to do with variability of erosion in nature and its implications for erosion prediction. There have been many studies of soil erosion model application and validation using measured erosion data, but it is difficult to get a general or broad perspective and quantification of variability until you have relatively large data sets to work with, which is rare. It is also difficult to address variability unless one can replicate experiments, which is a challenge for watersheds. Here we will focus on plot scale erosion and variability associated with hillslope erosion. The basic message is that erosion in nature at the hillslope scale is quite variable, and that variability has major implications for models and prediction. On the other hand, patterns are evident, and we will discuss those patterns that are observable.

The second issue has to do with the importance of continuous simulation for erosion modeling. By continuous simulation, we refer to a model that calculates erosion through the year and over many years. Most importantly, it is a model that has the capability to update the parameters that define the state of the system as it influences erosion resistance or susceptibility, such as standing plant biomass that acts as soil cover, plant residues in contact with the surface, soil moisture, soil consolidation, etc. We will argue that one cannot effectively evaluate land use scenarios without a reliable form of continuous simulation. Lastly, we will argue that soil erosion models can be used as effective tools for many purposes, as long as they are used with the understanding of their capabilities and limitations.

Materials and Methods
In order to evaluate variability of measured soil erosion and expectations for model predictions we introduce the concept of the physical model of soil erosion. One can argue that the best possible model for erosion from a given plot will be the physical model, that is, a replicate of the plot on the same hillslope with the same slope, soil, land use, and weather input. By "the same", it is meant that we would characterize the plots the same for the purposes of modeling the erosion. Another way of looking at it is that the measured erosion values from the two plots would be samples from the same treatment distribution.

A large number of experimental natural rainfall-erosion plot data were used for the analyses presented. For each section of the information presented, data from some or all of the following locations in the United States were used: Holly Springs, MS; Madison, SD; Morris, MN; Presque Isle, ME; Watkinsville, GA; Bethany, MO; Guthrie, OK; Castana, IA; Tifton, GA; Pendleton, OR; Geneva, NY; and Kingdom City, MO. The experimental erosion plots used here represent a wide range of cropping systems, including fallow, cotton, grass-corn-oats, alfalfa, wheat-clover-cotton, bermuda grass, red clover, winter rye, fall-tilled corn, conservation-tilled corn, no-till corn, oats, no-till corn & soybeans, no-till soybeans, conventional-tilled soybeans, and potato.

Results and Discussion

Using the data from replicated erosion plots, we were able to first of all obtain an idea of the variance associated with the erosion data from the plots. The details of the methodology that was used to generate the graph (Fig. 1) of the coefficient of variation in the measured data, expressed as a fraction, versus the magnitude of soil loss measured on the plots is presented in Nearing et al. (1999). There are several key points to be made. One is that the level of variance in the measured data is high in general. At a measured soil loss level of 0.1 kg/m2, which translates to 1 ton per hectare, the coefficient of variation is approximately 1, or in other words, 100%.

The second obvious point here is that the level of variance between plots was dependent on the magnitude of soil erosion that was measured (Fig. 1). At low erosion levels, variance was quite large. As erosion level increases, we see the coefficient of variation reduced to tens of percent. Implicit in this, but not stated in the graph, is that other system parameters such as the geographic location, type of soil, and crop type, did not enter into the picture for explaining the differences in variance found in the data. Variance was, as far as was discernable from the data, a function only of the magnitude of soil loss measured.

What is not clear in this graph (Fig. 1) is that the x-axis of this graph represents measured soil loss for the plots over three different time scales: events, individual year values, and average annual erosion values. In other words, the variance level depended on the magnitude of soil erosion measured, but it did not matter over what time period the erosion was measured. An overlay of the same graph for event data exactly overlaps the same graph for annual average erosion, though event data values on average were less than the annual data values.

Figure 1. Coefficients of variation between replicated plots as a function of magnitude of measured soil loss (from Nearing et al., 1999).

