A GIS-based Soil Erosion Risk Map for New Mexico

Bulut, G.G.[1]; Cal, M.P.[2]; Richardson, C.P.[3]; Gallegos, J.B.[4]

Abstract: A soil erosion risk map was developed for the State of New Mexico using *ArcGIS Desktop 9.3* software and its extensions with application of *FuzzyCell*, a fuzzy logic-based Geographic Information System (GIS) tool. A 2-yr, 6-hr precipitation depth, slope gradient, Normalized Difference Vegetation Index (*NDVI*), and soil erodibility were used as surrogate metrics for the respective index parameters (*R* factor, *LS* factor, *C* factor, and *K* factor) of the *Universal Soil Loss Equation* (*USLE*). A second iteration utilized a 1-dimensional length slope (*LS* factor) in lieu of slope gradient. The fuzzy generated risk maps were virtually identical statewide. Risk varied between 0.05 and 0.87 on a color scale from 0 to 1, providing a visual perspective of potentially problematic areas. This research was supported by the New Mexico Department of Transportation (*NMDOT*) in an effort to better understand design of drainage structures with respect to hydraulic capacity under sediment load from upland soil erosion. Areas of concern identified by the *NMDOT* relative to culvert deposition and clogging were generally associated with regions of medium to high risk.

***Keywords**:* Erosion risk; Geographic Information Systems; fuzzy logic.

[1]Gaye G. Bulut, Graduate Research Assistant, gayegul@nmt.edu, Department of Civil and Environmental Engineering, New Mexico Tech, Ridge Park Condo #315, 505 Oppenheimer Dr. Los Alamos, NM, 87544, PH 575-418-9175

[2] Mark P. Cal, P.E., BCEE, Chair and Professor, mcal@nmt.edu, Department of Civil and Environmental Engineering, New Mexico Tech, 801 Leroy Place Socorro, NM, 87801

[3] Clint P. Richardson, P.E., BCEE, Professor, h2odoc@nmt.edu, Department of Civil and Environmental Engineering, New Mexico Tech 801 Leroy Place Socorro, NM, 87801

[4]Jose B. Gallegos, Graduate Research Assistant, joeyg@nmt.edu, Department of Civil and Environmental Engineering, New Mexico Tech, 801 Leroy Place Socorro, NM, 87801

Introduction

Wischmeier (1959) developed the *Universal Soil Loss Equation* (*USLE*) for soil erosion estimations. The *USLE* estimates annual soil loss and is based on five empirical input variables: rainfall erosivity, soil erodibility, length slope, land use cropping factor, and erosion prevention practice factor. The objective of this study was to generate a potential soil erosion risk map for the State of New Mexico using the *USLE* model as a template. The methodology presented here was originally conceived as part of a New Mexico Department of Transportation (*NMDOT*) research project to address culvert deposition and clogging resulting from upland soil erosion and sediment transport.

Soil Erosion via the *Universal Soil Loss Equation* (*USLE*)

The *USLE* is an index method. The idea is formulated as:

$$A = K \times L \times S \times C \times R \times P \qquad \text{Eq. 1}$$

where A is average annual soil loss in metric ton/ha/yr, K is soil erodibility factor (metric ton·ha·hr/ha·MJ·mm), L is slope length factor (dimensionless), S is slope steepness factor (dimensionless), C is cover-management factor (dimensionless), R is rainfall erosivity factor (MJ mm/ha·hr·yr), and P is soil erosion prevention practice factor (dimensionless). The product of L and S is known as the length slope factor (LS).

For this study four attributes were selected to represent the respective *USLE* parameters: soil erodibility (K), slope gradient (θ), vegetative coverage ($NDVI$), and average precipitation ($_2P_6$). The soil erosion prevention practice factor was not considered. A composite 1-dimensional length slope factor (LS) was also considered in lieu of slope gradient. These variables are discussed subsequently.

