



Stichting Onderzoek Wereldvoedselvoorziening van de Vrije Universiteit

Centre for World Food Studies

**A NON-PARAMETRIC/PARAMETRIC ANALYSIS OF THE UNIVERSAL SOIL
LOSS EQUATION**

by

B.G.J.S. Sonneveld and M.A. Nearing

Contents

1.	The USLE	1
	The model	1
	Improving the USLE	2
2.	Data and methods	4
	Data	4
	Non-parametric estimation	4
	Parametric techniques	6
3.	Model robustness and modifications	6
	Robustness	6
	Improving the model fit	7
	Tenfold cross validation	8
4.	Summary and conclusions	10
	References	11

Abstract

Due to its modest data demands and transparent model structure the Universal Soil Loss Equation (USLE) remains the most popular tool for water erosion hazard assessment. However, the model has several shortcomings, two of which are likely to have prominent implications for the model results. First, the mathematical form of the USLE, the multiplication of six factors, easily leads to large errors whenever one of the input data is misspecified. Second, the USLE has a modest correlation between observed soil losses and model calculations, even with the same data that were used for its calibration. This raises questions about its mathematical model structure and the robustness of parameters that are implicitly assigned to the model. This paper, therefore, analyzes if the USLE could benefit from mathematical model transformations that, on one hand, mitigate the impact of incorrect input factors and, on the other hand, result in a better fit between model results and observed soil losses. For the analysis we revisit the original data set and consider the USLE factors as variables rather than their common interpretation as parameters. We first use both non-parametric and parametric techniques to test the robustness of the implicit parameter assignments in the USLE equation. Next, we postulate alternative mathematical forms and use parametric test statistics to evaluate parameter significance and model fit. A tenfold cross validation of the model with the best fit tests the sensitivity of the parameters for inclusion or exclusion of the data. The results show that the USLE model is not very robust, however, only slight model improvements are obtained by drastic modifications of its functional form, thereby sacrificing the simple model structure that was intended by its designers.

1. The USLE

The model

The USLE is the most widely used model for prediction of water erosion hazards and planning of soil conservation measures. It was adopted in 1958 by the Soil Conservation Service in the USA to make long-term assessments of soil losses under different cropping systems and land management practices. On the basis of a considerable experience with more than 10 000 plot years, 20 years later (Wischmeier and Smith, 1978) an updated equation was formulated which product form bears a resemblance to a Cobb-Douglas function with parameters of the value 1:

$$A = R \times K \times L \times S \times C \times P,$$

where **A** represents the soil loss, commonly expressed in tonnes.ha⁻¹.year⁻¹. **R** refers to the rainfall erosivity factor, calculated by the summation of the erosion index EI30 over the period of evaluation. EI30 is a compound function of the kinetic energy of a storm and its 30-minute maximum intensity. The latter factor is defined as the greatest average rainfall intensity experienced in any 30-minute period during a storm. It can be computed from automatic rain gauge charts, by locating the greatest amount of rainfall in any 30-minute period and then doubling this amount to get the same dimensions as normal rainfall intensity i.e. rainfall per hour. **K** is the soil erodibility factor reflecting the susceptibility of a soil type to erosion. It is expressed as the average soil loss per unit of the R factor. **L** is an index of slope length, expressed as the ratio of the expected soil loss to that observed for a field of 22.6 metres length. **S** is a slope gradient index, the ratio of the expected soil loss to that observed for a field of specified slope of 9 per cent. **C** is an index for the protective coverage of canopy and organic material in direct contact with the ground. It is measured as the ratio of soil loss from land cropped under specific conditions to the corresponding loss from tilled land under continuous fallow conditions. Finally, the protective factor **P** represents the soil conservation operations or other measures that control the erosion, such as contour farming, terraces and strip cropping. It is expressed as the ratio of soil loss with a specific support practice to the corresponding loss with up-and-downslope culture.

The simple structure of the USLE formula makes it easy to formulate transparent policy scenarios by changing the land use types (C and P factors) under given ecological conditions (R, K, L, and S factors). This, together with the low data requirements compared with physical based models, such as WEPP and EUROSEM, explains the popularity of the USLE in small-scale water erosion studies at a continental (UNEP/RIVM/ISRIC, 1996b; Knijff, et al. 2000), nationwide (Knijff et al., 1999; Schaub and Prashun, 1998; UNEP/RIVM/ISRIC, 1996a UNEP/RIVM/ISRIC, 1997; Bionnais et al., 1999), statewide (Hamlett et al. 1992), regional (Folley, 1998) and watershed level (Mellerowicz, 1994; Merzouk et al., 1998;

Kinnel, 1999; Dostal and Vrana, 1998). The USLE is also popular in (nationwide) land evaluation studies where it is linked with rule-based procedures to determine the decrease in productivity (Kassam, 1991; Struif Bontkes, 2001) or to estimate changes in nutrient balances (Smaling, 1993).