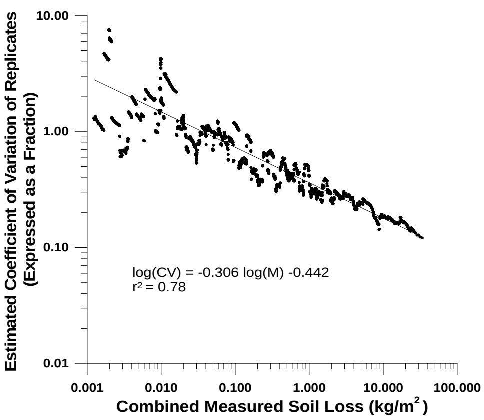
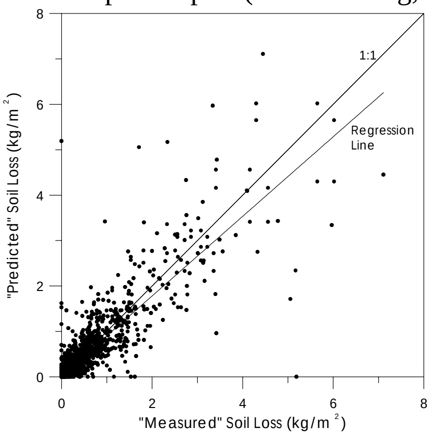


Figure 2. "Measured" vs. "predicted" soil loss for the physical model as represented by the the replicate plot (from Nearing, 1998).



These results have significant implications for modeling erosion. One immediate implication of data variability is that there is a limit in terms of accuracy for models. For example, using the data from replicated plots, Fig. 2 represents prediction accuracy for the best-case "physical model" (Nearing, 1998). All of the plot data was paired by replication, and event soil loss was plotted with one plot assigned as the treatment, and the other plot as the physical model. The level of fit obtained, in this case an r^2 of 0.77, can be considered as a benchmark, or "best-case" prediction scenario. One cannot reasonably expect a simulation model to fit better than this. Using the information from the previous graph, the coefficient of variation for measured data in the range of 1 to 5 kg/m^2 is of the order of 30 to 50%. If the measured data for the physical model in this graph were lesser in magnitude, the fit would reduce accordingly.

Another implication of variance in data is that it is much harder to predict low erosion rates than to predict high ones. So, for example, even though we know that the *relationship* between variance and soil erosion magnitude does not appear to depend on whether the erosion is from single storms or from long term averages, erosion values on average will tend to be lower for individual storms, and erosion predictions will tend to be poorer.

Figure 3 shows single storm predictions using the WEPP model (Zhang et al., 1996). WEPP is a process-based, continuous simulation model for soil erosion. It contains a model for predicting soil erosion for daily storm events, but also auxiliary models for plant growth and canopy cover; residue production, decay and burial; tillage; soil consolidation; soil moisture; infiltration; runoff; and many other system dynamics. Here Zhang et al. (1996) have used the model, un-calibrated, for 2119 storm events, and received a fit of approximately 0.4, which compares to a fit of 0.77 for the physical model results. One way of interpreting this is that the WEPP model is predicting the events approximately half as well as would our ideal physical model of the replicate plot. There obviously may be some room for improvement here in the WEPP predictions, but here at least we have a more realistic idea of where we stand and how much we could improve if our model was "perfect". It is important to stress that these predictions were using the model in the un-calibrated state. One finds very few published evaluations of un-calibrated erosion models, though this is how we most often need to apply them for solving problems.

Figure 3. Measured vs. predicted soil loss for daily results of the WEPP model for 2119 storm events (data from Zhang et al. ,1996).

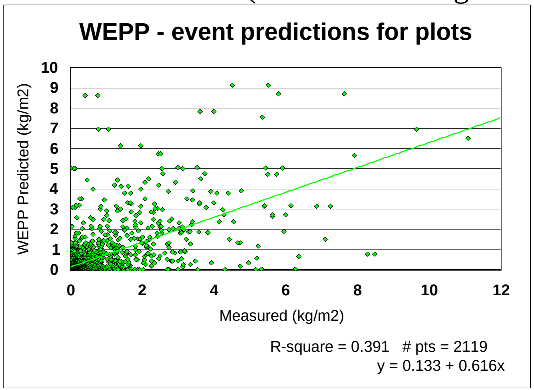
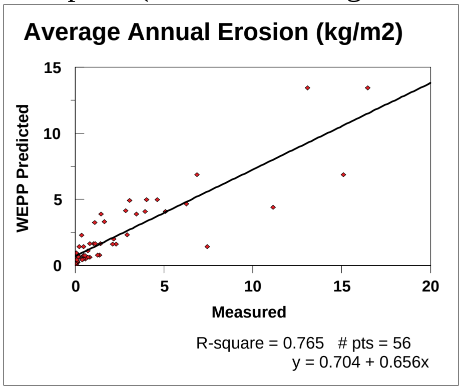


Figure 4. Measured vs. predicted soil loss for average annual results of WEPP for 56 natural runoff plots (data from Zhang et al., 1996).



