This paper was peer-reviewed for scientific content.
Pages 919-923. In: D.E. Stott, R.H. Mohtar and G.C. Steinhardt (eds). 2001. Sustaining the Global Farm. Selected papers from the 10th International Soil Conservation Organization Meeting held May 24-29, 1999 at Purdue University and the USDA-ARS National Soil Erosion Research Laboratory.

Using Fuzzy Logic-Based Modeling to Improve the Performance of the Revised Universal Soil Loss Equation

L.T. Tran, M.A. Ridgley, M.A. Nearing, L. Duckstein and R. Sutherland*

ABSTRACT

This paper reports the application of fuzzy logic-based modeling (FLBM) to improve the performance of the Revised Universal Soil Loss Equation (RUSLE). The FLBM approach was to make the RUSLE's structure more flexible in describing the relationship between soil erosion and RUSLE factors and in dealing with data and model uncertainties while not requiring any further information. The approach used in this study consists of two techniques: multiobjective fuzzy regression (MOFR) and fuzzy rule-based modeling (FRBM). First, MOFR was used to derive the relationship between soil loss and a combination of RUSLE factors. These MOFR models were in turn linked together in a FRBM framework. Then these fuzzy rules were applied to adjust the RUSLE prediction corresponding to each combination of RUSLE factors.

The Nash-Sutcliffe model efficiency of the fuzzy model on a yearly basis was 0.70 while RUSLE's was 0.58. On an average annual basis, the efficiency was 0.90 and 0.72 for the fuzzy model and RUSLE, respectively. With several good characteristics, the FLBM approach can be used to improve the performance of RUSLE with little effort and modification to the existing RUSLE model.

INTRODUCTION

This paper reports the application of fuzzy logic-based modeling (FLBM) to improve the performance of the Revised Universal Soil Loss Equation (RUSLE). For the purpose of conservation planning, the prediction accuracy of RUSLE is very important in making sound decisions on how soil should best be protected from erosion. However, an analysis of over 1700 plot-years of data, taken from more than 200 plots at 21 sites in the U.S., showed that soil erosion was not adequately described merely by the multiplication of six RUSLE factors in all cases. For instance, data indicated that the relationship between rainfall erosivity factor (*EI*) and soil loss, when other RUSLE factors were held constant, was not always linear. The aim of the FLBM approach was to make the RUSLE's structure more flexible in describing the relationship between soil erosion and RUSLE factors and in dealing with data and model uncertainties while not requiring any further data. The paper is organized as follows – RUSLE and its limitations are discussed in the next section. Data and methodology are presented in the section 3. Section 4 is devoted to results and discussion followed by conclusion in the last section.

The Revised Universal Soil Loss Equation

RUSLE is an empirical equation derived from a large amount of field data and has been widely used as an erosion prediction and conservation planning tool in the U.S. as well as worldwide. The model computes soil erosion using values representing major factors influencing erosion, including climate erosivity, soil erodibility, topography, and land use and management. Keeping the same format of the Universal Soil Loss Equation (USLE), RUSLE is expressed as follows:

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

Where A is the mean annual soil loss (t ha^{-1}); R is the rainfall and runoff erosivity (MJ mm ha^{-1} h^{-1}); K is the soil erodibility factor (t ha h (ha MJ mm)$^{-1}$); LS is the combined dimensionless slope length and slope steepness factor; C is the dimensionless cover-management factor, and P is the dimensionless supporting practices factor.

The rainfall and runoff factor (R) is the average annual total of the storm EI values, which equal the total storm energy (E) times the maximum 30-min intensity (I_{30}). The relation between soil losses to the EI parameter was assumed to be linear (Renard et al., 1997). Compared to USLE, RUSLE included more precise values of R for the western half of the United States; and more corrections, more refined smoothing, and the filling of data gaps for the eastern United States (Renard et al., 1997).

Generally K in RUSLE was computed in a similar manner as in USLE. The difference of K between the two models is an adjustment added in RUSLE to take into account seasonal changes, such as freezing and thawing, soil moisture and soil consolidation (Renard et al., 1994a). On the other hand, compared to USLE, LS in the RUSLE was refined by assigning new equations based on the ratio of rill to interrill erosion and accommodating complex slopes (Renard et al., 1994b).