Fuzzy Logic

The *USLE* model or a facsimile thereof can be integrated into *GIS* where the data stored and processed is a representation of an abstract form of the real world. While the real world consists of fuzziness, conventional *GIS* software works using crisp logic. In other words, there are shades of gray in the physical world (fuzzy logic), whereas there is only black and white in the traditional *GIS* environment (crisp logic). Due to the continuous nature of the landscape, *GIS* software cannot model the physical world in an exact way without some loss in the detail (Yanar and Akyürek 2006).

Fuzzy means vagueness and it stands for any uncertain, imprecise, or ambiguous real world knowledge. Since most systems are designed to deal with classic set data, which denotes a belonging to a set by 1 and not belonging to a set by 0; they are unable to reflect fuzzy information, which falls between 1 and 0. Fuzzy architecture has degrees of membership for a better definition of vagueness resulting from the human thinking process.

Fuzzy logic was derived from the fuzzy set theory (Zadeh 1965) and is able to define linguistic forms of the memberships such as 'very', 'quite' and 'slightly'. Fuzzy logic is broadly recognized as a tool with an ability to compute with words for modeling qualitative human thought process in the analysis of complex systems and decisions.

Fuzzy Logic, Soil Erosion, and the *USLE*

Yanar and Akyürek (2006) introduced fuzzy set theory into *GIS* that allowed users to obtain more flexibility and capability to effectively handle and process imprecise information about the real world. Using fuzzy logic, as a linguistic approach, they developed a fuzzy rule-based software system called *FuzzyCell*.

Fuzzy rules, also known as *IF-THEN* rules, include two parts: the antecedent part of the rule, the part starting with *IF* up to *THEN*, and a consequent part, the part starting with *THEN* up to the end (Tayfur *et al.* 2003). For instance, as applied herein, *IF* the precipitation depth is high; the slope gradient is steep; the soil is compact and resistive to erosion; and vegetation is sparse, *THEN* the potential erosion risk is high.

In addition to fuzzy rules, membership functions must be assigned to every element of the universe and a degree of membership specified to a fuzzy set. This membership value must be between 0 (no membership) and 1 (full membership), indicating to which degree an element belongs to the fuzzy set. Common memberships include triangular, trapezoidal, Guassian, and sigmoidal.

Software

The *GIS*-based software package used throughout this study is the Environmental Systems Research Institute's (ESRI) *ArcGIS Desktop 9.3* with primarily *ArcToolbox*. This *ArcGIS* application performs data-based manipulations such as spatial analysis, raster calculations, geoprocessing, and data merging. Another program used for this study is a fuzzy inference system that is implemented into the *ArcGIS*, known as *FuzzyCell* (Yanar and Akyürek 2006). *FuzzyCell* provides for eight possible membership functions, inference methods for rule aggregation, operators for set operations, and defuzzification methods.

Datasets

Digital Elevation Model (*DEM*): A *DEM* is a common method of representing a continuous surface in a raster form. *DEMs* may have varying resolutions resulting from the distance between sampling points, which can be considered as the accuracy of the data. As provided by the U.S. Geological Survey (*USGS*) these rasters come in 3-m, 10-m, and 30-m resolutions (http://ned.usgs.gov/ned). For this macro-scale study, a 30-m *DEM* data was used to quantify the topographical characteristics of New Mexico. The fine resolution 3-m dataset is currently not available for New Mexico and a 10-m dataset would require 9 times longer computer processing and manipulation time. Based on this dataset, New Mexico has a slope gradient that ranges between 0 and 77 degrees.

Normalized Difference Vegetation Index (*NDVI*): The *NDVI* is a standardized index that displays greenness. The differential in the reflections of different land coverage types under two spectral bands, red and infrared (*IR*), is monitored and used to produce an *NDVI* dataset. The *NDVI* calculations are based on the fact that green plants absorb radiation in the visible range and they reflect radiation in the near *IR* range. The *NDVI* values vary between -1 and 1, where -1 represents clouds, water, and snow and 0 represents rock and bare soil. The *NDVI* process in the *GIS* environment creates a one-band, 8-bit image scaling the output values from -1 to 1 to 0 to 255. The *NDVI* raster dataset used herein represents the average maximum annual *NDVI* values from 1995-2009. The *NDVI* values for New Mexico ranged between 109 and 181.