When looking at the predictions of the annual average erosion rates using the same data, things look much better (Fig. 4) (data from Zhang et al., 1996). The fit in this case was much higher, at an r^2 of 0.77. This is a result of the fact that the random variation inherent in the plot data tends to smooth out when we look at higher erosion rates, which also often happens to be correlated to measuring over longer time scales. But the key point is that it is not, apparently, the time of record itself that governs the level of accuracy that can be and is achieved, but rather the magnitude of erosion measured. The effect is correspondent to Fig. 1, wherein the variance in the measured data decreased as a function of measured soil erosion magnitude.

Figure 5 shows similar data for average annual erosion using the Universal Soil Loss Equation. The accuracy of the results is approximately similar as that for WEPP. According to its developers, the USLE was never intended to be used as a event model, but only for predicting annual averages. However, if one looks at the structure of the model one sees that the average annual erosivity factor, R, is simply an average annual summation of individual storm erosivities, or EI_{30}s. Hence there is no fundamental reason why the USLE could not be used as an event model. The probable reason that the USLE was designated to be used only for annual averages was that the developers had access to enough data to know that predicting erosion for individual events, particularly with an uncalibrated model on a routine basis, was simply not possible with any reasonable level of accuracy. The individual event predictions using the USLE probably would not differ in accuracy much from the WEPP predictions for the same data.

Another approach that can be taken to the problem of validation, application, and calibration of models is the use of the event soil loss frequency distributions. Here (Fig. 6) we have a frequency distribution of measured and predicted soil erosion plotted in terms of recurrence interval (Baffaut et al., 1998). They found that even though the fit for measured vs. predicted events is usually relatively low, such as the r^2 of 0.39 shown in Fig 3, the frequency distribution of soil loss may compare well with the measured data. Baffaut et al. (1998) also showed that the frequency distribution of events can be used for calibration purposes. In general, the lower end of the frequency curve, or the small events, tend to be dominated by splash erosion. Thus the lower part of the curve can be used to calibrate the data for the splash or interrill parameters. The upper end of the curve is dominated by rill erosion, and correspondently the rill erodibility parameters can be calibrated on that portion of the curve.

Figure 5. Coefficients of variation between replicated plots as a function of magnitude of measured soil loss (data from Risse et al., 1993).

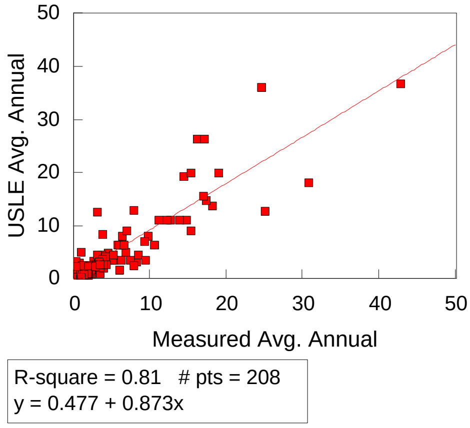
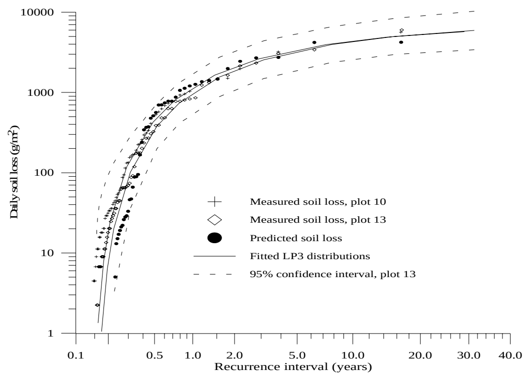


Figure 6. Comparison of distributions of measured and predicted daily soil losses on the fallow plots in Morris, MN from 1962-1971 (from Baffaut et al., 1998).

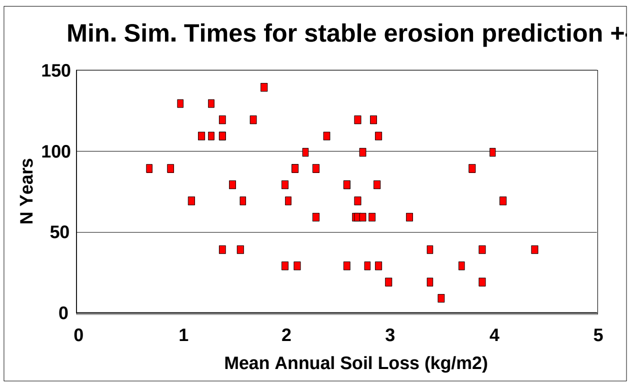


Much of the discussion above, including the idea of using long-term modeling averages such as average annual and frequency distributions of individual events, relies implicitly on the idea of utilizing continuous simulation models for predicting erosion. A large number of the models that are being used are not continuous simulation, but rather single event models. By a continuous simulation model, we refer to a model that calculates erosion through the year and over many years. Most importantly, it is a model that has the capability to update the parameters that define the state of the system as it influences erosion resistance or susceptibility, such as standing plant biomass that acts as soil cover, plant residues in contact with the surface, soil moisture, soil consolidation, etc.