The C factor in USLE was computed based on cropping sequence, surface residue, surface roughness, and canopy

*L.T. Tran, M.A. Ridgley, and R. Sutherland, University of Hawaii, Geography Department, 2424 Maile Way, Honolulu, HI 96822, USA; M.A. Nearing, National Soil Erosion Laboratory, USDA-ARS, West Lafayette, IN 47907-1196, USA; L. Duckstein, Ecole Nationale du Genie Rural, des Eaux et des Forets, 19 avenue du Maine, 75732 Paris Cedex 15, France. *Corresponding author: LTT1@psu.edu. Current address: Center for Integrated Regional Assessment, Penn. State University; 2217 Earth & Engineering Science Building, University Park, PA 16802, USA.

cover, weighted by the percentage of erosive rainfall during the six crop stages. Compared to USLE, RUSLE included some more factors in determining *C*, such as prior land use, canopy cover, surface cover, surface roughness, and soil moisture. However, the key difference in computing *C* between the two models is the use of time-varying computation with a 15-day interval in RUSLE (Renard et al., 1997), which was considered a major improvement in estimating soil loss (Renard et al., 1994b).

P in RUSLE, which was also considered more advanced than those in USLE, was computed with a combination of empirical and process-based erosion technology, based on hydrologic soil group, slope, row grade, ridge height, and the 10-year single storm erosion index value (Renard et al., 1997). A complete description of all RUSLE factors can be found in USDA Agricultural Handbook Number 703 (Renard et al., 1997).

Risse et al. (1993) carried out a comprehensive analysis to assess the error associated with the USLE. Results from this study showed that the overall Nash-Sutcliffe model efficiency was 0.75 on an average annual basis and 0.58 when compared on a yearly basis. These values are considered reasonable as the same data set showed an annual erosion variability of ±35% between replicated plots (Yoder et al., 1998). However, looking at more detail, USLE over predicted soil loss on plots at low erosion rates while underpredicted for plots at high erosion rates.

Relying heavily on Risse et al.'s work and using data from the same sites and the same periods, Rapp (1994) found a similar result for RUSLE with a model efficiency of 0.73 on an average annual basis and 0.58 on a yearly basis. Similar to USLE, RUSLE tended to overpredict on plots with low erosion rates and under predict on plots with high erosion rates. For this similarity in performance of USLE and RUSLE, Yoder et al. (1998) explained that RUSLE was developed from the basic USLE for the purpose of extending its application rather than increasing its predictive accuracy for normal cropping situations. As a consequence, they would be expected to provide a similar degree of accuracy.

With more data and many improvements, RUSLE was considered scientifically superior to USLE (Renard et al., 1994b). Consequently, it was expected that RUSLE would perform better than USLE. However, this was not the case as presented above. Besides the issue of data limitation, the relatively moderate performance of RUSLE can be attributed to several theoretical problems. First of all, to make the model more practicable, runoff, an important factor to which soil loss is closely related, was not explicitly dealt with but incorporated within the *R* factor (Morgan, 1995). On the other hand, although soil losses are directly proportional to the *EI* parameter, this relation is generally nonlinear. Scattergram of soil losses versus *EI* values at barley sites in