Precipitation: A *GIS*-compatible dataset of *rainfall erosivity* (*R*) is not currently available for New Mexico. However, Wischmeier and Smith (1978) suggested that the 2-yr, 6-hr precipitation ($_2P_6$) shows reasonable accuracy to *rainfall erosivity* for the western states. This precipitation metric was obtained through the Hydrometeorological Design Studies Center (*HDSC*), a division of the National Oceanic and Atmospheric Administration (NOAA)'s National Weather Service (http://www.ncdc.noaa.gov). New Mexico is located in a semi-arid region with a minimum of 0.77-inches and maximum of 2.19-inches of $_2P_6$ precipitation, averaging at 1.38-inches.

Soil Erodibility: The Soil Survey Geographic (*SSURGO*) and State Soil Geographic (*STATSGO*) are two soil property databases produced by the Natural Resources Conservation Service (*NRCS*) (http://soils.usda.gov/survey/geography/statsgo). Since the *SSURGO* data for New Mexico is not complete, the *STATSGO* data was selected for this study. Its resolution is 30-m. The *K* factor, or soil erodibility, is available in *STATSGO*. The *K* factor for New Mexico has a scale between 0 and 0.64.

Methods

ArcGIS® Toolsets: The *ArcCatalog Export* tool converts outputs from .img extension, which is the *FuzzyCell* output, to a raster format for further processing in *ArcMap*. The *ArcCatalog Project* tool assigns the data to the specified projection, which was Albers Equal Area Conic Projection (1805) due to its minimal distortion. In order to get the adjusted shape of New Mexico, some modifications were necessary, including changes in assigning the central meridian to -106.894440, standard parallel 1 to 32.274999, standard parallel 2 to 36.055000, and latitude of origin to 34.256940.

The *Mask* tool, which provides the user with an opportunity to work only with the related cells or regions and mark the remaining areas as "NoData", is used when the smaller blocks of data are needed to be processed due to the limited capabilities of *FuzzyCell*. After all the blocks of the soil erosion risk are calculated, the *Merge* tool seams the data into one layer to get a single-piece risk map for New Mexico.

Using the *Spatial Analyst Surface* and *Hydrology* tools and the 30-m *DEM*, a slope gradient (θ) raster and a 1-dimensional *LS* factor raster were developed. The equations used for the *LS* factor are provided in Equations 2, 3, 4, 5, and 6.

$$L = (\lambda / 22.13)^m \qquad \text{Eq. 2}$$

$$m = (1 + \beta) \qquad \text{Eq. 3}$$

$$\beta = (\sin\theta / 0.0896) / \left[3 \times (\sin\theta)^{0.8} + 0.56\right] \qquad \text{Eq. 4}$$

$$S = 65.4 \sin^2\theta + 4.56 \sin\theta + 0.0654 \qquad \text{Eq. 5}$$

$$LS = L \times S \qquad \text{Eq. 6}$$

where L is the length factor, m is a β-based exponent, θ is the slope gradient in radians, and S is the slope factor. The resultant *LS* factor values range between 0 and 31877 for New Mexico.

Fuzzy Logic and *FuzzyCell*: Liu *et al.* (2009) outlined the four main steps necessary for fuzzy reasoning, which are computing compatibilities, truncating conclusions, aggregating truncated conclusions, and defuzzification process.

The first step of the fuzzy reasoning is compatibility computation, which designates a crisp value to each antecedent (the part starting with *IF* up to *THEN*). *FuzzyCell* operates on the input raster datasets pixel by pixel to generate an output. For each input variable (contributing factor) the tool reads its raw value for that pixel from the raster data and uses its membership function to calculate a membership value. After finding all four membership values, a fuzzy operator is applied to them to produce a single membership value. For this study, the trapezoidal-norm (t-norm) operator "product" is used, which forces the tool to draw conclusions sensitive to every input. *FuzzyCell* repeats this procedure and calculates a compatibility value for each rule.

