5.7 IMPLEMENTING QUALITY CONTROL TECHNIQUES FOR RANDOM NUMBER GENERATORS TO IMPROVE STOCHASTIC WEATHER GENERATORS: THE CLIGEN EXPERIENCE.

Charles R. Meyer*
USDA-Agricultural Research Service
West Lafayette, Indiana

Chris Renschler
Department of Geography, University at Buffalo
Buffalo, New York

Roel C. Vining
USDA-Natural Resource Conservation Service
West Lafayette, Indiana

1. INTRODUCTION

The CLIGEN stochastic weather generator (Nicks et al. 1995) produces time series of daily weather parameters from monthly values observed at the site, like monthly mean, standard deviation (SD), and skewness. This approach permits generation of representative weather patterns, for selectable time intervals, from a relatively small amount of input data. Proper functioning of the climatic equations in weather generators absolutely depends upon receiving numeric inputs which adhere to the specified distributions that the equations are designed for. It is also desirable that the series of numbers used have an element of chance or randomness; i.e., that they not be related to each other. For this reason random number generators are used to produce them. This paper will demonstrate that in stochastic modeling, we have typically gotten our priorities reversed, emphasizing randomness more than achieving the required distribution. We also present a solution to the dilemma which is relatively simple to implement.

The reader may well ask what all the fuss is about. Aren't our standard random number generators supposed to produce a uniform distribution? The problem lies in the length of run required to achieve this. Simple tests on uniform random number generators show that 30,000 to 60,000 numbers must be generated to yield a population approaching the one desired. That is equivalent to a 1,000 to 2,000 year run with a model using monthly parameters like CLIGEN. Using the technique outlined in this paper, we have consistently been able to converge upon the historic values within 30 years.

Other stochastic models which probably are impacted by this effect include GEM (Johnson 2001), WINDGEN (Wagner 1999), SWAT (Arnold et al. 1995), USCLIMATE (Hanson et al. 1994), SWRRB (Arnold and Williams 1994), GLEAMS (Knisel 1993), EPIC (Sharpley and Williams 1990), WGEN (Richardson and Wright 1984), and CREAMS (Knisel 1980).

HOW CLIGEN WORKS:

In CLIGEN, each of eight parameter distributions (maximum temperature, minimum temperature, solar radiation, precipitation, wind direction, wind velocity, temperature dew point, and time to peak intensity) is generated by feeding the output from a uniform pseudo Random Number Generator (RNG) (Fig. 1a) into a Standard Normal deviate Generator (SNG) (Fig. 1b), to produce sets of "deviates" that approximate the Standard Normal (SN) distribution (mean = 0 and Standard Deviation (SD) = 1).

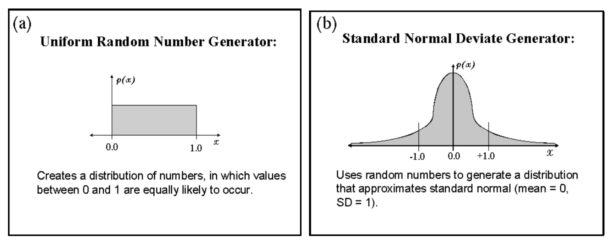

Figure 1

In CLIGEN the values of each SN distribution are then scaled by the corresponding observed historical monthly mean, SD, etc. to produce daily weather values. Figure 2 illustrates this process for the monthly temperature distribution, which converts a *standard normal* distribution (Fig. 2a) to the temperature distribution (Fig.2b) as indicated by the illustration below.

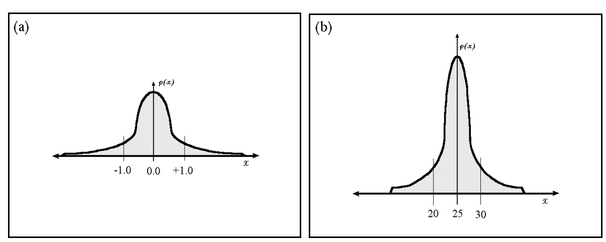

Figure 2

---

* Corresponding author address: Charles R. Meyer, 1196 Bldg. SOIL, Purdue Univ., West Lafayette, IN 47907; e-mail: meyerc@ecn.purdue.edu

As illustrated in Figure 3 the quality of the daily weather values (Step 3 -- Circled numbers) produced by CLIGEN depend directly upon the quality of the distributions produced by the RNG (Step 1) and SNG (Step 2). Lack of quality assurance for these distributions has potentially serious implications for CLIGEN and simulation models (like WEPP and WEPS) using its outputs to produce their own estimates (Step 4). As observed in Numerical Recipes (2000) "... a reliable source of random uniform deviates, ... is an essential building block for any sort of stochastic modeling".