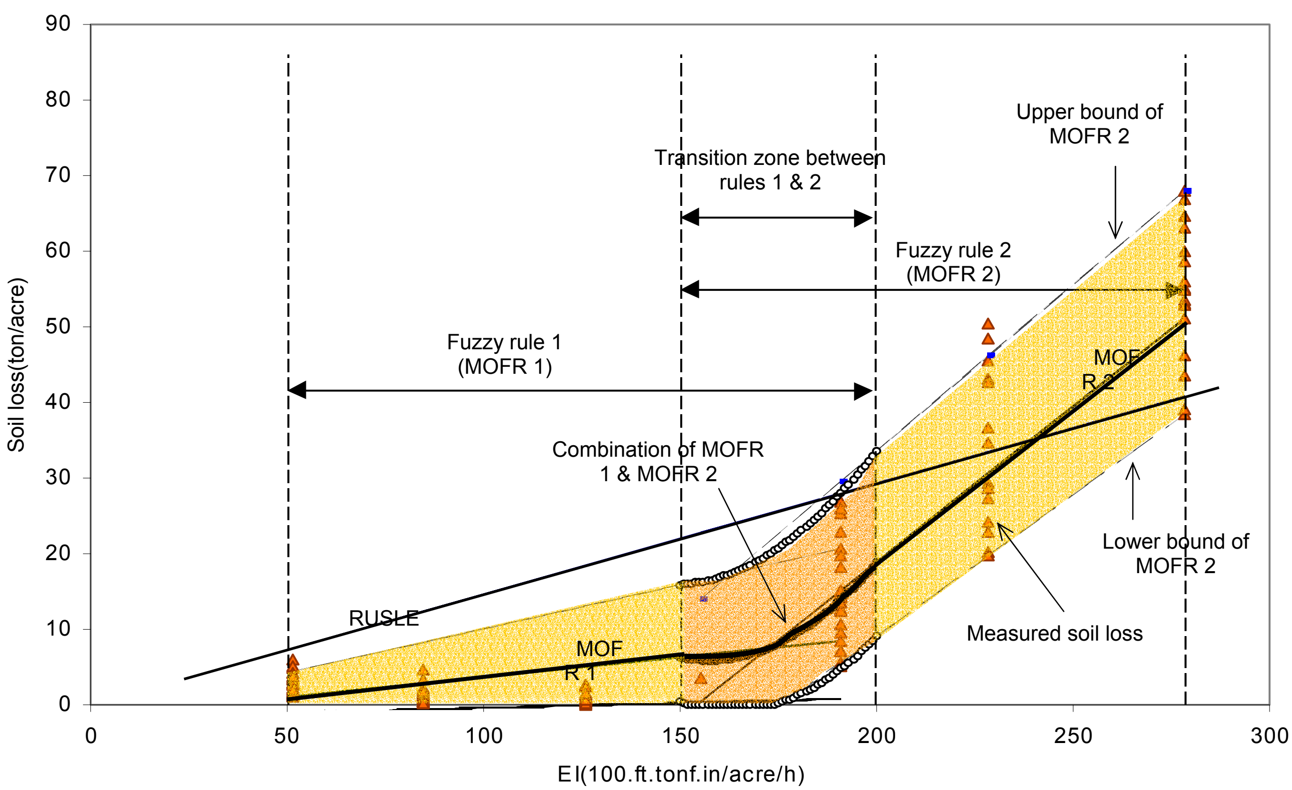

Figure 1. Measured and RUSLE-predicted soil loss versus EI values for barley sites.