Why is this so important? The issue revolves around the temporal variability in the system characteristics that influence so dramatically the erosion rates for a given storm event. One can think of the erosional response as being a function of the overlap of two distributions, the driving force of rainfall (in this case) and the state of the system in terms of its resistance to the driving forces (Fig. 7). Obviously, the reality is very complex, and Fig. 7 simplifies the reality to a conceptual level. For example, there is no guarantee that the resistance distribution itself is independent of the driving force distribution.

Figure 7. Schematic diagram representing the overlap of two distributions: one representing the driving force of erosion (e.g., rainfall) and the second representing the system resistance to erosion.

Figure 8. Minimum simulation times required to reach stable long-term average predictions of soil loss using the WEPP model applied to fallow conditions (Baffaut et al. FFFF).



For a given set of force and resistance distributions we can expect a specific erosion response distribution. If we look, for example, at a 50 year, return frequency storm occurring on a field that has recently been planted, the impact of that event will be a relatively large erosion response. For the same storm on the same crop rotation except offset by one year, thus occurring at a time during the rotation when the system resistance is very high, the erosion response may be very small. In this case the first field might be devastated by this storm, where in the second field there was no visible sign of erosion at all, even though both fields are under the same cropping system.

The most common attempt to deal with the questions that continuous simulation attempts to address, without actually doing long term simulation, is to run a single storm model on a distribution of individual events. The limitation is that this does not in itself address the fact that the erosional response is a complex overlap of the two dynamic distributions, rather than just the storm distribution alone. The evidence is clear that this is not sufficient for characterizing erosional response differences among land-used treatments or for producing accurate long term erosion estimates, and certainly it has never been verified that the process is effective at doing either.

With the WEPP model we also attempted to develop a technique to use three representative years of continuous simulation to obtain good estimates of long-term averages of erosion. The idea was to select representative wet, dry, and average precipitation years that, when used in the WEPP model, would result in erosion estimates that mimicked long term average trends. We found that one might be able to do this, as long as one was dealing with nearly identical systems, such as summer crops with similar planting and harvest dates. If one changed to winter crops, such as wheat for example, the process no longer worked. One had to choose a different representative three years. Relative erosion rates from the long term averages compared to the three year averages were not internally consistent between management systems, which were what we were trying to differentiate. The fact that the WEPP effort was unsuccessful does not mean that using rainfall distributions with an event model to determine long term erosion averages can't be done, but it has not been accomplished to date.

The reason for our lack of success in the abovementioned problem is evident in Fig. 8. Even with a continuous simulation model one has to deal with extreme variability in the erosion predictions. Baffaut et al (1996) conducted a simulation study using the WEPP model to determine how many years of simulation were necessary in order to obtain a stable long term erosion estimate. The predicted erosion rate was considered to be stable when it its value was within 10% of the 200 year average soil loss, and remained within that interval for all subsequent years. The results (Fig. 8) indicated a need for periods of simulation approaching 150 years in order to reach stability. In general, the simulation time tended to decrease as function of the magnitude of measured soil loss, but as one can see from the graph the relationship was a weak one. The simulations were primarily done using fallow conditions, and in that sense probably reflect conservative values since erosion magnitude for cropped conditions would tend to be less, and hence variability greater.

It is useful to mention again that WEPP model predictions of event distributions were shown to mimic distributions of measured erosion quite well (Baffaut et al. 1998). We do not believe that the long simulation times reported here are a function of the instability of the model. If anything, the model may not reflect the variability that actually could exist over decadal time periods. These types of results of model simulations leave one feeling discouraged about the possibility of constructing a method of using a single storm model in conjunction with rainfall frequency distributions to obtain accurate long-term erosion estimates or for quantifying erosion differences between land uses.