However, the USLE has some intrinsic model limitations which require attention. A non-parametric analysis of the original USLE data set (Keyzer and Sonneveld, 1998) reveals that observations are largely concentrated around the lower soil losses, where the model gives reliable estimates. However, data on higher soil losses are scarce and large errors can be expected for high rainfall erosivity in combination with steep slopes and the lower and higher K values. The results confirm earlier findings where the USLE model shows a rather poor statistical fit when used to explain annual soil losses ($R^2 = 0.57$) of the same data that were used for its calibration (Risse et al., 1993) and, moreover, consistently overestimates small soil losses and underestimates the higher ones (Nearing, 1998). Furthermore, the mathematical form of the USLE, the multiplication of six factors, leads to large errors whenever one of the factors is misspecified (Wischmeier, 1976). This raises questions about the model specification of the USLE and robustness of the implicit parameter assignments. Therefore an approach seems justified to investigate the impact of an equation transformation of the USLE model that, on one hand, could mitigate the impact of erroneous input factors and, on the other hand, result in a better fit between model results and observed soil losses.

Improving the USLE

Various attempts to improve the predictive capability of the USLE technology have been made in recent years, both along traditional and nontraditional avenues of inquiry. Along the traditional route, whereby we seek to improve the prediction capabilities of the model by focusing on better parameter estimations, the most extensive work is undoubtedly the Revised Universal Soil Loss Equation (RUSLE) (Renard et al., 1997). The changes from the USLE to the RUSLE generally fit into two categories: a) incorporation of new or better data, and b) consideration of selected erosion processes. Recent data from the western United States was used to develop a new map for the rainfall erosivity factor, R. Obviously, this improvement and other similar improvements in the RUSLE that are based on new or better data will have impact only for regional or other applications for which the new data is relevant.

The incorporation of selected erosion processes into the RUSLE model has the potential for broader prediction improvements. Some of these improvements include functions for the seasonal variability in the soil erodibility factor, K; slope length and steepness factors that are dependent on rill to interrill erosion ratios, inclusion of support practice, P, factors for subsurface drainage, rangelands, off-grade contouring, and strip-crop rotations; and the dependence of the contour P-factor on storm severity, (Renard et al., 1991,

1997). Perhaps the most extensive changes in the model are the inclusions of sub-factors in the cover-management factor, C, for the effects of prior land use, canopy cover, ground surface cover, surface roughness, and soil moisture. It is difficult to assess the improved function of the RUSLE as compared to the USLE because so many of the changes in the new model are targeted for specific applications. For example, improvements in the P factor for subsurface drainage have significance only for the specific case of drained fields, and don't affect the application of the model elsewhere. The increased prediction capability of the RUSLE might best be stated in terms of the increase in the scope of application, rather than its increased prediction accuracy for cases in which the USLE was developed. It should also be noted that the RUSLE focuses on application within the United States, and adaptation and use elsewhere have as yet been limited. In one study using data that was collected on natural runoff plots located primarily in the eastern half of the United States, the RUSLE model did not outperform the USLE in its prediction accuracy (Tiwari et al., 2000).

Also recently, two nontraditional approaches to improving the USLE have been undertaken. One involves the use of fuzzy logic-based modeling (Tran et al., 2002). In this case a data set consisting of 1700 plot-years of information from 200 individual plots at 21 sites in the United States were used. These were data that were used, in part, to develop the USLE model. RUSLE parameters estimated from a previous study (Rapp, 1994) were listed with the measured soil loss data from each plot year. The data were divided into several subsets based on ranges of some of the RUSLE parameters, and then the measured soil loss data was related to the RUSLE parameters using the fuzzy logic approach. The results indicated that the new fuzzy logic-based model provided a better fit to the data than did RUSLE. Of course, in this case the data used to develop the fuzzy logic-based model were the same as those used to evaluate the model. The same data were also used, in part, to develop the USLE that was precursor to the RUSLE. It is unclear if the relative rankings of the two models (fuzzy logic and RUSLE) would remain the same were the models applied to other, independent data. Independent data from a significant number of locations is difficult to find.