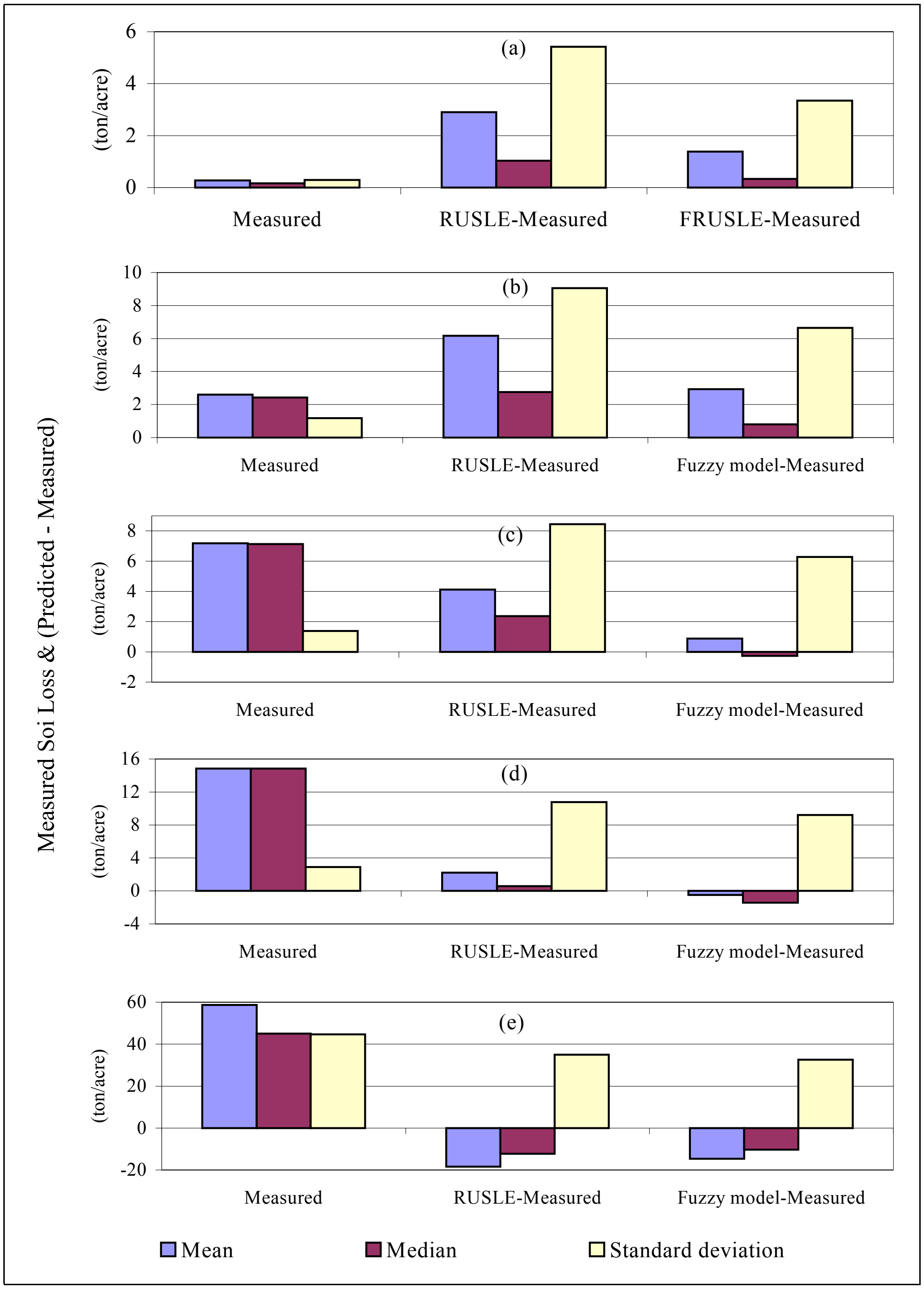

Figure 2. Expected accuracy of the RUSLE and the fuzzy model for different categories of measured soil loss (ton/acre) on a yearly basis; (a), <1; (b), 1-5; (c), 5-10; (d), 10-20; and (e), >20.

Figure 1 illustrates this point. Perhaps the linear assumption in the RUSLE is only applicable within a certain range but not the whole scope of all *EI* values. With the structure of a simple multiplication of several factors, RUSLE, however, cannot accommodate such a nonlinear relationship.

Another theoretical problem is that the interdependence between different factors is quite considerable. For instance, the relationships between *L* and *S* in USLE were derived from soil-loss measurement from mostly medium-textured, poorly aggregated surface soils (Wischmeier and Smith, 1978). These relationships, in turn, were used to determine *K* values. As a consequence, errors and shortcomings from these relationships would carry over into *K* (Renard et al., 1997). There is also a similar problem for the rainfall erosivity factor *R* (Wischmeier and Smith, 1978).

Furthermore, interactions between *K* and *C* caused problems in delineating values of these two factors (Renard et al., 1997).

DATA AND METHODOLOGY

Data

The data set used in this study was obtained from Rapp (1994). It was originally supplied by the USDA-ARS Southwestern Watershed Research Center, containing year-by-year information of over 1700 plot-years from more than 200 individual plots at 21 sites. It included previously determined USLE factor values, crop types and yields, and rotation sequences for each year and the plot dimensions. The number of plots and duration were different from site to site. The average is nearly eight years per plot for all 21 sites. Many plots had duplicated measurements to take into account the natural variability of soil loss. With the original data set, Rapp (1994) used the RUSLE computer program to identify the RUSLE factor values by providing all necessary information to the program. This data set is considered reliable since this work was supervised by Ken Renard, the team leader of the RUSLE project. The complete data set as well as a list of sites with individual plot conditions can be found in Rapp (1994).

Methodology

As analyzed in the previous section, the RUSLE had several problems related to model structure and parameters. To make the RUSLE's structure more flexible in describing the relationship between soil erosion and the RUSLE factors, and in dealing with uncertainties of parameters, while not requiring any additional information, a FLBM approach was applied. This approach includes two techniques: multi-objective fuzzy regression (MOFR) and fuzzy rule-based modeling (FRBM).