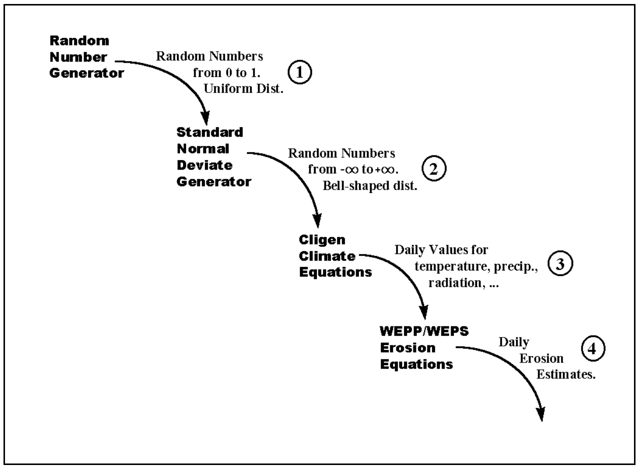

Figure 3

Johnson et al. (1996) observed that while CLIGEN reproduced historical long-term average values for the simulation period reasonably well, it reproduced year-to-year variance in average annual values poorly, and did a dismal job of reproducing the monthly standard deviations, failing all 72 tests performed. If the problem with CLIGEN were due to the scientific climate equations used, the historic means probably would not have been successfully reproduced; however, since the problem seems to involve the distribution of the outputs relative to the mean, it seems more likely that the distribution of random numbers fed to the equations is suspect. A RNG should produce sets of numbers having a uniform distribution; however, our tests on the RNG in CLIGEN showed serious problems, as did our tests on the numbers subsequently produced by the SNG. Our tests on the RNG from GEM showed similar deficiencies.

2. METHODS

The ***Chi-square test*** can be used to determine the probability that a set of numbers fits a specified distribution. This is done by ranking the numbers, sorting them into specified intervals, and comparing mathematically to the number expected for each interval.

The Central Limit Theorem of statistics states that means of samples approach a normal distribution as the number of samples becomes large, regardless of the underlying distribution (Ross 1993). This powerful theorem justifies ***confidence interval (CI) testing*** on sample means, and even on the sample SD's if the underlying distribution is known.

2.1. TESTING THE QUALITY OF THE UNIFORM DISTRIBUTIONS GENERATED:

Thus, it is possible using the Chi-square test to measure the probability that the RNG is producing the uniform distributions required (Fig. 3, Step 1). Because CLIGEN generates its outputs from monthly parameters, the random numbers produced by Cligen for each parameter, for each month of the year, were subjected to a Chi-square test to measure the probability that they were NOT from a uniform distribution. We made runs for 5, 10, 30, and 100 years. One would expect the results from a 100 year run to be quite good. They were not (Table 1). GEM's RAN3 fared badly as well.

Table 1

100 yr run. Probability distribution is NOT uniform. Chi-square test.

Parameter:	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
MaxTemp	.333	.365	.230	.413	**.535**	.150	.193	.023	.179	.101	.132	.244
MinTemp	.406	.491	**.634**	.207	.123	**.780**	**.818**	.395	.477	**.679**	.310	**.916**
Radiation	.358	**.637**	.468	**.897**	**.726**	**.552**	**.517**	.348	**.891**	**.903**	**.649**	.393
PrecipAmt	**.755**	.408	**.865**	**.581**	.351	.067	.248	.197	**.831**	**.860**	**.964**	.017
WindVel	**.706**	.250	.132	.410	.410	**.862**	**.930**	**.964**	**.944**	**.923**	**.679**	**.919**
TDP	**.713**	.497	.437	.491	.471	**.705**	**.515**	**.536**	**.669**	**.693**	**.536**	.185
PrecipPrb	**.777**	**.838**	**.939**	**.886**	**.810**	.048	.109	.232	.165	.167	**.550**	.204
DSTG	**.924**	.307	.247	.143	.216	.166	.163	.225	.223	.381	**.627**	.175
WindDir	.500	.340	**.578**	.072	.175	.119	.280	.090	.015	.063	.320	.100
TimeToPk	.499	**.689**	**.796**	.355	.332	**.620**	**.532**	**.502**	**.632**	**.629**	**.823**	**.787**

2.2. CONTROLLING QUALITY OF STANDARD NORMAL DISTRIBUTIONS AS THEY ARE GENERATED:

The confidence interval approach is commonly used in manufacturing to ensure that goods meet the desired production quality standards. If CLIGEN's RNG and SNG together are thought of as a "factory" producing random numbers, having mean = 0, SD = 1, for CLIGEN to consume, the quality of the distribution can be both measured and controlled using confidence interval testing on the mean and the SD.