A second nontraditional approach to improving the USLE was made by Licznar and Nearing (in review). This approach involved using artificial neural networks to predict soil loss from natural runoff plots. Data from 2879 erosion events from eight locations in the United States were used. In this case, neither USLE nor RUSLE parameters were considered, nor were the results directly compared to the USLE or RUSLE results. However, the results were compared to the Water Erosion Prediction Project (WEPP) model (Flanagan and Nearing, 1995). We include a review of this work here because it used the same natural runoff plot data source as was used to develop the USLE and the fuzzy logic-based model discussed above, and it falls within the realm of essentially empirically based approaches to erosion modeling. Neural networks were created using commercially available software that related soil loss to 10 data elements: precipitation [mm], duration of precipitation [h], canopy

cover [ND], interrill cover [ND], effective hydraulic conductivity [$\text{mm}\times\text{h}^{-1}$], adjusted interrill soil erodibility K_i [$\text{kg}\times\text{s}\times\text{m}^{-4}$] as computed by the WEPP model (Flanagan and Nearing, 1995), adjusted baseline rill erodibility K_r [$\text{kg}\times\text{s}\times\text{m}^{-4}$] as computed by the WEPP model (Flanagan and Nearing, 1995), number of days since last disturbance [day], slope steepness [$\text{m}\times\text{m}^{-1}$], slope length [m]. The neural network was trained on a subset of the data, then applied to the remainder of the data set. Overall, the neural network performed as good as or better than did the WEPP model in predicting soil loss.

The approach that we follow in this paper is to revisit the original USLE data set and interpret the individual factors as variables instead of their usual interpretation of parameters. This allows us to postulate alternative mathematical forms that accommodate the USLE factors to test if we can improve the model fit. We start the analysis by testing the robustness of the original mathematical form of the USLE using non-parametric reliability estimates and parametric test results. We will also use the flexible form of the non-parametric regression to evaluate the ‘best’ possible model fit with observed soil losses and compare these with the parametric model fits. Finally, for the model with the highest fit we will test the sensitivity of the parameters for the inclusion or exclusion of the data by a tenfold cross-validation (Weiss and Kulowski, 1991). In this procedure, the data set is subdivided, at random, into 10 sets of about equal size. The model is estimated each time with 9 subsets of the data, keeping the remaining 10 per cent as an evaluation set.

This paper is organized as follows. In section 2 we introduce the data set and the non-parametric and parametric technique that are used for the analysis. Section 3 presents the results and section 4 concludes.

2. Data and methods

Data

We use a representative sample of the original data set to review relationships between observed soil loss, the model error and explanatory variables. The data set was used by Risse et al. (1993) to evaluate model efficiency and comprises 1704 observations on annual soil loss, collected from 208 natural run-off plots at 22 experimental stations in the USA, during the period 1930-1980. Each variable of the USLE equation was determined on these sites.

Non-parametric estimation

The non-parametric estimation technique that is applied in this study yields a flexible form that evaluates the ‘best’ possible fit between observed soil losses and model results. Furthermore, we use statistics on the reliability of the estimate to evaluate the robustness of

the model. The non-parametric technique is effectuated by a kernel density regression (e.g. Bierens, 1987) according to the following stochastic model

$$\hat{y}(x) = \int y(x + \varepsilon) \psi(\varepsilon) d\varepsilon,$$

where y is the dependent variable (e.g. observed soil loss), x is a vector of explanatory variables and ε denotes measurements errors in x . The function $y(x+\varepsilon)$ is the unknown (erosion) function, and the regression takes the expected value of this function. For an infinite sample of observations spread evenly over the domain of x , it would be possible to evaluate this expected value. However, with a finite sample of size S , the value of y can only be estimated and for this the kernel density regression uses the Nadaraya-Watson estimator:

$$\tilde{y}(x) = \sum_s P_s(x) y^s,$$

where y^s denote observations and x can be any point. This is a probability weighted sample mean. The probabilities are computed on the basis of the distance of x^s from the given point x , attributing higher weight to nearby points. Weights are assigned using a postulated density function (the kernel) for ε whose spread is controlled by the window size parameter θ , which standard value is determined by,

$$\theta = \left(\frac{4}{n(d+2)} \right)^{\frac{1}{d+4}},$$

with n the number of observations, and d the number of exogenous variables (Silverman, 1986). The mollifier program supposes that all the elements of ε are independently and normally distributed. The mollifier also calculates several statistics that reflect the estimates' reliability. In this exercise we use the partial derivative of the regression curve with respect to the explanatory variable x_k^t at data point x^t , as well as a measure of reliability for it. For this, it evaluates at every data point:

$$\frac{\partial \tilde{y}(x^t)}{\partial x_k^t} = \sum_s \frac{\partial P_s(x^t)}{\partial x_k^t} y^s.$$

from which we derive the probability of a wrong sign of the first derivative, the mathematical explanation of which is relegated to the Annex. A value of .5 indicates that on average the slope information is uninformative, above .5 that it has the wrong sign, and the more below .5, the more reliable the average slope.