Each indicator is considered to have three different membership functions (low, medium, and high) that quantify the grade of its membership. Since there are 4 indicators each having 3 different possible memberships, there are 3^4 or 81 *IF-THEN* rules that must be constructed. For instance:

> *IF* slope gradient is high (steepness), soil erodibility is high, *NDVI* is medium, and precipitation depth is high, *THEN* risk is very significant.

The second step of fuzzy reasoning is truncating conclusions. Once the compatibility for a rule is calculated, the membership function of the consequent part (from *THEN* up to end) is truncated by the compatibility value using a fuzzy implication operator, which is '*min*' for this study.

The next step is aggregating truncated conclusions, which generates a new single fuzzy set representing the truncated conclusions for each rule. For this study, the '*max*' aggregation operator is used to truncate the conclusions.

The last step is defuzzifying the overall conclusion to convert a fuzzy set back into a crisp value. There are several defuzzification methods, such as the weighted average, maximum membership, average maximum membership, and center of gravity. In this study, the *center of gravity* method is employed (Tayfur *et al.* 2003), which takes the centroid of the area under the membership function curve of a fuzzy set as the result (Liu 2009).

Fuzzy Membership and Cutoffs: A partially overlapping trapezoidal membership function was used for each indicator. Based on the study by Ahamed *et al*. (2000), cutoff values of 0.1, 0.3, and 0.5 were chosen for the soil erodibility distribution. Ahamed *et al.* (2000) also evaluated the *LS* factor cutoff values for fuzzy memberships and suggested cutoff values as 0.5, 3.5, 9 and 16, respectively, for slope gradient classes of 0-5, 5-15, 15-30, and greater than 30%. Using these classes as a guide, cutoff values for the slope gradient (θ) were taken as 5, 10, and 15%.

In order to calculate the *NDVI* cutoff values, *C* factor cutoff values were assigned 0.2, 0.5, and 0.8 (Ahamed *et al.* 2000). *NDVI* values were then back calculated as 110, 125, and 145 using a *NDVI-C* factor relationship developed by van der Knijff *et al*. (1999).

The $_2P_6$ membership cutoff values were assigned as 0.63, 1.32, and 1.73 inches using the relationship developed by Wischmeier and Smith (1978) and *R* factor cutoffs of 10, 50, and 90 as hundreds of ft·ton$_f$ ·in /ac·hr for low, medium, and high rainfall erosivity. This inverse method closely matched the minimum, average, and maximum $_2P_6$ data for New Mexico.

The soil erosion risk membership was defined so that it would scale the outcome from 0 to 1. A partially overlapping triangular membership function with cutoff values spanning 7 fuzzy ranges was assumed. These soil erosion risk consequences were defined as *very insignificant* (vi), *insignificant* (i), *slightly insignificant* (si), *neutral* (n), *slightly significant* (ss), *significant* (s) and *very significant* (vs).

Results and Discussion

Figure 1 is the statewide erosion risk map based on the *LS* factor. The risk varied between 0.05 and 0.87 on a scale from 0 to 1. The red to blue color scale provides a visual perspective of potentially problematic areas. Note that *Spatial Analyst* was used to evaluate the difference between this and the erosion risk map based on θ. In general on a statewide basis, the erosion risk calculated using θ as an attribute is similar to the erosion risk determined with *LS*. However, the authors propose that erosion risk based on a combined attribute incorporating flow length and slope more appropriately models erosion potential.

