USING CLIGEN TO GENERATE RUSLE CLIMATE INPUTS

B. Yu

ABSTRACT. *CLIGEN is a stochastic weather generator that produces continuous daily variables to drive process–based runoff and erosion prediction models such as WEPP. To test CLIGEN's ability to generate precipitation–related variables, which are particularly important to runoff and erosion prediction, algorithms were developed to compute the R–factor, its monthly distribution, and 10–year storm erosion index (EI) needed to apply the Revised Universal Soil Loss Equation (RUSLE). Measured R–factor and 10–year storm EI for 76 sites in the U.S. were used for calibration, and 89 additional sites were used for validation. It was found that the generated R–factor using CLIGEN is highly correlated with the measured R–factor for the calibration sites (r^2 = 0.96), although the generated R–factor is systematically larger than the measured R–factor. The predicted R–factor for validation sites has a model efficiency (E_c) of 0.92 and a root mean squared error of around 600 MJ mm ha^{-1} h^{-1} $year^{-1}$, or 24% of the average R–factor for the 89 sites. In addition, CLIGEN–generated precipitation data can also be used to predict 10–year storm EI (E_c = 0.75) and monthly distribution of rainfall erosivity for a wide range of climate environments (average discrepancy = 2.6%). This represents considerable improvement over existing methods to estimate R–factor and 10–year storm EI for locations with only monthly precipitation data, although the systematic over–estimation of the R–factor using CLIGEN–generated climate data suggests possible inadequacies in the assumed storm patterns in CLIGEN and WEPP.*

Keywords. *WEPP, Storm pattern, Stochastic weather generator.*

CLIGEN is a stochastic weather generator to produce, among other things, continuous daily climate files to run WEPP for runoff and soil loss predictions (Nicks et al., 1995; Flanagan and Nearing, 1995; Laflen et al., 1991, 1997). Ten weather variables are generated for each day of the simulation period. The quality of the four precipitation–related variables is of particular importance because previous studies have shown that predicted runoff and soil loss are most sensitive to these precipitation variables (Nearing et al., 1990; Chaves and Nearing, 1991). Yu (2000) noted a critical coding error in relation to storm generation in CLIGEN, modified the algorithm to simulate the peak rainfall intensity, and tested CLIGEN using break–point rainfall data for 14 sites in the U.S. in terms of predicted runoff and soil loss. Subsequently, the entire CLIGEN program was re–coded to conform to WEPP coding conventions, with the modified algorithm for generating peak rainfall intensity implemented and command line options introduced (Meyer, 2001; Flanagan et al., 2001). As a result, the program structure has been simplified considerably, and code readability vastly improved. In addition, standard normal deviates used to generate climate variables by CLIGEN are "quality–controlled" by testing monthly batches of normally distributed deviates as the simulation progresses (Meyer, 2001; Flanagan et al., 2001). Meanwhile, the CLIGEN parameter database has been expanded to include about 2600 sites in the U.S. (Flanagan et al., 2001).

Tests of CLIGEN in terms of runoff and soil loss predicted with WEPP, although highly relevant, were confounded by the fact that specific soil, topography, management, and infiltration and erodibility parameter values have to be used. Performance of CLIGEN was thus conditional upon other input requirements for WEPP. It is therefore desirable to consider other performance indicators that rely solely on rainfall characteristics. If CLIGEN can generate climate data for process–based erosion prediction models such as WEPP, then CLIGEN would logically be expected to provide climate input for other erosion prediction models, such as the R–factor for RUSLE (Revised Universal Soil Loss Equation, Renard et al., 1997). In fact, Nicks and Gander (1994) apparently calculated the R–factor for the USLE for the eastern U.S. (east of the 105th meridian) and found that "while there is not exact agreement between the contour lines constructed using CLIGEN and those given in the USLE handbook, the pattern is quite similar...". However, it is not clear how the R–factor was calculated and which sites in the eastern U.S. were used. It is the objective of this article to use the latest corrected CLIGEN to test its ability to generate the R–factor, its monthly distribution, and 10–year storm EI values for RUSLE for 165 sites across the U.S. This not only provides a further reality check on CLIGEN but, when validated, will allow CLIGEN to be used as an appropriate tool for generating climate data for both RUSLE and WEPP in an integrated modeling environment such as MOSES (Meyer et al., 2001).

Renard and Freimund (1994) reviewed methods to estimate the R–factor using climate data that are readily available, such as monthly precipitation amounts. They also developed regression equations to estimate the R–factor from mean annual precipitation and 10–year storm EI from the measured R–factor for 132 sites in the U.S. Their results

Article was submitted for review in August 2001; approved for publication by the Soil & Water Division of ASAE in March 2002.

The author is **Dr. Bofu Yu**, Faculty of Environmental Sciences, Griffith University, Nathan, Queensland, 4111, Australia; phone: 617–3875–5258; fax: 617–3875–7459; e–mail: b.yu@mailbox.gu.edu.au.

  ISSN 0001–2351

provide an important yardstick against which performance of CLIGEN will be assessed.

The metric unit for the R–factor is MJ mm ha^{-1} h^{-1} $year^{-1}$, and MJ mm ha^{-1} h^{-1} for 10–year storm EI. Throughout this article, wherever appropriate "R–factor in SI units" instead of "R–factor in MJ mm ha^{-1} h^{-1} $year^{-1}$" is used for simplicity. The same also applies to 10–year storm EI. To obtain the R–factor in U.S. customary units of hundreds of foot tonf inch $acre^{-1}$ h^{-1} $year^{-1}$, the R–factor in SI units needs to be divided by a factor of 17.02 (Foster et al., 1981; Renard et al., 1997).

Two different versions of CLIGEN are considered in this article: Version 5.101 (Meyer, 2001), and Version 5.101 with "quality–control" measures disabled. Instead of the original random number generators used in Version 5.0 and prior versions, subroutines ran1 and gasdev from Numerical Recipes (Press et al., 1992) were used to generate uniform and normal random deviates for CLIGEN. Again for simplicity, these versions are referred to as V5.101 and NR1, respectively, in this article.