A fuzzy rule-based model comprises several single fuzzy rules. Each fuzzy rule generally consists of a set of fuzzy set(s) as argument(s) A_k and a consequence B also in the form of a fuzzy set such that

If A_1 and A_2 and..... and A_K then B

A detailed technical discussion of FRBM can be found in Bárdossy and Duckstein (1995). In this study, each fuzzy rule was derived by the means of MOFR (discussed later) to describe the relationship between soil loss and the *EI* factor within a certain range of other of RUSLE factors. There are different methods for combining fuzzy rule consequences. The method of additive combination of fuzzy rule responses was used in this study (Bárdossy and Duckstein, 1995). Often a consequence from applying a fuzzy rule system is a fuzzy set. The task of transforming a fuzzy consequence into a crisp number is called defuzzification. For this study, the maximum-weighted sum defuzzification method developed by Tran (1999) was utilized.

MOFR is a fuzzy regression model developed by Tran (1999) and Tran and Duckstein (accepted), which is capable of combining central tendency and possibilistic properties of statistical and fuzzy regression, respectively. MOFR overcomes several shortcomings of fuzzy and statistical regression approaches (e.g., sensitivity to data outliers of fuzzy regression, difficulties of verifying distribution assumptions, insufficient and/or inaccurate input and/or output data, vagueness of the relationship between input and output variables in statistical regression). Furthermore, MOFR is robust with respect to y-direction outlier (often referred to simply as outlier, as distinct from the x-direction outlier often referred to as the leverage point). Hence, the model can be used when only few data are available, while this is not the case for least-squares regression. This feature is essential for MOFR to be applied to a small subset of RUSLE data. A technical description of this MOFR can be found in Tran (1999), and Tran and Duckstein (accepted).

In fact the FLBM approach used in this study can be viewed as a fuzzy piecewise linear regression model (FPLRM), in which MOFR was applied to derive the linear equation for each segment of data and FRBM was used to link those segments together. However, in contrast with conventional piecewise regression models, FPLRM does not have a single joint point but a nonlinear curve in the transition zone between two consecutive segments (Figure 1).

To apply the fuzzy rule-based modeling approach to RUSLE, the following steps were carried out:

1. Divide the data set into several subsets of certain RUSLE parameter ranges;
2. Apply MOFR to derive the relationship between soil loss and *EI* for each certain range of other RUSLE factors; and
3. Use FRBM to link these MOFR models and apply the fuzzy rule set to compute predicted soil loss corresponding to a combination of RUSLE factors.

A detailed description of these steps can be found in Tran (1999). Figure 1 illustrates the fuzzy logic-based model applied for barley data.

Predictions from the RUSLE and the fuzzy model were compared with measured data in several ways, including r^2 and the model efficiency defined by Nash and Sutcliffe (1970), which is calculated as follows:

$$R^2 = 1 - \frac{[\sum_{i=1}^{N} (Q_{mi} - Q_{ci})^2]}{[\sum_{i=1}^{N} (Q_{mi} - \bar{Q}_m)^2]}$$

where Q_{mi} and Q_{ci} are the measured and computed values, respectively, of event i, and $\bar{Q}_m$ is the mean of measured values.

RESULTS AND DISCUSSION

Results showed that the fuzzy model performed better than RUSLE. The Nash-Sutcliffe model efficiency of the fuzzy model on a yearly basis was 0.70 while RUSLE's was 0.58 (Table 1). On an average annual basis, the efficiency was 0.90 and 0.72 for the fuzzy model and RUSLE, respectively, an improvement of >25% with respect to the performance of RUSLE. Similar picture can be seen if r^2 is used to compare the performance of RUSLE and the fuzzy model (Table 1). Furthermore, the problem of over prediction at low soil loss rates and under prediction at high soil loss rates was reduced with the fuzzy model. Fig. 2 shows that the over prediction at soil loss rates <10 ton ac^{-1} was decreased significantly. On the other hand, the fuzzy model lessened considerably the over prediction at soil loss rates >20 ton ac^{-1}. It should be mentioned that over prediction at low soil loss rates is less critical than under prediction at high soil loss rates.

The FLBM approach applied to RUSLE did not only make the structure of the model more flexible and more realistic in describing the relationship between soil loss and rainstorm parameter, but it also overcame the problems of uncertainty in the RUSLE parameters. For instance, due to

Table 1: model on a yearly basis and on an average annual basis.