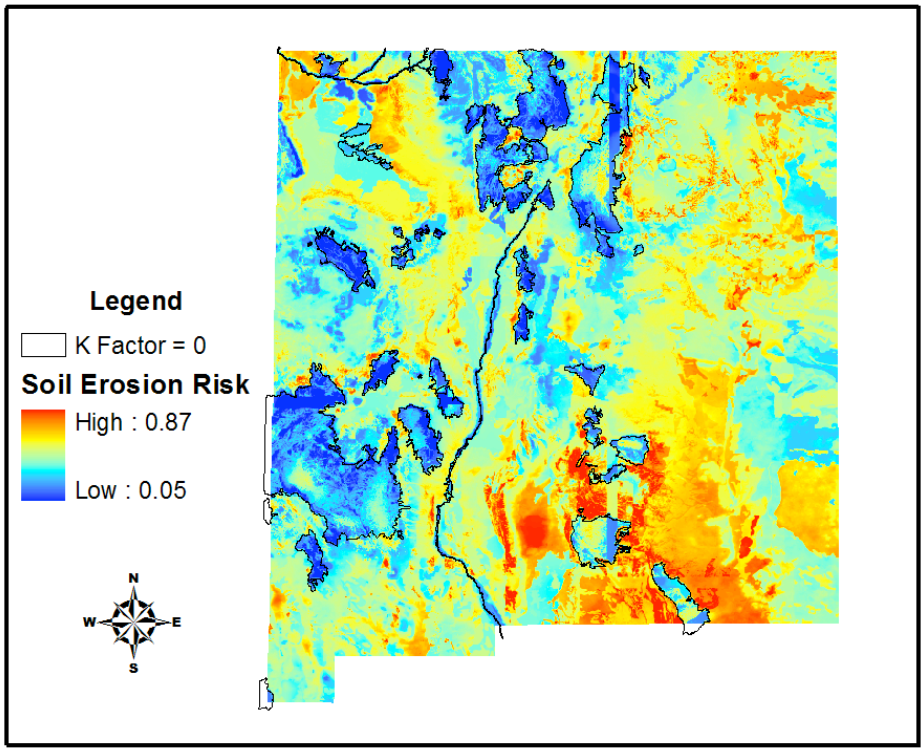

Figure 1: Soil Erosion Risk Map for New Mexico.

With the superposition of individual rasters for *NDVI*, θ, *K*, and $_2P_6$, (not shown), one can start to interpret the combinations of one or more indicators that contribute to a given high or low erosion risk. For example, far northeast section of New Mexico shows moderate to high soil erosion risk. Here precipitation is high, soil erodibility is high, vegetative coverage is medium, and the terrain is flat. Along the eastern border, the soil erosion risk is low. Soil erodibility is low; however, the other three indicators are the same as above. The mid-south and southern regions show the highest erosion risk, having erodible soils, medium to dense vegetative cover, and low precipitation.

It is expected that low *NDVI* values result in high soil erosion risk due to the fact that there is not enough vegetation or land cover to inhibit soil movement. Steep slope gradients generally result in high soil erosion risk. Erodible soils are prone to detachment and movement by water downslope by intense precipitation. The combinations of indicators and the dominance by one or more are complex and difficult to interpret and generalize on a region by region basis.

The black framed areas in Figure 1 delineate areas there is either no soil erodibility data or the value is zero. This tends to bias those areas to have lower defuzzified soil erosion risk values. When the *K* factor is zero, the membership function automatically becomes 1 (full belonging to *low soil erodibility range*). Since the defuzzification method used in this study is *center of gravity*, the outcome tends to be lower when one factor is zero. In addition, when soil erodibility is low, there

are only 27 (3^3) rules left that can be applied to impact the outcome, depending on the pixel values of *LS*, *NDVI* and $_2P_6$. As evident in Figure 1, the black framed areas tend to have lower soil erosion risk (blue).

Figure 1 identifies potentially problematic regions within New Mexico for soil erosion by water. Interestingly, a shapefile of problematic areas for culvert clogging was provided by *NMDOT*. This shapefile was superimposed upon the statewide soil erosion risk map. A region specific example is given in Figure 2. Major roads are shown for reference. The relative correlation of a problematic area with high risk (red to yellow) is vividly evident for some *NMDOT* locations. For other areas (areas of low risk), the opposite is apparent (blue to turquoise); however, a general consensus of the utility of this overlay exercise is that the problem areas are bias to high risk (red to yellow). Other factors, such as culvert design not addressed in the soil erosion risk methodology herein, may be a controlling influence on sediment deposition within the culvert barrel.