Parametric techniques

The parametric technique used in this paper is the conventional regression techniques of Ordinary Least Squares. We use t-scores to test robustness of assumed parameter values of the USLE factors and evaluate the correlation coefficient (R^2) between the observed soil losses and model results when alternative mathematical forms of the USLE are used.

3. Model robustness and modifications

Robustness

The robustness of the functional form is tested both in a non-parametric and a parametric way. The factor values of R, K, L, S, C and P form here the independent variables and observed soil losses constitute the dependent variable.

The *non-parametric* test calculates the probability of a wrong sign of the first derivative of the individual parameters which provides a measure of significance of the associated variable with the dependent variable.

Table 1. Results of the non-parametric regression on the USLE data set

Variable	Probability of wrong sign for 1 st derivative
R-factor	0.47312
LS-factor	0.48187
K-factor	0.49746
C-factor	0.46322
P-factor	0.50082

Table 1 presents the probability of a wrong sign of the first derivative for the USLE factors. All factor values are close to 0.5 while the P-factor even exceeds this level, indicating a rather low association of the variables with soil loss.

However, there is no formal test statistic such as $a\chi^2$ to relate this statistic to the measured one and, hence, the results have to be interpreted without a formal test. Therefore, we will also analyse the robustness of the USLE form with a parametric technique that allows for a more common and formal interpretation. The approach is to interpret the mathematical formula of the USLE as a Cobb-Douglas function with power coefficients of 1. We then evaluate the significance of added parameters α_i , as:

$$\log A = \alpha_0 + \sum_{i=1}^6 (1 + \alpha_i) \log X_i ,$$

where α_0 is a constant and α_i a parameter value that belongs to the vector X_I , which represents the five USLE factors. The results of this estimation are given in Table 2.

All parameters are significant at the 1 per cent level, except for the K-factor, which is highly insignificant. Parameter values of K and LS factors are almost zero indicating that the exponents are close to 1, approximating the assumed values in the USLE model. The estimated parameter values for C, R and especially the P-factor show, however, that the assumptions are not very robust. Together with the results of the non-parametric analysis, this implies that possible model improvements can be obtained by estimating different parameter values for the individual USLE factors

Table 2. Results of the parametric regression on the USLE data set

Parameter	Estimate
Constant	0.3947
R	0.1641
K	-0.0702
LS	0.0243
C	-0.1387
P	-0.6692

Improving the model fit

We start the exercise on model improvement by applying a non-parametric regression on the USLE factors and soil losses. The flexible form of the non-parametric model provides us with a ‘theoretical’ best fit between model results and observed soil losses. The R^2 of the non-parametric regression, again between factor values of R, K, L, S, C and P as independent variables and observed soil losses as dependent variable, shows a value of 0.76, which is, for the given window size, the best achievable fit between this model and the observed data that can be obtained. However, this non-parametric model has as disadvantage that while the model derived in this manner is highly flexible and elaborate, it only reflects theoretical restrictions through the choice of variables and not through properties of the functions, precisely because the moulding of the model into a particular shape is almost exclusively driven by the data. In the context of policy modeling, in this case on decisions of land management, this is considered undesirable and a parametric approach is therefore preferred.

We therefore aim at a better model fit by testing different analytical forms in which model calculations (A) and individual factor estimations (Xb) are taken as empirical data for independent variables, while observed soil losses form the dependent variable. The selection of non-linear deviations of individual factors is based on the results of the non-parametric analysis (Keyzer and Sonneveld, 1997) and correlation coefficients is presented in Table 3.

Table 3. Mathematical forms and correlation coefficient.

Name	Functional format	R ²
Linear	$S = a(A) + b$	0.5700
Quadratic	$S = a(A)^2 + b(A) + c$	0.5773
Polynomial	$S = a(A)^3 + b(A)^2 + c(A) + d$	0.5774
Cobb-Douglas	$\ln A = a + \sum_i b_i \ln X_i$	0.6035
Exponential I	$S = \exp^{aA+b}$	0.5762
Exponential II	$S = \exp^{a \sum_i b_i \ln X_i}$	0.4387
Mitscherlich-Baule	$S = a(1 - \exp^{-\sum_i b_i X_i})$	0.5768

From table 3 we see that the Cobb-Douglas function has the best fit and its parameter estimates are reported in Table 4.