Materials and Methods

Data

Although CLIGEN parameter files are available for more than 2600 sites in the U.S., parameter values for solar radiation and storm intensity variables originated from 142 sites using the method of triangulation. For each site in the CLIGEN database, the weighted average of parameter values from three nearby sites was used. When a CLIGEN site coincides with one of the original 142 sites, a weighting factor of 1 would appear, indicating that parameter values from other sites were not used. This weighting scheme is an important consideration when selecting sites for the R–factor comparison.

Mean annual rainfall, R–factor, and 10–year storm EI values were extracted for 165 sites from the RUSLE database. These constitute all the sites from the RUSLE database for which CLIGEN parameter files are available, except that Honolulu is the only site selected from 17 available sites in Hawaii. The remaining 16 sites in Hawaii were discarded because they all had an identical 10–year storm EI value of 160 (U.S. customary units) although the R–factor varied from 100 to 400 (U.S. customary units) among these sites. For each of the 165 sites, a corresponding CLIGEN parameter file was extracted either for the site or for the nearest site. The 165 sites were separated into two groups. The first group contains all the sites with a weighting factor of 1, indicating that the R–factor, 10–year storm EI, and CLIGEN parameters values are likely to have originated from the same data source. These sites, 76 in total, were used for model calibration. The remaining 89 sites were used for model validation. For the validation sites, the average distance between the RUSLE sites and CLIGEN sites is about 21 km, with a maximum distance of about 79 km between Scoby, Montana (RUSLE site) and Wolf Point, Montana (CLIGEN site). Figure 1 shows the distribution of these calibration and validation sites in the contiguous U.S. and Hawaii.

In addition, WEPP climate input files were used to calculate the R–factor for 14 USDA–ARS sites. These climate files in WEPP input format were prepared using observed precipitation data for previous WEPP validation studies (Risse et al., 1995; Zhang et al., 1995a, 1995b; Liu et al., 1997). These files have been used to evaluate CLIGEN's performance in terms of predicted runoff and soil loss with WEPP (Yu, 2000). While measured R–factor values were available from the RUSLE database for some of the 14 sites, most of the R–factor values for these sites had to be estimated from isoerodent maps for the U.S. (Renard et al., 1997).

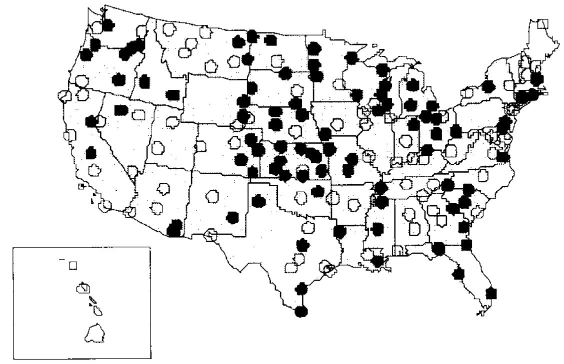

Figure 1. Calibration (open circles) and validation (dots) sites in the contiguous U.S. and Hawaii.

Algorithm

CLIGEN generates four precipitation–related variables for each wet day (Nicks et al., 1995): precipitation amount P (mm), storm duration D (h), time to peak as a fraction of storm duration, t_p, and the ratio of peak intensity over the average intensity, i_p. In CLIGEN, time is normalized by storm duration, D, and rainfall intensity is normalized by the average intensity, P/D. Therefore both t_p and i_p are dimensionless variables, and they can be regarded as normalized time to peak and normalized peak intensity, respectively.

A double exponential function is used to describe the normalized intensity pattern as:

$$i(t) = \begin{cases} i_p e^{b(t-t_p)} & 0<t<t_p \\ i_p e^{-d(t-t_p)} & t_p<t<1 \end{cases} \qquad (1)$$

where b and d are parameters for the storm pattern. In essence, rainfall intensity is assumed to rise exponentially with time from 0 to t_p, and then decrease exponentially from t_p to 1. Two further assumptions were made (Nicks et al., 1995):

1. Rainfall intensity at the beginning and at the end of the storm is the same, i.e., i(0) = i(1).
2. The area under the curve defined by equation 1 from 0 to t_p is assumed to be equal to t_p, so:

$$t_p = i_p \int_0^{t_p} e^{b(t-t_p)} dt \qquad (2)$$

With these assumptions, the storm pattern can be uniquely described by:

$$i(t) = \begin{cases} i_p e^{b(t-t_p)} & 0<t<t_p \\ i_p e^{-bt_p(t-t_p)/(1-t_p)} & t_p<t<1 \end{cases} \qquad (3)$$

where b is the only parameter that satisfies:

$$i_p(1-e^{-bt_p})-bt_p=0 \quad (4)$$

Nicks et al. (1995) restricted b to values less than 60 so that Newton's method could be used to solve for b with the microcomputers available at the time. In fact, it is much better to regard the product of b and t_p as a new parameter B and solve for B instead of the original parameter b because the solution to equation 4 in terms of B is guaranteed to lie in the range between $\ln(i_p)$ and i_p. The new parameter B thus can be determined readily and efficiently for any given value of i_p using Newton–Raphson's method within the range specified above.