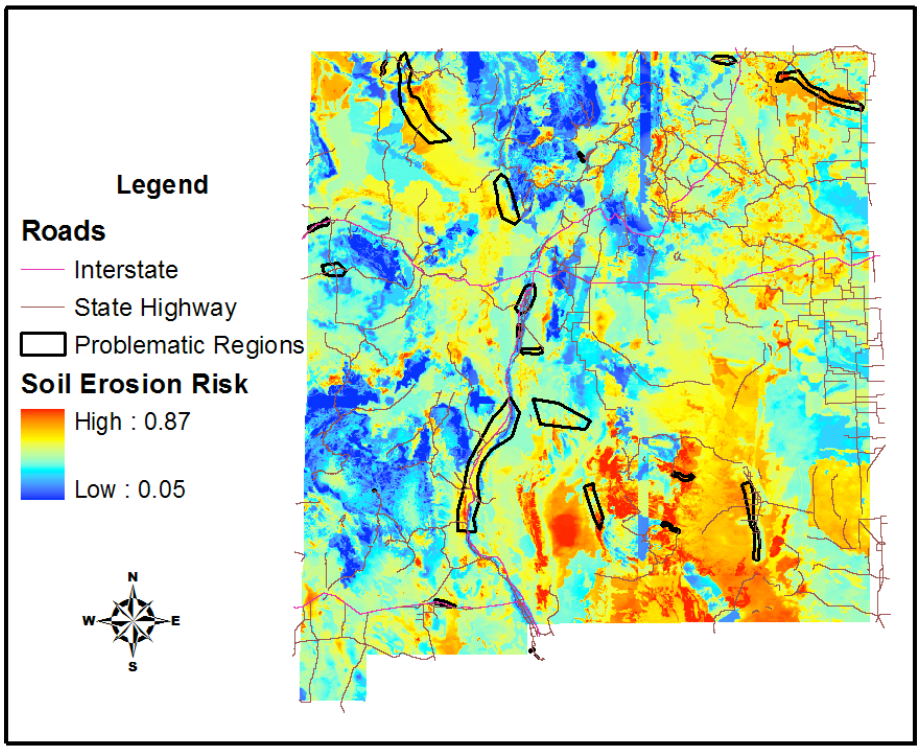

Figure 2: Problematic Areas Provided by *NMDOT*.

Summary and Conclusion

Quantifying the risk of soil erosion within a geographic region would provide a useful engineering management and design tool. To this end, a first-iteration soil erosion risk map was generated for the State of New Mexico. The method assumed that soil erodibility, precipitation, vegetation, and topography were primary factors

that contribute to soil erosion. Areas for soil erosion were identified based on a color scale from 0 (low risk) to 1 (high risk). However, the value of the statewide map as a management or design tool is somewhat diminished as no field verification has been conducted to date to support the results other than a digital shapefile provided by *NMDOT* showing several problematic areas for culvert deposition and clogging.

Recommendations for Future Work

Using finer resolutions for each indicator, introducing more or different indicators, and redefining fuzzy memberships can be performed to validate results and improve methods used in this study. Site specific field reconnaissance can be used to verify the accuracy of the results, as well as adjust the method and revise the erosion risk output. Problematic areas can be evaluated via laboratory and field analyses and *FuzzyCell* input can be adjusted using this data. Once the data and *FuzzyCell* output are compatible, the same adjustment can be used for the rest of the state to generate a calibrated soil erosion risk map. These adjustments could be performed by changing the indicator membership cutoff values or membership functions.

The *GIS* resolution employed herein was a 30-m by 30-m grid. Using different resolutions with fuzzy models, such as a 10-m by 10-m grid, to study problematic regions can increase the value of the soil erosion risk estimation approach used in this study. Regions with the highest risks can be analyzed as separate Hydrologic Unit Code (*HUCs*) or individual watersheds in order to get a finer definition of risk and connection between appropriate contributing factors. This micro-scale analysis can lead to a macro-scale statewide soil erosion risk map using extrapolation techniques based on a reduced level of critical contributing factors.