Interestingly, most parameter values are close to unity approximating the original USLE exponents, while the intercept is near zero. The P-factor is an exception with a much lower value. The R² of 0.6045 is only slightly higher than the USLE model fit (R² = 0.57).

Table 4. Parameter values of USLE/log regression

Variable	Estimate*
Intercept	0.395113
R-factor*	1.163904
K-factor*	0.929675
LS-factor*	1.024114
C-factor*	0.861252
P-factor**	0.330219

* Parameter values were at the 0.0001 significance level.

** Parameter value at the 0.001 significance level.

Tenfold cross validation

The estimated parameters of the Cobb-Douglas function are tested for their sensitivity to the inclusion or exclusion of observations by a tenfold cross-validation procedure. Figure 1 presents the estimates. All parameters retain their original sign, but vary in the deviations from their mean value.

Parameters for rainfall erosivity, soil erodibility and coverage factor exhibit minor fluctuations, and are highly significant throughout the estimation rounds. The topography

index (LS-factor) is also highly significant but its value is less stable reporting maximum deviations from its mean value of 15 per cent. The protection factor is most sensitive for the exclusion of the data and has a maximum of 100 per cent deviation from its mean value and has the lowest significance of all parameters. The R² differs only with 0.01 or 0.02 from the 0.6035 value that was obtained for the entire data set. The results are sufficient to deduce that the parameters are stable except for the protection factor which shows a high sensitivity to the inclusion or exclusion of observations.

We therefore conclude that, although the parametric form of the USLE is not robust, postulating different parametric forms only slightly increases the robustness of the model but does not result in substantial improvements of the correlation with observed soil losses and goes at the expense of the simple USLE model structure.

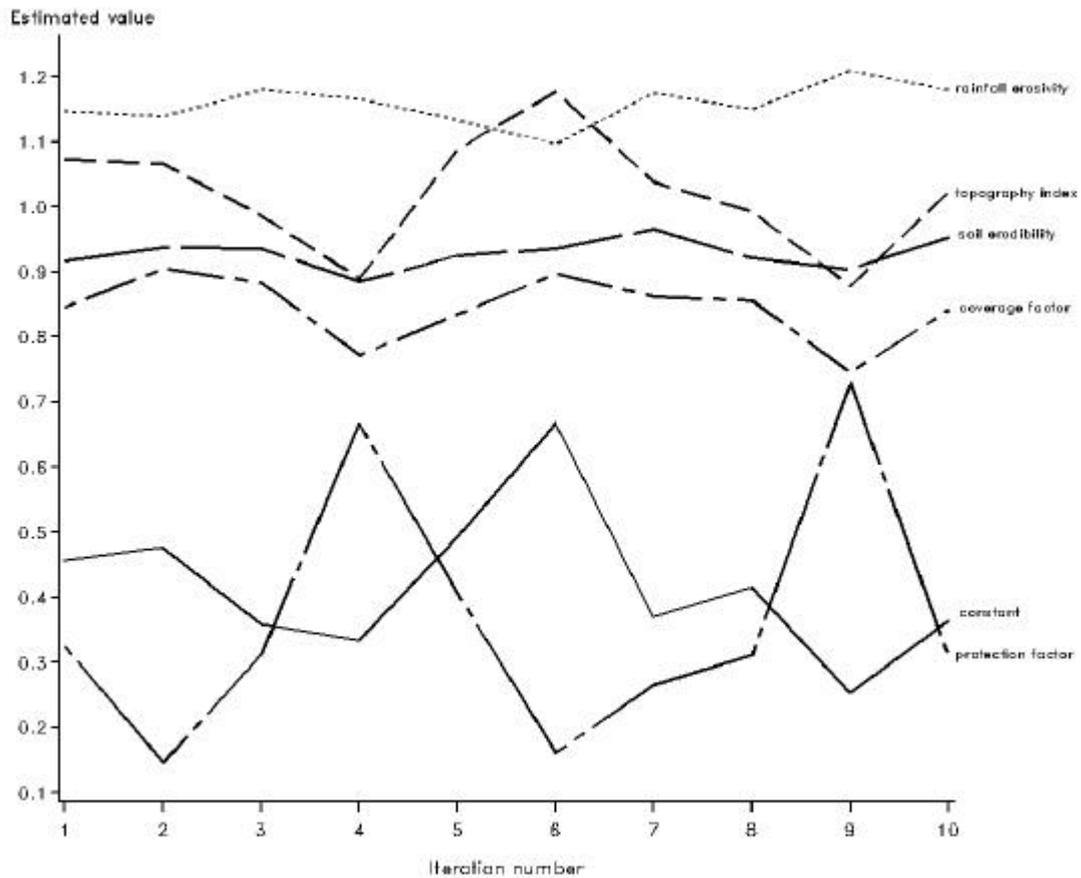


Figure 1. Tenfold cross-validation.