For each day when precipitation occurs and when mean air temperature is greater than 0° C, peak 30–min intensity is calculated as follows. If D is less than or equal to 30 min, then I_{30} = 2P (mm h^{-1}) by definition. If D is greater than 30 minutes, maximizing the amount of rain in a given time interval (Δt), then the peak intensity is:

$$i_{\Delta t}=\frac{i_p}{B\Delta t}(1-e^{-B\Delta t}) \quad (5)$$

Re–normalization yields the 30–min peak rainfall intensity as:

$$I_{30}=\frac{2Pi_p}{B}\left(1-e^{-\frac{B}{2D}}\right) \quad (6)$$

The unit energy equation recommended for RUSLE is given by:

$$e(i)=e_o(1-\alpha\, e^{-I/I_o}) \quad (7)$$

where e_o = 0.29 MJ ha^{-1} mm^{-1}, α = 0.72, and I_o = 20 mm h^{-1} (Brown and Foster, 1987; Renard et al., 1997). The total storm energy, E, can be derived by integrating the unit energy over the double exponential storm pattern, which leads to:

$$E=Pe_o\left[1-\frac{\alpha\, i_p}{B}\frac{I_o}{I_p}\left(e^{-\frac{I_p}{I_o}e^{-B}}-e^{-\frac{I_p}{I_o}}\right)\right] \quad (8)$$

where I_p is peak intensity (mm h^{-1}).

The daily storm erosion index, EI, is the product of equation 6 and equation 8. These are accumulated for each month, and the R–factor, by definition, is the sum of mean monthly EI values. A program, CLG2RF, was written to implement the algorithm described above. For any WEPP climate input file(s), including those generated by CLIGEN, CLG2RF calculates daily storm EI values whenever liquid precipitation occurs, and outputs (1) mean annual precipitation, (2) R–factor, (3) monthly distribution of rainfall erosivity, and (4) 10–year storm EI. The 10–year storm EI value is determined from an annual series of maximum storm EI values. Each value in the annual series is assigned an average recurrence interval using Weibull's formula (Maidment, 1993). The 10–year storm EI value can be determined either directly or using the linear interpolation technique. The program can handle climate data for either single or multiple sites. In addition, users can specify a precipitation threshold below which the storm EI values are excluded from calculations. Yu (1999) investigated the effect of using different precipitation thresholds on calculated R–factor and found that the effect can be noticeable, especially for areas with low mean annual precipitation. For this article, all liquid precipitation was included in R–factor calculations. This is consistent with the method used for preparing the isoerodent map for the western U.S. and with the recommendations for calculating the R–factor for RUSLE (Renard et al., 1997).

PROCEDURE

CLIGEN V5.101 was used to generate climate data for a period of 100 years for each of the 165 sites. V5.101 was then modified to use random number generators, namely ran1 and gasdev, from Numerical Recipes (Press et al., 1992) to create a new version called NR1. All of the "quality control" procedures in V5.101 were disabled in NR1. Another period of 100 years' data was generated with NR1 for each of the sites. For each site, a random seed for V5.101 was selected in the range from 1 to 10,000. Larger random seeds were not used because V5.101 first generates random numbers for a number of times that equals the random seed specified before the calculation proper commences. Larger random seeds would use unnecessarily large amounts of computation time. For NR1, random seeds were selected in the range from 1 to 2^{31}–1 as allowed by ran1. The random seeds used for both types of simulations were recorded for each site so that the results can be readily reproduced if needed.

CLIGEN generates storms on a daily basis. Only one storm is generated on wet days. The double exponential storm pattern is assumed for all storms. Because of these assumptions, hence limitations, R–factor and 10–year storm EI produced by CLIGEN and CLG2RF are termed generated R–factor and generated 10–year storm EI. It is hypothesized that the measured and generated erosivity values are the same, or at least that a good relationship between the two exists and that the relationship is consistent across all climate regions. For calibration sites, standard linear regression technique was used to examine the relationship between measured and generated R–factor and 10–year storm EI. For validation sites, predicted R–factor or 10–year storm EI values were simply compared with the measured values. The coefficient of efficiency (Nash and Sutcliffe, 1970), E_c, was used to quantify the model performance. The coefficient of efficiency has lately become the standard measure for model validation purposes and has been widely used for WEPP validation studies (e.g., Zhang et al., 1996; Tiwari et al., 2000; Yu et al., 2000).

Monthly or half–monthly percentage distribution of rainfall erosivity is needed to calculate the weighted cover factor and is also useful for identifying periods of high erosion risk. In the RUSLE database, the half–monthly distribution is tabulated for each of 149 zones in the U.S. Six sites were selected to cover a wide range of precipitation regimes. The monthly distribution for the zone in which the selected site was located was compared with the monthly distribution predicted by CLG2RF using CLIGEN–generated climate files. It is worth noting that no calibration is required when comparing monthly distribution in terms of percentage contribution to the R–factor. The mean absolute difference in the monthly distribution of the R–factor was used as a measure of the discrepancy between the measured and predicted monthly distributions. This measure has been used to characterize the performance of daily rainfall erosivity models (Yu and Rosewell, 1996; Yu, 1998; Yu et al., 2001).

Results

Calibration

Figure 2 shows the relationship between the generated R–factor using CLIGEN–generated climate files for WEPP and the measured R–factor from the RUSLE database. It is clear that the generated R–factor is systematically larger than the measured R–factor for the calibration sites. From figure 2, it can also be seen that the data point for Baton Rouge, Louisiana, appears to be an outlier. In the RUSLE database, Baton Rouge has a R–factor of 11 914 MJ mm ha^{-1} h^{-1} year^{-1}, the same as that for New Orleans, but much higher than that for Mobile, Alabama, with R = 10 212 MJ mm ha^{-1} h^{-1} year^{-1}. The isoerodent map for the eastern U.S., however, shows similar R–factor values for Baton Rouge and Mobile. If we exclude the data point for Baton Rouge, then the calibration equation for the R–factor becomes:

$$R = 0.576 R_{gen}, \quad r^2 = 0.96 \qquad (9)$$

using NR1. The calibration equation for 10–year storm EI values:

$$EI = 0.609 EI_{gen}, \quad r^2 = 0.86 \qquad (10)$$

is shown in figure 3. Calibration results using V5.101 in addition to those presented above are summarized in table 1 for comparison. There is no difference of any significance between NR1 and V5.101 in terms of r^2 and root mean squared errors, although V5.101 produces slightly better calibration results.