These results also leave one questioning the use of a "design storm" for erosion prediction. Land use sytems cannot be evaluated with the concept of the design storm. One could take the "worst case" scenario of the designated return storm frequency modeled at the least resistant time of the year for designing erosion control structures, for example. However, the only way to determine the probability of occurrence of that magnitude of an erosion event occurring within a given return period would be to run the continuous simulation model and determine the predicted frequency distribution curves for erosion. In other words, using the worst-case-design-storm method one will predict an erosion rate for the storm. But what is the probability that that level of erosion might occur in any given year? A slightly larger storm at a slightly different time of year, and hence theoretically more resistant condition of the system, would produce the same level of erosion. On the other hand, with a continuous simulation model one can design structures and conservation practices for a design erosion event. If one runs the simulation model for 200 years, then it is possible to pick from the event distribution data the 10 year or 20 year return frequency erosion event.

Another way of summarizing this is that with a single storm model one can plan for a certain return frequency *storm precipitation* for the system, but only in a single specific system state. With a continuous simulation model one can plan for a design, return frequency *erosion event*. Those are two very different things, and the former is not useful relative to erosion.

To this point we have discussed the limitations and the problems of natural variability in erosion and long simulations required for obtaining stable erosion values. Erosion models can, nonetheless, act as very valuable tools for a variety of purposes. A) They can help the land owner or manager choose suitable conservation practices, because they are able to assess relative effects of land use even in individual hillslope cases where the accuracy of any given prediction might be uncertain. B) They can be used to make broad scale erosion surveys in order to understand the scope of the problem over a region and to track changes in erosion over time, because if the model is predicting the mean erosion well for a given land use, then it will reflect the mean of the population of erosion values for land use treatments. C) Models can be used to regulate activities on the land for purposes of conservation compliance, because they can provide a consistent and fair evaluation system to compare agricultural fields. The model might not give the exact quantification of erosion on every field for every year, but in the long term the predictions are fair and reasonably accurate. D) They can be used to estimate long term loadings to streams and other water bodies, because as the time period increases, the accuracy increases. E) If used properly, they are useful as storm-response design tools. But in this case the storm design must be done within the context of a continuous simulation model run over a sufficiently long period of time to obtain a clear quantification of the size of the erosion event for a specified return frequency.

All of the above applications require explicitly the use of a continuous simulation model. One cannot accurately assess changes or differences in land use scenarios or conservation practices without continuous simulation. USLE and RUSLE are included in this class of models, since erosion is calculated based on time variations in the cropping and erosivity factors. However, one cannot use USLE or RUSLE for the last two applications (D and E). They cannot be used for estimating long term off-site loadings, because they do not include concentrated flow routing for off-site sediment yields, and they cannot be used for determining design storms, since they are not event models. All of the above applications also implicitly take into account the issue of variability that we talked about earlier. Simulations must be made for long time periods in order to make reliable quantitative assessment.

Where mistakes are most often made in the application of models is when we do not recognize the inability of models to accurately predict erosion at low levels, such as for events or even for a few years of erosion in cases where erosion rates are low and events infrequent. We really can't accurately *measure* erosion as a function of treatments when we try to do it over short time periods. Natural variation is huge, and our model variation is even greater. Secondly, we also are working in the realm of fantasy when we try to use single storm models to assess land use treatments or to define design erosion events. Single storm models do not have the capability to function in this capacity.

This presentation was limited to the discussion to hillslope scale erosion in a simple context of sheet and rill erosion. But even at hillslope scales the situation can be much more complex. There is enormous complexity in the many processes that take place in a real landscape, as well as complexity of surface morphology and the interactions of the morphology with processes, which our models do not attempt to take into account. Along the same line, we are also finding that the basic concepts of our process-based models appear not to function in natural areas and rangelands. The entire concept of rill and interrill erosion breaks down in these areas.

References

Baffaut, C., Nearing, M.A. and Govers, G. (1998). Statistical distributions of daily soil loss from natural runoff plots and WEPP model simulations. *Soil Sci. Soc. Am. J.* 62:756-763.

Baffaut, C., Nearing, M.A. and Nicks, A.D. (1996). Impact of climate parameters on soil erosion using CLIGEN and WEPP. *Trans. Am. Soc. of Agric. Eng.* 39:447-457.

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., Govers, G. and Norton, L.D. (1999). Variability in soil erosion data from replicated plots. *Soil Sci. Soc. Am. J.* 63:1829-1835.

Risse, L.M., Nearing, M.A., Nicks, A.D. and Laflen, J.M. (1993). Assessment of Error in the Universal Soil Loss Equation. *Soil Science Soc. of Am. J.* 57:825-833.

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