CLIGEN V-5.101 was revised to incorporate confidence interval software to test and control the numbers produced by the SDG (Fig. 3, Step 2) for use by CLIGEN's equations (Meyer 2001). This is carried forward to all subsequent versions, which we refer to here collectively as CLIGEN V-5.x. We arbitrarily enforce a 50% probability they are SN. This extra control should help ensure that the numbers we get out of CLIGEN much more closely match the parameters we fed in. Table 3 compares the annual means of daily precipitation values from four climatically diverse U. S. sites (listed in Table 2), to the values from which CLIGEN V-4.2 and V-5.x supposedly computed them (in bold type). The generated weather data were subsequently used to compare results from the Water Erosion Prediction Project (WEPP) soil erosion model (Flanagan and Nearing, 1995). The same hillslope topography, soils and land use management were used for the WEPP simulations in all four locations (Table 3).

Table 2

Climatic Classification of Sites Used.

Station, State	Abbreviation	Precipitation	Temperature	Köppen Climate Classification*
Indianapolis, IN	IND	Uniform	Warm Summer	Cfb
College Station, TX	COS	Uniform	Hot Summer	Cfa
Moscow, ID	MOS	Dry	Warm Summer	Dsb
Tucson, AZ	TUC	Dry	Hot Summer	Csa

* (Griffiths and Driscoll 1982)

Table 3

Summary of 30 year CLIGEN precipitation generation and WEPP simulation.

Station	CLIGEN Version / Observed	Avg. Annual Precipitation [mm]	Avg. Annual Runoff [mm]	Avg. Annual Sediment Yield [t ha^{-1}]
IND	V-4.2	**1034.9**	307.6	65.6
	V-5.x	**1011.9**	265.4	60.8
	Observed	**1013.1**	n.a.	n.a.
COS	V-4.2	**1040.4**	307.6	214.9
	V-5.x	**987.7**	265.4	174.1
	Observed	**959.5**	n.a.	n.a.
MOS	V-4.2	**643.0**	21.7	11.3
	V-5.x	**644.4**	21.4	8.4
	Observed	**621.5**	n.a.	n.a.
TUC	V-4.2	**310.1**	25.5	21.9
	V-5.x	**291.1**	21.2	19.4
	Observed	**293.2**	n.a.	n.a.

n.a. = not available.

2.3. CLIGEN V–4.2

Using CI tests we observed that the monthly means of the daily values (Fig. 3, Step 3) generated for maximum temperature (Fig. 4a) and precipitation (Fig. 4b) fell outside the expected limits an abnormally high percentage of the time.

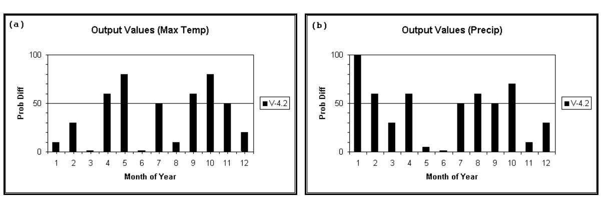

Figure 4

Because the effect of a single aberrant number should be diluted by adding more “good” numbers to the pool, when a mean goes out of range one would expect that extending the duration of the run might cause the mean to return to the expected range. Runs were made increasing the duration by one-year steps from 1 to 100 years, and then in 100 year steps to 1,000 years. Generally, once the runs went outside the confidence interval , extending the length of simulation did not correct the problem.

2.4. QUALITY CONTROL IN CLIGEN V-5.X:

In CLIGEN V-5.x, standard normal numbers are generated for each parameter a month at a time, and subjected to CI tests. If the probability that the current set of numbers is not within specifications exceeds 50%, for either mean or SD, they are rejected and a new set is generated for testing.

Controlling the means and SD’s of CLIGEN V-5.x’s standard normal deviates proved sufficient to bring the CLIGEN outputs within the expected bounds for all parameters that were tested (maximum temperature, minimum temperature, precipitation, and solar radiation), <u>if range checks and other mechanisms that alter the distributions after generation were disabled</u>. For example, minimum temperature is forced to be less than maximum temperature for each day, and negative precipitation is not allowed. Constraints are also imposed upon radiation values after they are generated.

3. RESULTS

Results shown here for College Station, TX, are typical of all other locations tested. “Prob Diff” in the following figures indicates the probability that the population of numbers output differs from the target population.