Spatially consistent generation of weather variables is an important requirement for CLIGEN. Residuals from the calibration equations given above were examined in this context to determine whether systematic errors occur spatially. Rank correlation between the residuals and latitude, longitude, and altitude, respectively, were calculated for the calibration sites. The average rank correlation coefficient was 0.090, and the average t–value was 0.769, with the number of degrees of freedom being 73 for the calibration sites. None of the six rank correlation coefficients was significantly different from zero (at the 5% level) for the R–factor as well as the 10–year EI values. These results show that CLIGEN can produce geographically consistent climate data to compute the R–factor and 10–year storm EI across the contiguous U.S.

Validation

Calibration equations 9 and 10 were used to predict the R–factor and 10–year storm EI using CLIGEN–generated climate files for the 89 validation sites. Validation of the relationship between the measured and CLIGEN–generated R–factor, 10–year storm EI is presented in table 1 and figures

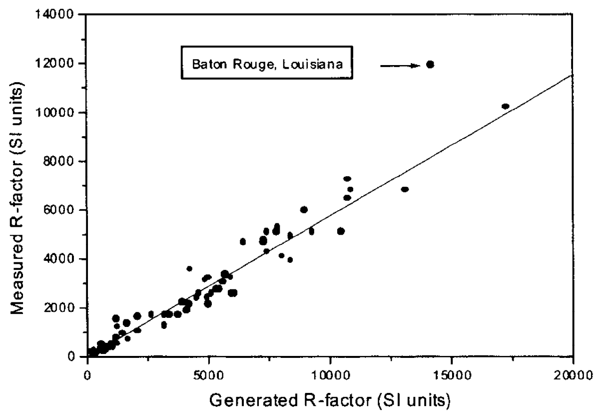

Figure 2. Relationship between R–factor generated with NR1 and measured R–factor from RUSLE database for 76 calibration sites.

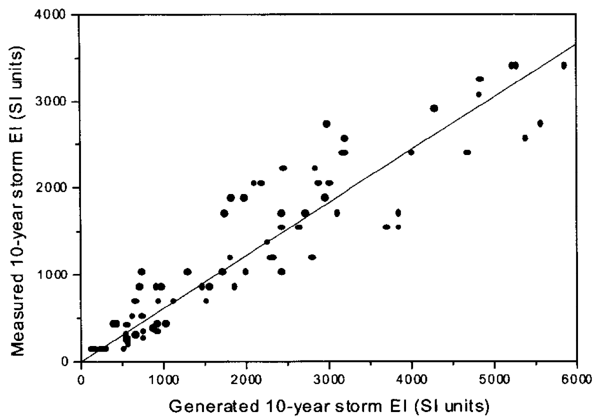

Figure 3. Relationship between 10–year storm EI generated with NR1 and the measured value from RUSLE database for 76 calibration sites.

Table 1. Calibration and validation results for R–factor and 10–year storm EI using CLIGEN and CLG2RF in comparison with those using mean annual rainfall: n = sample size, bias = the ratio of measured to generated values, E_c = model efficiency (Nash and Sutcliffe, 1970), RMSE = root mean squared error, and SEE = standard error of estimation (Renard and Freimund, 1994).

	Calibration				Validation			Nonlinear regression (Renard and Freimund, 1994)		
	n	bias	r^2	SEE (metric units)	n	E_c	RMSE (metric units)	n	r^2	SEE (metric units)
R–factor								132	0.81	1075
NR1[a]	75	0.576	0.96	455	89	0.92	604			
V5.101[a]	75	0.586	0.96	447	89	0.92	599			
10–year storm EI								132	0.63	565
NR1	76	0.609	0.86	366	89	0.75	422			
V5.101	76	0.680	0.87	347	89	0.74	425			

[a] Excluding Baton Rouge, Louisiana.

4 and 5. It can be seen from table 1 that validation results are not as good as calibration results in terms of root mean squared errors. It can also be seen that for validation sites, NR1 performed slightly better than V5.101 with respect to 10–year storm EI, and V5.101 performed slightly better than NR1 with respect to R–factor. The difference between the two random number generation schemes is small and insignificant again in terms of model efficiency and root mean squared errors. Figure 4 shows good agreement between measured and predicted R–factor for validation sites, although the predicted R–factor tends to be larger than the measured R–factor for regions of low R–factor, a pattern not clearly evident for calibration sites (fig. 2). For 10–year storm EI, the measured EI values tend to assume a few discrete values. For example, 9 of the 89 sites have an EI value of 1362 MJ mm ha^{-1} h^{-1} (80 U.S. customary units) and 6 sites have an EI value of 3404 MJ mm ha^{-1} h^{-1} (200 U.S. customary units). This strongly suggests that the original EI values in the RUSLE database lack precision and they may have been estimated or interpolated in a crude manner. It opens to question what model efficiency and root mean squared error would be, had better EI data been available for validation purposes.

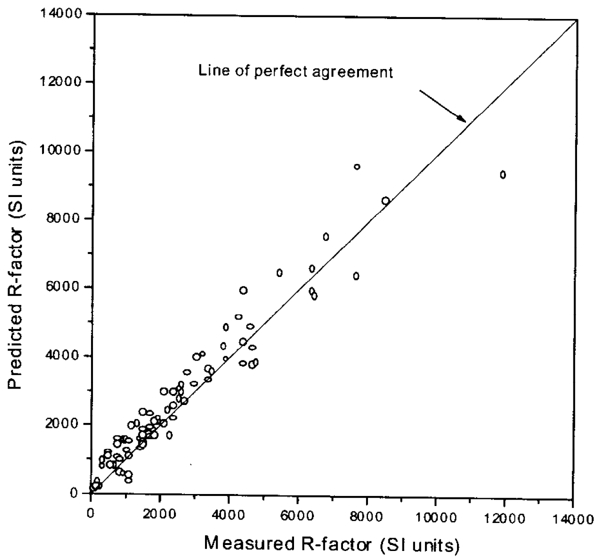

Figure 4. Comparison of predicted R–factor and measured R–factor from RUSLE database for 89 validation sites.