4. Summary and conclusions

This paper investigates if the USLE could benefit from mathematical model transformations that, on one hand, mitigate the impact of incorrect input factors and, on the other hand, result in a better fit between model results and observed soil losses. For the analysis the original USLE data set is revisited whereby we consider the individual USLE factors as variables rather than their common interpretation as parameters. We first use both non-parametric and parametric techniques to test the robustness of the implicit parameter assignments in the USLE equation. Next, we postulate alternative mathematical forms and use test statistics to evaluate parameter significance and model fit. A tenfold cross validation of the model with the best fit evaluates the sensitivity of the parameters for inclusion or exclusion of the data.

Both the non-parametric and parametric results show that the USLE formula is not very robust. However, we could only obtain a slight model improvement in robustness of parameter estimates and model fit by modifying its functional form, thereby sacrificing the simple model structure that was intended by its designers (Wischmeier, 1976).

Further research should not aim at refinements of the functional format but, for example, concentrate on the inclusion of other explanatory variables like a runoff coefficient as proposed by Kinnell (1995).

References

- Bierens, H.J. (1987) Kernel density estimations of regression functions. *Advances in Econometrics* 6, Cambridge University Press.
- Bissonais, Y Le, V. Auzet and C. Valentin (1999) Erosion in France a nationwide assessment. Paper presented at the COST 623 meeting in Brussels.
- Dostal, T. and K. Vrana (1998) Linkage of USLE and the IDRISI-GIS for estimation of sediment yield from watersheds in the Czech Republic. In: A. Klik (ed.) *Experiences with soil erosion models*, Wiener Mitteilungen no. 151, pp. 243-256.
- Flanagan, D.C. and M.A. Nearing. 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.
- Folley, A. (1998) Erosion risk assessment in Ghana using USLE: The role of GIS. In: A. Klik (ed.) *Experiences with soil erosion models*, Wiener Mitteilungen no. 151, pp. 273-286.
- Hamlett, J.M., D.A. Miller, R.L. Day, G.W. Peterson, G.M. Baumer and J. Russo. (1992) Statewide GIS-based ranking of watersheds for agricultural pollution prevention. *Journal of Soil and Water Conservation*, vol 47, no. 5, pp. 399-404.
- Kassam, A.H., H.T. van Velthuizen, A.J.B. Mitchell, G.W. Fischer and M.M. Shah (1991) Agroecological assessment for agricultural development planning. A case study of Kenya. Resources data base and land productivity. Technical Annex 2. Soil erosion and productivity. FAO/IIASA.
- Kinnell, P.I.A. (1995). The IXEA index: An index with the capacity to give more direct consideration of hydrology in the USLE modelling environment. *Journal of Soil and Water Conservation*, vol. 50, pp. 507-512.
- Kinnell, P.I.A. (1999) USLE-M and the modeling of within catchment erosion. Abstract A-0011, Abstracts 10th ISCO conference. Paper presented at the 10th ISCO conference 23–27 May. West Lafayette, Indiana, USA.
- Keyzer, M.A. and B.G.J.S. Sonneveld (1998) Using the mollifier method to characterize datasets and models: the case of the Universal Soil Loss Equation. *ITC Journal*, vol. 3-4, pp. 263-272.
- Knijff, Van der J.M., R.J.A. Jones, L.Montanarella (1999) *Soil erosion Risk Assessment in Italy*. European Soil Bureau, Space Applications Institute. European Commission. European communities.
- Knijff, Van der J.M., R.J.A. Jones, L. Montanarella (2000) *Soil erosion Risk Assessment in Europe*. European Soil Bureau, Space Applications Institute. European Commission. European communities.
- Liczner, P and M.A. Nearing. 200X. Artificial neural networks of soil erosion and runoff prediction at the plot scale. *Catena* (in review)
- Mellerowicz, K.T., Rees H.W., Chow T.L. and Ghanem I. (1994) Soil conservation planning at the watershed level using the Universal Soil Loss Equation with GIS and microcomputer technologies; a case study. *Journal of Soil and Water Conservation*, vol. 49, no. 2, pp. 194-200.
- Merzouk, A and H. Dhman (1998) Shifting land use and its implication on sediment yield in the Rif Mountains (Morocco). In: H.P. Blume, H. Eger, E. Fleischhauer, A Hebel, C. Rey and K.G. Steiner. *Towards sustainable land use*, vol. 1. Advances in Geoecology 31. Catena Verlag, Reiskirchen, Germany, pp. 333-340.
- Nearing, M.A. (1998) Why Soil Erosion Models Over-predict Small Soil Losses and Under-predict Large Soil Losses. *Catena*, vol. 32, pp.15-22.
- Rapp., J.F. 1994. Error assessment of the Revised Universal Soil Loss Equation using natural runoff plot data. M.S. Thesis. School of Renewable Natural Resources, Univ. of Arizona, Tucson, AZ.