The number and types of indicators in the fuzzy model can be changed, and the results can be compared. Variables affecting soil erosion can be selected or modified based on literature; for example, the use of *LS* in lieu of θ as an attribute. The number of indicators can be increased to study new correlations or decreased to eliminate insensitive ones. In this sense, a 2-dimensional *LS* factor (Moore and Burch, 1986), in lieu of a 1-dimensional *LS* factor, can be used to show the effect of flow convergence on the erosion risk estimate, incorporating indirectly the attribute of upslope contributing area. A more physiographic-based estimate of potential soil erosion should result. The *LS* factor decreases when estimated with upslope contributing area as compared with the 1-dimensional *LS* (Rodriguez and Suárez 2010). The resultant annual soil erosion estimate would decrease based on Equation 1. The erosion risk estimate using *FuzzyCell* would be modified as well. However, an appropriate membership would need to be evaluated, perhaps contrasting slope gradient (%) with *LS* as was done by Ahamed *et al.* (2000).

Different membership cutoff values for *NDVI*, θ, *K*, and ${}_2P_6$ can be processed in the *FuzzyCell* environment to further evaluate the sensitivity of the variables and/or to generate new soil erosion risk estimates. Additionally, coupled with field verification, this is more likely to result in a defensible soil erosion risk map for the State of New Mexico.

For the study herein, three levels for a fuzzy set of *NDVI*, θ (or *LS*), $_2P_6$, and *K* were specified. This relegates the fuzzy inference methodology to being a rather coarse resolution of vagueness and, consequently, the soil erosion risk estimate; however, it is computationally simpler to handle. For a finer resolution, for example with five levels each, the number of rules required becomes 5^4 *versus* 3^4. A finer resolution for some attributes may be required based on field verification and sensitivity analysis; for other attributes a coarser specification may suffice. An increase in the number of rules requires more calculation time. Although doable with *FuzzyCell*, this would entail a significant effort to divide the state into small areas for computation and then stitch them together by the *Mosaic* tool.

References

Ahamed, T.R., K. Gopal Rao, and J.S.R. Murthy. (2000). "Fuzzy Class Membership Approach to Soil Erosion Modelling." *Agricultural Systems*, **63**(2), 97-110.

Liu, K.F.R., Liang, H. H., Yeh, K. and Chen, C. W. (2009). "A Qualitative Decision Support for Environmental Impact Assessment Using Fuzzy Logic." *Journal of Environmental Informatics*, **13**(2), 93-103.

Moore, I.D. and G.J. Burch (1986). "Physical Basis of the Length-Slope Factor in the Universal Soil Loss Equation." *Soil Sci. Soc. Am. J.,* **50**(5), 1294-1298.

Rodriguez, J.L. and Suárez, M.C.G. (2010). "Historical Review of Topographical Factor." *Aqua-LAC*, **2**(2), 56-61.

Tayfur, G., Ozdemir, S., and Singh, V.P. (2003). "Fuzzy Logic Algorithm for Runoff-induced Sediment Transport from Bare Soil Surfaces." *Advances in Water Resources*, **26**(12), 1249-1256.

van der Knijff, J.M., Jones, R.J.A, and Montanarella, L. (1999). "Soil Erosion Risk Assessment in Italy." European Soil Bureau, European Commission Joint, L.Research Centre Institute for Environment and Sustainability.

Wischmeier, W.H. (1959). "A Rainfall Erosion Index for a Universal Soil-Loss Equation." *Soil Sci. Soc. Am. J,* **23**(3), 246-249.

Wischmeier, W.H. and Smith, D.D. (1978). "Predicting Rainfall Erosion Losses: A Guide to Conservation Planning" U.S. Dept. of Agriculture, Science and Education Administration and Purdue University, Washington, D.C.

Yanar, T.A. and Akyürek, Z. (2006). "The Enhancement of the Cell-based *GIS* Analyses with Fuzzy Processing Capabilities." *Information Sciences,* **176**(8), 1067-1085.

Zadeh, L.A. (1965). "Fuzzy Sets. Information and Control." *Information and Control*, **8,** 338-353.

Acknowledgements: This research was funded in part by the New Mexico Department of Transportation (*NMDOT*) Research Division.