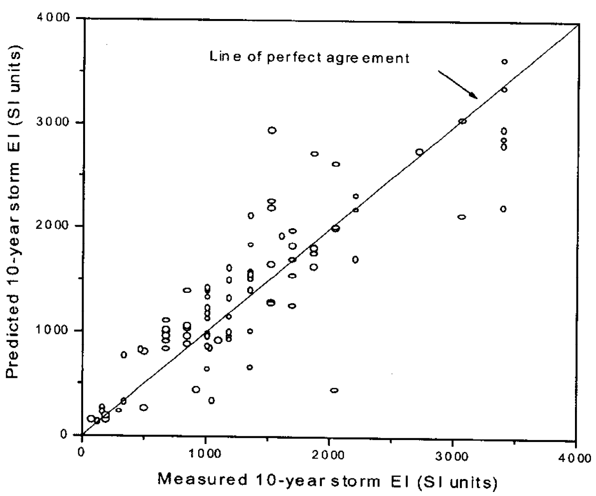

Figure 5. Comparison of predicted 10–year storm EI and measured values from RUSLE database for 89 validation sites.

The validation results (table 1) show considerable improvement over the non–linear regression technique used to estimate the R–factor (Renard and Freimund, 1994). It is worth noting that data from all 132 sites were used to develop the regression equation, and there was no independent test of the regression equation developed. In this study, the R–factor values for validation sites were not used to develop calibration equation 9. The root mean squared error using CLIGEN and CLG2RF is much smaller than the standard error of estimation using the regression technique, taking into account that the two measures of model error are essentially the same, differing only in terms of the number of degrees of freedom. Now that CLIGEN parameter files are readily available, there is essentially no additional cost in using CLIGEN and CLG2RF instead of the monthly precipitation data and regression equations as far as predicting the R–factor is concerned.

Renard and Freimund (1994) found that mean annual precipitation is a poor predictor of the 10–year storm EI for 132 sites in the U.S. They then developed a relationship between the R–factor and 10–year storm EI for these sites with $r^2 = 0.90$ and standard error of estimation of 297 SI units. To use the relationship to predict 10–year storm EI presumes that the R–factor is known *a priori*. Therefore, the standard error of estimation is not strictly comparable to the validation results reported in this article. For comparison purposes, the power functional form used by Renard and Freimund (1994) was fitted to the measured R–factor and 10–year storm EI for the 89 validation sites. The identical approach was then applied to the R–factor and 10–year storm EI predicted by NR1 and V5.101, respectively, using calibration equations presented above and in table 1. Figure 6 shows the power functions fitted to the measured and predicted 10–storm EI. It can be seen that the predicted relationship between R–factor and 10–year storm EI, especially that using NR1, is quite similar to the observed relationship between the two variables for the 89 validation sites.

Predicted monthly distribution of the R–factor is shown in figure 7 for six sites. Also shown in figure 7 is the measured

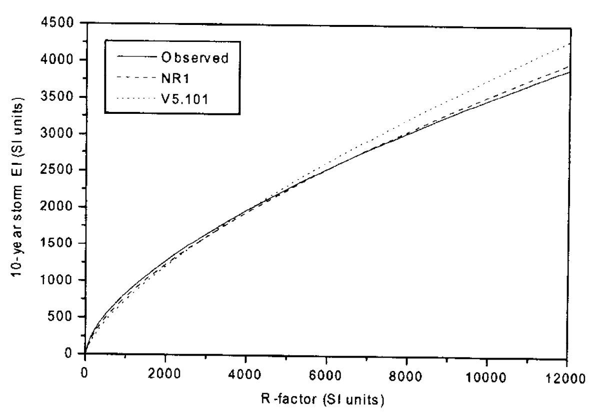

Figure 6. Comparison of observed and CLIGEN–generated relationship between R–factor and 10–year storm EI for 89 validation sites.

monthly distribution for the zone in which the selected sites are located. These sites were selected to represent different rainfall regimes in North America (Trewartha, 1981). Table 2 summarizes the rainfall regime and mean absolute difference between measured and predicted monthly distribution. It is clear from figure 7 that the predicted monthly distribution using CLIGEN and CLG2RF captures the seasonal distribution of erosivity for a wide range of precipitation regimes. The discrepancy between measured and predicted monthly distribution averages 2.6% for the six sites. It is also worth noting that Renard and Freimund (1994) did not attempt to estimate the monthly distributions from annual or monthly precipitation. Yu et al. (2001) argued that, as a minimum, daily rainfall data would be needed to adequately estimate the monthly distribution of erosivity.

14 USDA-ARS Sites

R–factor values for the 14 USDA–ARS sites used in Yu (2000) were related to the R–factor calculated with CLG2RF and observed precipitation data prepared in WEPP climate input format. Figure 8 shows a scatter plot for these 14 sites and the line of best fit through the origin. The regression equation for these 14 sites based on observed precipitation data is:

$$R = 0.581 R_{WEPP}, \qquad r^2 = 0.91 \tag{11}$$

with a standard error of estimation of 582 SI units. It is clear that the relationship is not as good as the calibration equation for the 76 sites described above (see also table 1). The large variations among the 14 sites are likely to be a reflection of