- Renard, K.G., G.R. Foster, G.A. Weesies, D.K. McCool, and D.C. Yoder, Coordinators. 1998. Predicting soil erosion by water: A guide to conservation planning with Revised Universal Soil Loss Equation (RUSLE). U.S. Dep. Agric., Agric. Handbook 703. U.S Government Printing Office, Washington, D.C.
- Renard, K.G., G.R. Foster, G.A. Weesies, and J.P. Porter. 1991. RUSLE: Revised Universal Soil Loss Equation. *J. Soil and Water Cons.*, 46, 30-33.
- Risse, L.M., M.A. Nearing, A.D. Nicks and J.M. Laflen (1993) Error assessment in the Universal Soil loss equation. *Soil Sci. Soc. Am. J.*, vol. 57, pp. 825-833.
- Schaub, D. and V. Prashun (1998) A map on soil erosion on arable land as a planning tool for sustainable land use in Switzerland. In: H.P. Blume, H. Eger, E. Fleischhauer, A Hebel, C. Rey and K.G. Steiner. *Towards sustainable land use*, vol. 1. Advances in Geoecology 31. Catena Verlag, Reiskirchen, Germany, pp. 161-168.
- Smaling, E.M.A. (1993) *An agro-ecological Framework for integrated nutrient management*. PhD. Thesis. Agricultural University Wageningen. Wageningen.
- Struif-Bontkes, T. (2001) Agricultural prices and land degradation in Koutiala, Mali. A regional simulation model based on farmers' decision rules. In: N. Heerink, H. van Keulen and M. Kuipers (eds.) *Economic policy reforms and sustainable land use in LDCs*. Physica Verlag pp. 151-168.
- Tiwari, A.K., L.M. Risse, and M.A. Nearing. 2000. Evaluation of WEPP model and its comparison with USLE and RUSLE. *Trans. Am. Soc. of Agricultural Eng.* 43:1129-1135.
- Tran, L.T., M.A. Ridgley, M.A. Nearing, L. Duckstein, and R. Sutherlan. 2002. Using Fuzzy Logic-Based Modeling to Improve the Performance of the Revised Universal Soil Loss Equation. p. 919-923. In: D.E. Stott, R.H. Mohtar, and G.C. Steinhardt (eds.) *Sustaining the Global Farm—Selected papers from the 10th International Soil Conservation Organization Meeting, May 24-29, 1999, West Lafayette, IN*. Published by the International Soil Conservation Organization. CD-ROM available from the USDA-ARS National Soil Erosion Laboratory, West Lafayette, IN. Web site <http://topsoil.nserl.purdue.edu/nserlweb/isco99/pdf/isco99pdf.htm> (verified March 13, 2002).
- UNEP/RIVM/ISRIC (1996a) A Qualitative Assessment of Water Erosion Risk using 1:5 M SOTER Data. An application for Northern Argentina, South-East Brazil and Uruguay. Report 96/04, ISRIC, Wageningen
- UNEP/RIVM/ISRIC (1996b) Global Assessment of Land Vulnerability to Water Erosion on a $\frac{1}{2}^\circ \times \frac{1}{2}^\circ$ grid. Report 96/08, ISRIC, Wageningen.
- UNEP/RIVM/ISRIC (1997) The Impact of Land Degradation on Food Productivity. Case studies of Uruguay, Argentina and Kenya. Report 97/01, ISRIC, Wageningen, 1997.
- Wischmeier, W.H. (1976) Use and misuse of the universal soil loss equation. *Journal of Soil and Water Conservation*, vol. 31, pp 5-9.
- Wischmeier, W.H. and D.D. Smith (1978) *Predicting rainfall erosion losses*. Agricultural Handbook 537, USDA, Washington DC.