Parameter	Subset	n	RUSLE†	RUSLE‡	fuzzy model
				On a yearly basis	
Nash & Sutcliffe model efficiency	Calibration	957	-	0.612	0.666
	Validation	745	-	0.576	0.753
	Whole	1702	0.586	0.599	0.698
r^2	Calibration	957	-	0.617	0.674
	Validation	745	-	0.576	0.753
	Whole	1702	0.604	0.603	0.704
				On an average annual basis	
Nash & Sutcliffe model efficiency		203	0.719	0.721	0.896
r^2		203	0.736	0.740	0.905

† No recalibration was made for the RUSLE.
‡ Recalibration was made for the RUSLE using the calibration subset (957 data points).

interdependency between variables, values of a particular RUSLE parameter used in combination with other parameters within certain ranges might be incorrect. Through multiplication of all factors and nothing else to compute soil loss, this kind of error directly affects the result of the RUSLE. In contrast, the fuzzy model took into account the issue of interdependency and provided the best fit between soil loss and rainstorm energy.

In addition to central values, the fuzzy model provided lower and upper bounds on the predicted range of soil loss (Fig. 1). These bounds are valuable information for both scientific understanding and for management decisions. For instance, the level of uncertainty associated with a given range of RUSLE factors can be evaluated from these predicted ranges of soil loss.

In terms of modeling, FRBM makes the tasks of model development and updating quite easy. It is because each fuzzy rule can be developed or updated independently. This feature is not available in conventional modeling techniques, as the work must be done over for the entire model.

CONCLUSION

This analysis showed that the FLBM approach has several good features. For instance, the approach is quite simple, no other data outside RUSLE are needed, and the main structure of RUSLE is maintained. Moreover, the performance of RUSLE was improved significantly with the use of this approach. Hence it would be worth a try to employ the whole data set used in RUSLE to develop the fuzzy RUSLE model.

REFERENCES

Bárdossy, A. and L. Duckstein. 1995. Fuzzy rule-based modeling with applications to geophysical, biological and engineering systems. CRC Press, Boca Raton.

Morgan, R.P.C. 1995. Soil erosion and conservation (2nd Edition). Longman, England.

Nash, J.E. and J.E. Sutcliffe. 1970. River flow forecasting through conceptual model. Part 1 – A discussion of principles. Journal of Hydrology (Amsterdam) 10: 282-290.

Rapp, J.F. 1994. Error assessment of the Revised Universal Soil Loss Equation using natural runoff plot data. Master Thesis, School of Renewable Natural Resources, The University of Arizona.

Renard, K.G., G.R. Foster, D.C. Yoder and D.K. McCool. 1994a. RUSLE revisited: Status, questions, answers, and the future. Journal of Soil and Water Conservation 49 (3): 213-220.

Renard, K.G., J.M. Laflen, G.R. Foster and D.K. McCool. 1994b. The Revised Universal Soil Loss Equation. p. 105-124. *In* R. Lal (ed.) Soil Erosion, Research and Methods (2nd Edition). Soil and Water Conservation Society, Ankeny, Iowa.

Renard, K.G., G.R. Foster, G.A. Weesies, D.K. McCool and D.C. Yoder. 1997. Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equation *(RUSLE)*. USDA Agricultural Handbook, No. 703. U.S. Gov. Print. Office, Washington, DC.

Risse, L.M., M.A. Nearing, A.D. Nicks and J.M. Laflen. 1993. Error assessment in the Universal Soil loss Equation. Soil science Society of America Journal 57: 825-833.

Tran, L. and L. Duckstein. Accepted. Multi-objective fuzzy regression with central tendency and possibilistic properties. Fuzzy Sets and Systems.

Tran, L. 1999. Multi-objective Fuzzy Regression and its Application to Erosion Prediction. Ph.D. Dissertation, Department of Geography, The University of Hawaii.

Wischmeier, W.H. and D.D. Smith. 1978. Predicting rainfall erosion losses - A guide to conservation planning. USDA Agricultural Handbook No. 537. U.S. Gov. Print. Office, Washington, DC.

Yoder, D.C., G.R. Foster, G.A. Weesies, K.G. Renard, D.K. McCool and J.B. Lown. 1998. Evaluation of the RUSLE soil erosion model. Proceedings of the 1998 ASAE Annual International Meeting, Orlando, Florida, July 11-16, 1998.