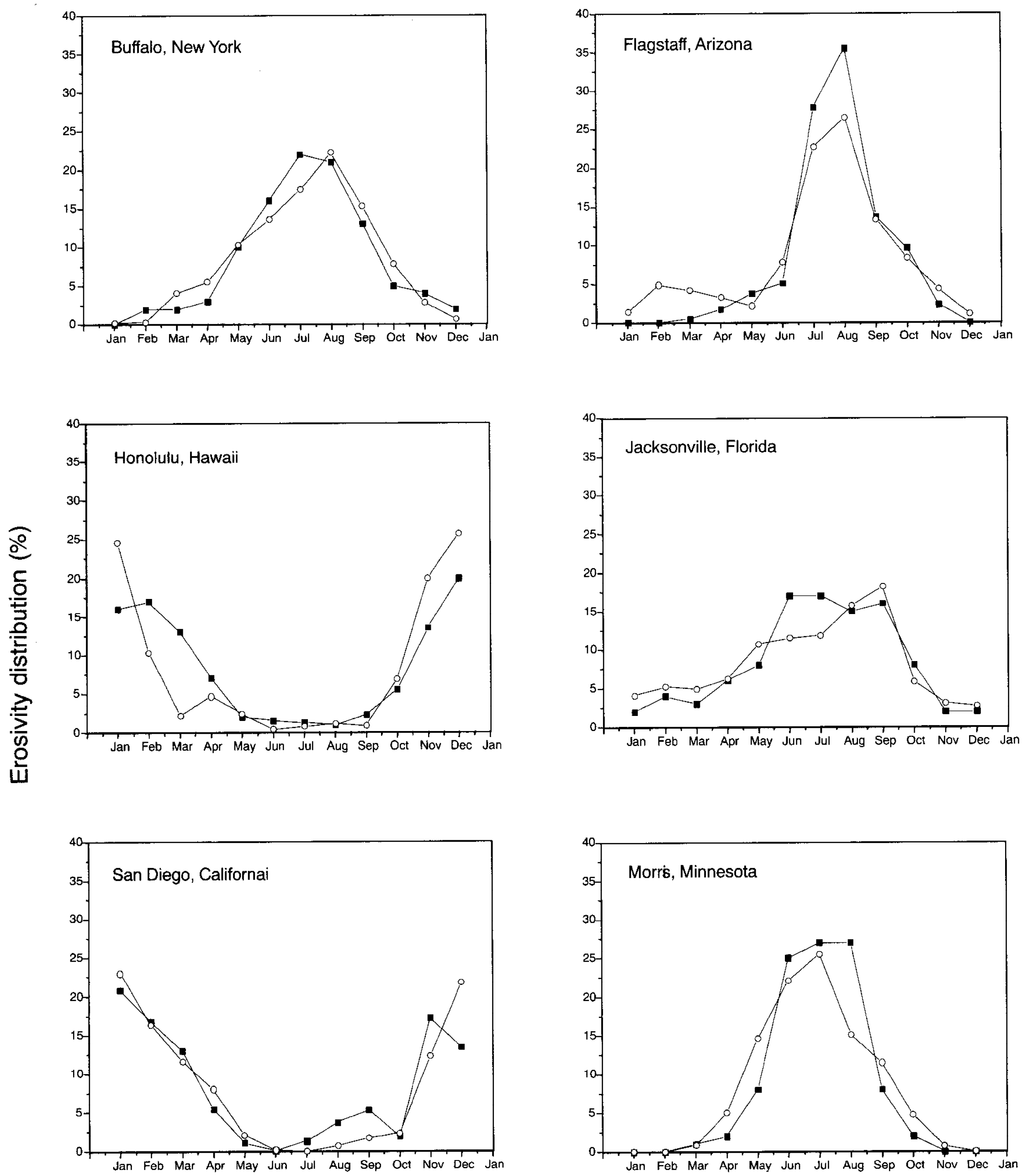

Figure 7. Measured (solid squares) and predicted (open circles) monthly distribution of rainfall erosivity for selected sites in the U.S.

Table 2. A comparison of the monthly distribution of erosivity for selected sites with different rainfall regimes. Identification of rainfall regime was based on Trewartha (1981).

Site	Rainfall Regime	Discrepancy (%)
Buffalo, New York	Eastern Maritime	1.9
Flagstaff, Arizona	Intermontane Transitional	2.9
Honolulu, Hawaii	Tropical Oceanic	3.8
Jacksonville, Florida	Sub–tropical Oceanic	2.1
Morris, Minnesota	Interior (Complex)	2.7
San Diego, California	Mediterranean	2.4

the natural climate variability, because these WEPP climate files prepared using observed precipitation data have an average record length of only about 10 years (Yu, 2000). The average record length is considerably shorter than the 100–year simulation period used for all calibration and validation sites. However, the relationship between the R–factor using observed precipitation data in WEPP input format and the measured R–factor for the 14 sites is remarkably similar to the calibration equation for the 76 sites (fig. 8).

The nearly identical relationship between the measured R–factor and that calculated using observed climate data for WEPP has two important implications. First, this shows that CLIGEN is able to produce P, D, t_p, and i_p values that are statistically similar to those based on observed climate data in terms of generated R–factor. Thus, the conclusions by Yu (2000) were corroborated without having to rely on runoff and soil loss parameters and a particular modeling framework. Second, the consistent relationship between measured and generated R–factors shows that the discrepancy between the two is not a result of CLIGEN *per se* but of the way storm patterns are represented in CLIGEN and WEPP. In particular, the assumptions of a single storm on wet days and all storms having a double exponential distribution are likely to be the main causes for the discrepancy.

Discussion

This article shows that CLIGEN is able to generate the

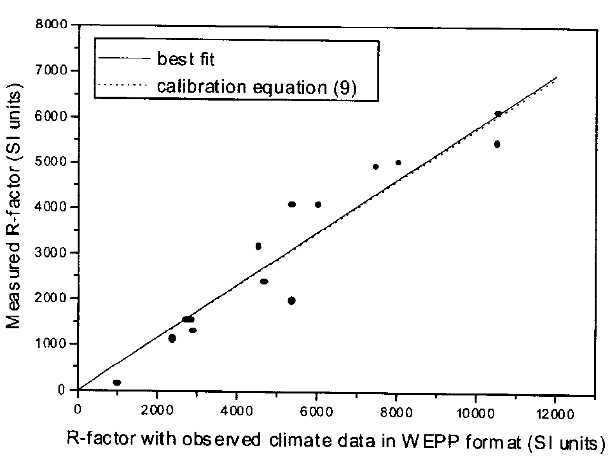

Figure 8. Relationship between measured R–factor and that calculated using observed precipitation data in WEPP input format for 14 USDA–ARS sites. The same data set was used in a previous attempt to evaluate CLIGEN's performance in terms of predicted runoff and soil loss (Yu, 2000). The dashed line shows the calibration equation for the 76 sites presented in figure 2.