Annex

Statistical reliability of mollifier estimates

The mollifier program assesses the partial derivative of the regression curve as well as a measure of its reliability. For this, it calculates the first partial derivative of

$$y_{\theta}^s(x) = \sum_s y^s P_{\theta}^s(x) \quad (1)$$

to x_k at point x , where k represents an explanatory variable, at all data points x^t .

$$\frac{\partial \tilde{y}(x^t)}{\partial x_k} = \sum_s \frac{\partial P_{\theta}^s(x^t)}{\partial x_k} y^s. \quad (2)$$

Since $\sum_s \frac{\partial P_{\theta}^s(x^t)}{\partial x_k} = 0$ we can write

$$\frac{\partial \tilde{y}(x^t)}{\partial x_k} = \sum_s \frac{\partial P_{\theta}^s(x^t)}{\partial x_k} (y^s - y^t), \quad (3)$$

where y^t refers to the t^{th} observation. As by definition, $\frac{\partial P_{\theta}^s(x^t)}{\partial x_k} = P_{\theta}^s \frac{\partial \log P_{\theta}^s(x^t)}{\partial x_k}$, it follows that

$$\frac{\partial \tilde{y}(x^t)}{\partial x_k} = \sum_s P_{\theta}^s(x^t) \left[\frac{\partial \log P_{\theta}^s(x^t)}{\partial x_k} (y^s - y^t) \right]. \quad (4)$$

Let us now rewrite and interpret the term in square brackets.

$$\frac{\partial \log P_{\theta}^s(x^t)}{\partial x_k} = \frac{\partial \log \psi_s(x^t)}{\partial x_k} - \sum_{h=1}^s P_{\theta}^h(x^t) \frac{\partial \log \psi_h(x^t)}{\partial x_k} \quad (5)$$

Now for a density $\psi_s(x^t) = \psi\left(\frac{x^s - x^t}{\theta}\right)$ where ψ is a normal joint density with diagonal variance matrix and variance σ_k^2 around x^t it follows that

$$\frac{\partial \log \psi_s(x^t)}{\partial x_k} = \frac{x_k^s - x_k^t}{\sigma_k^2}. \quad (6)$$

Hence the term in square brackets can be rewritten as

$$\frac{\partial \tilde{y}(x^t)}{\partial x_k} = \sum_s P_\theta^s(x^t) [\xi_k^s \delta_k^s], \quad (7)$$

where $\xi_k^s = \frac{x_k^s - x_k^t}{\sigma_k^2} - \sum_h P_\theta^h(x^t) \frac{(x_k^h - x_k^t)}{\sigma_k^2}$ and $\delta_k^s = (y^s - y^t)$.

In other words, the term in square brackets is the contribution of observation s to the slope.

For given x^t this enables us to define the probability of a positive sign for the slope as

$$P_k^+(x^t) = \sum_s P_\theta^s(x^t) \mathbb{1}_{\xi_k^s \delta_k^s \geq 0}. \quad (8)$$

Hence the probability of a wrong sign can be calculated as

$$P_k^\#(x^t) = P_k^+(x^t), \quad \text{if } \frac{\partial \tilde{y}(x^t)}{\partial x_k} < 0, \text{ and} \quad (9)$$

$$1 - P_k^+(x^t), \quad \text{if } \frac{\partial \tilde{y}(x^t)}{\partial x_k} \geq 0.$$

The Centre for World Food Studies (Dutch acronym SOW-VU) is a research institute related to the Department of Economics and Econometrics of the Vrije Universiteit Amsterdam. It was established in 1977 and engages in quantitative analyses to support national and international policy formulation in the areas of food, agriculture and development cooperation.

SOW-VU's research is directed towards the theoretical and empirical assessment of the mechanisms which determine food production, food consumption and nutritional status. Its main activities concern the design and application of regional and national models which put special emphasis on the food and agricultural sector. An analysis of the behaviour and options of socio-economic groups, including their response to price and investment policies and to externally induced changes, can contribute to the evaluation of alternative development strategies.

SOW-VU emphasizes the need to collaborate with local researchers and policy makers and to increase their planning capacity.

SOW-VU's research record consists of a series of staff working papers (for mainly internal use), research memoranda (refereed) and research reports (refereed, prepared through team work).

Centre for World Food Studies
SOW-VU
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands

Telephone (31) 20 - 44 49321
Telefax (31) 20 - 44 49325
Email pm@sow.econ.vu.nl
www <http://www.sow.econ.vu.nl/>