required climate inputs for RUSLE, so long as calibration equations are used to adjust the generated R–factor and 10–year storm EI values. This, however, should not overshadow the fact that the R–factor generated by CLIGEN based on daily EI values is systematically higher than the measured R–factor. While no detailed investigation has yet been undertaken into the cause of these large generated R–factor values, it is quite likely that the over–estimation occurs because of accumulation of rain into a single storm on wet days and the assumption of a double exponential storm pattern for all storms. Brown and Foster (1987) used a single exponential function to represent re–ordered intensity data for 54 storms in the U.S. to simplify calculations of storm energy and peak 30–min intensity. Their results showed that on average, estimated I_{30} was about 35% greater than the measured I_{30} values. Their results also suggested that storm energy was slightly over–estimated (3.5% to 4.7% for the top half the data range for storm energy). They ascribed the over–estimation of I_{30} to multiple peaks of similar magnitude within a single storm. Although their results are not strictly comparable to what is reported in this article, because Brown and Foster (1987) used a single rather than double exponential function and used the energy equation for the USLE to calculate the storm energy, there is sufficient evidence to suggest that the bias in the generated R–factor is systematic rather than accidental. Thorough testing is called for to determine whether daily P, D, t_p, and i_p are adequate to represent actual storm patterns, not only in the context of reproducing daily storm energy and daily I_{30}, but also in the broader context of runoff and soil loss prediction using WEPP.

The issue of random number generators also requires further comments. While there is no significant difference between V5.101 and NR1 in terms of predicted R–factor and 10–year storm EI for the 165 sites tested, it is important to note that the random number generator used in CLIGEN Version 5.0 and prior versions is inadequate when subjected to rigorous statistical tests. Test results using DIEHARD developed by Marsaglia (1985, 2001) showed that the random number generator used in CLIGEN 5.0 and prior versions failed most of the tests. Ran1 from Numerical Recipes, by comparison, is able to pass all the tests except those involving the 32nd bit of the integer, as would be expected. It is therefore imperative to improve or replace the random number generators in CLIGEN Version 5.0 and prior versions.

I would argue for the use of ran1 in conjunction with gasdev in CLIGEN for the following reasons. First, these routines are readily available from a well–known source, and presumably these generators have been widely and critically examined. Second, ran1 would not fail to pass any statistical tests, except when the number of generated uniform deviates begins to exceed 10^8 (Press et al., 1992). Even for extremely wet areas where it rains every other day, CLIGEN would still require, on average, no more than 12 uniform deviates for each day of simulation. This implies that ran1 is quite adequate, unless the years of simulation start to approach 23,000. Superior generators with very long periods are available to replace ran1 if and when required (Marsaglia and Zaman, 1994; Press et al., 1996). Finally, although speed is rarely a major issue nowadays with ever–increasing computational power, NR1 is noticeably faster (by approximately a factor of 4) than V5.101.

Conclusion

CLIGEN is a useful tool for weather generation because it is relatively simple, it generates a wide range of weather variables, and more importantly, the required parameter values are available for a large number of sites. This article shows that CLIGEN not only can be used to supply simulated climate data on a daily basis for WEPP, it can also be used as an effective tool to generate the R–factor, its monthly distribution, and 10–year storm EI for RUSLE at minimum additional cost. Thus, CLIGEN is able to meet all RUSLE's climate input requirements. The quality of these estimates is superior to that using the existing methods and monthly precipitation data. When compared to observed climate data for 14 USDA–ARS sites, this article shows that CLIGEN is able to preserve the storm characteristics in terms of the generated R–factor for these sites. However, the systematic and consistent discrepancy between measured R–factor and the R–factor based on both CLIGEN–generated and observed precipitation data highlighted possible inadequacies in the assumed storm patterns in CLIGEN and WEPP.

Acknowledgements

The author is grateful to Charles Meyer of USDA–ARS for many useful discussions on the algorithm and the CLIGEN code, and for his enormous effort in re–coding CLIGEN and in initially pointing out the inadequacy of random number generators used in CLIGEN V5.0 and prior versions. The author is also grateful to Ken Renard for R–factor values for a number of USDA–ARS sites used in this article, and to Rebecca Trevithick of Griffith University for her enormous effort in number crunching and code verification. The author thanks ASAE reviewers for their constructive comments. This work on CLIGEN validation is partly supported through the Collaborative Research Centre for Catchment Hydrology Project 2.2: Managing pollutant delivery from dryland upland catchment.

References

Brown, L. C., and G. R. Foster. 1987. Storm erosivity using idealized intensity distributions. *Trans. ASAE* 30(2): 379–386.

Chaves, H. M. L., and M. A. Nearing. 1991. Uncertainty analysis of the WEPP Soil Erosion Model. *Trans. ASAE* 34(6): 2437–2444.

Flanagan, D. C., and M. A. Nearing, ed. 1995. USDA–Water Erosion Prediction Project: Hillslope profile and watershed model documentation. NSERL Report No. 10. West Lafayette, Ind.: USDA–ARS National Soil Erosion Research Lab.

Flanagan, D. C., C. R. Meyer, B. Yu, and D. L. Scheele. 2001. Evaluation and enhancement of CLIGEN weather generator. In *Soil Erosion Research for the 21st Century, Proc. ASAE Sym.*, 107–110. J. C. Ascough and D. C. Flanagan, eds. St Joseph, Mich.: ASAE.

Foster, G. R., D. K. McCool, K. G. Renard, and W. C. Moldenhauer. 1981. Conversion of the Universal Soil Loss Equation to SI metric units. *J. Soil Water Conserv.* 36(6): 355–359.

Laflen, J. M., L. J. Lane, and G. R. Foster. 1991. WEPP. A new generation of erosion prediction technology. *J. Soil and Water Conserv.* 46(1): 34–38.

Laflen, J. M., W. J. Elliot, D. C. Flanagan, C. R. Meyer, and M. A. Nearing. 1997. WEPP: Predicting water erosion using a process–based model. *J. Soil and Water Conserv.* 52(2): 96–102.

Liu, B. Y., M. A. Nearing, C. Baffaut, and J. C. Ascough II. 1997. The WEPP watershed model: III. Comparison to measured data from small watersheds. *Trans. ASAE* 40(4): 945–952.

Maidment, D. R., ed. 1993. *Handbook of Hydrology.* New York, N.Y.: McGraw–Hill, Inc.

Marsaglia, G. 1985. A current view of random generators. Keynote address. In *Proc. Computer Science and Statistics: 16th Symposium on the Interface.* Amsterdam, The Netherlands: Elsevier Science.

_____. 2001. DIEHARD: A battery of tests for random number generators. Available at: :http://stat.fsu.edu/~geo/diehard.html.

Marsaglia, G., and A. Zaman, 1994. Some portable very–long–period random number generators. *Computers in Physics* 8(1): 117–121.

Meyer, C. R. 2001. CLIGEN Weather Generator, expanded and improved. Available at: http://horizon.nserl.purdue.edu/Cligen/.

Meyer, C. R., L. E. Wagner, D. C. Yoder, and D. C. Flanagan. 2001. The Modular Soil Erosion System (MOSES), In *Soil Erosion Research for the 21st Century, Proc. ASAE Sym.*, 358–361. J. C. Ascough and D. C. Flanagan, eds. St Joseph, Mich.: ASAE.

Nash, J. E., and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models: Part I. A discussion of principles. *J. Hydrol.* 10: 282–290.

Nearing, M. A., L. Dear–Ascough, and J. M. Laflen. 1990. Sensitivity analysis of the WEPP Hillslope Erosion Model. *Trans. ASAE* 33(3): 839–849.

Nicks, A. D., and G. A. Gander. 1994. CLIGEN: A weather generator for climate inputs to water resource and other models. In *Computers in Agriculture 1994: Proc. 5th Int. Conf.*, 903–909. D. G. Watson, F. S. Zazueta, and T. V. Harrison, eds. St Joseph, Mich.: ASAE.

Nicks, A. D., L. J. Lane, and G. A. Gander. 1995. Weather generator. Chapter 2 in *USDA–Water Erosion Prediction Project: Hillslope Profile and Watershed Model Documentation.* D. C. Flanagan and M. A. Nearing, eds. NSERL Report No. 10. West Lafayette, Ind.: USDA–ARS Nat. Soil Erosion Research Lab.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical Recipes: The Art of Scientific Computing.* Cambridge, U.K.: Cambridge University Press.

_____. 1996. *Numerical Recipes in Fortran 90: The Art of Parallel Scientific Computing.* Cambridge, U.K.: Cambridge University Press.

Renard, K. G., and J. R. Freimund 1994. Using monthly precipitation data to estimate the R–factor in the revised USLE. *J. Hydrol.* 157: 287–306.

Renard, K. G., G. R. Foster, G. A. Weesies, D. K. McCool, and D. C. Yoder, coordinators. 1997. *Predicting Soil Erosion by Water: A Guide to Conservation Planning with the Revised Universal Soil Loss Equation (RUSLE).* Agriculture Handbook No. 703. Washington, D.C.: USDA.

Risse, L. M., B. Y. Liu, and M. A. Nearing. 1995. Using curve numbers to determine baseline values of Green–Ampt effective hydraulic conductivities. *Water Resour. Bull.* 31(1): 147–158.

Tiwari, A. K., L. M. Risse, and M. A. Nearing. 2000. Evaluation of WEPP and its comparison with USLE and RUSLE. *Trans. ASAE* 43(5): 1129–1135.

Trewartha, G. T. 1981. *The Earth's Problem Climates.* 2nd ed. Madison, Wisc.: University of Wisconsin Press.

Yu, B. 1998. Rainfall erosivity and its estimation for Australia's tropics. *Aust. J. Soil Res.* 36(1): 143–165.

_____. 1999. A comparison of the R–factor in the Universal Soil Loss Equation and Revised Universal Soil Loss Equation. *Trans. ASAE* 45(6): 1615–1620.

_____. 2000. Improvement and evaluation of CLIGEN for storm generation. *Trans. ASAE* 46(2): 301–307.

Yu, B., and C. J. Rosewell. 1996. An assessment of a daily rainfall erosivity model for New South Wales. *Aust. J. Soil Res.* 34(1): 139–152.

Yu, B., C. A. A. Ciesiolka, C. W. Rose, and K. J. Coughlan. 2000. A validation test of WEPP to predict runoff and soil loss from a pineapple farm on a sandy soil in subtropical Queensland, Australia. *Aust. J. Soil Res.* 38(5): 537–554.

Yu, B., G. M. Hashim, and Z. Eusof. 2001. Estimating the R–factor using limited rainfall data: A case study from Peninsular Malaysia *J. Soil and Water Conserv.* 56(2): 101–105.

Zhang, X. C., M. A. Nearing, and L. M. Risse. 1995a. Estimation of Green–Ampt conductivity parameters: Part I. Row crops. *Trans. ASAE* 38(4): 1069–1077.

_____. 1995b. Estimation of Green–Ampt conductivity parameters: Part II. Perennial crops. *Trans. ASAE* 38(4): 1079–1087.

Zhang, X. C., M. A. Nearing, L. M. Risse, and K. C. McGregor. 1996. Evaluation of WEPP runoff and soil loss predictions using natural runoff plot data. *Trans. ASAE* 39(3): 855–863.