3.1. Comparison of CLIGEN V-4.2 and CLIGEN V-5.x

The comparison of the means of the two CLIGEN versions’ generated standard normal deviates (Fig 3, Step 2), to standard normal, is shown for maximum temperature in Fig. 5a. Fig. 5b compares the means of the historical statistical parameters input and the daily maximum temperatures output (Fig. 3, Step 3) from CLIGEN. CLIGEN V-5.x brings the standard normal deviates into control, and the outputs also come into control.

Maximum temperature exactly follows the process of scaling the standard normal deviates outlined in Fig 2. Because no limits are applied after generation, which could alter the resulting distribution, one might expect Figures 5a and 5b to look identical. However, that is not exactly the case, because the temperature calculations utilize a SD computed from the observed data .

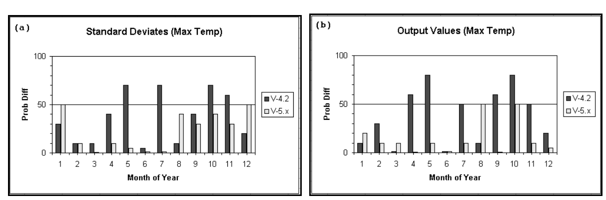

Figure 5

Fig. 6a compares the means of the standard normal deviates to standard normal for precipitation for both CLIGEN versions. Fig. 6b compares the means of the statistical parameters input to the daily precipitation output. CLIGEN V-5.x brings the standard normal deviates into control, and the outputs also come into control.

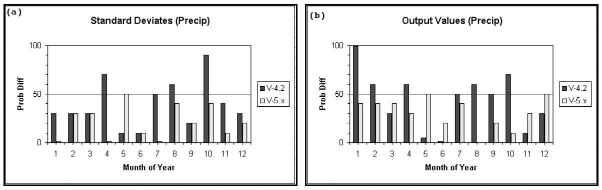

Figure 6

Fig. 7a compares the means of the standard normal deviates to standard normal for minimum temperature, and Fig. 7b compares the means of the statistical parameters input to the daily minimum temperature output. CLIGEN V-5.x brings the standard normal deviates into control, but the outputs do not come into control. The output values for minimum temperature show seasonality between winter and summer months with winter months being more likely out of control.

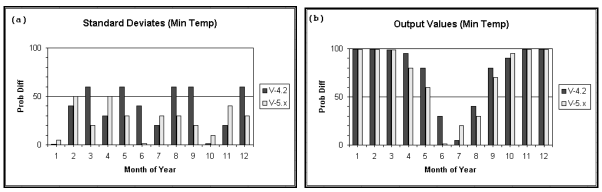

Figure 7

3.2. Effect of CLIGEN versions on WEPP erosion model results

Since CLIGEN's output reports the historic monthly means for the current site, it is easy to determine whether the version with quality control (QC) is reproducing the historical recorded mean annual precipitation better than the original version. The results of 30-year water erosion model runs summarized in Table 3 show that when CLIGEN V-4.2 differed significantly from the QC version CLIGEN V-5.x, the QC was closer to the observed average annual precipitation. Also note the <u>similarity</u> of average annual precipitation and runoff amounts for Moscow, ID (MOS), and the <u>difference</u> between the amounts of average annual sediment yield. This underscores the effect of accurately reproducing the historical <u>distributions</u>, not just the <u>total amounts</u>.

4. DISCUSSION

One potential problem with this quality control approach is that it may skew the results in <u>time</u>. The numbers for a given parameter and month of the year are generated and tested in combination with those previously accepted, as each new year is simulated. There are two offsetting factors involved in the test: 1) When very few numbers are involved, the confidence interval is very wide. There is not enough information to constrain the interval much. 2) As more numbers are added to the distribution tested, the confidence limits get tighter, but adding a really aberrant mean (or SD) does not affect the mean of the overall group very much. Currently it is not known how these offsetting factors play out against each other in the early stages of the run. It is conceivable that a divergent number might be rejected early in the run, but accepted later, once the system has enough "mathematical inertia" not to be grossly affected by it. More investigation is needed to examine this. Of course if it is determined that the first 5 years of a 30 year run are biased, one can simply make a 35 year run and discard the first 5 years.

5. CONCLUSIONS

The results of stochastic programs are sensitive to the quality of their randomly generated <u>distributions</u>. While randomness is a desirable characteristic because it mimics the conditions we assume to exist in the physical world, it is of secondary importance to the actual distribution in assuring that the equations in our models perform the work they were designed to do. Fortunately as shown in this research, this problem is relatively easy to remedy without major changes to the existing program structure. The quality control method and source code is publicly available on the CLIGEN website of the National Soil Erosion Research Laboratory (Meyer, 2001).

Acknowledgements
We would like to express our deep appreciation to Bofu Yu of Griffith University, Australia, and Jurgen Garbrecht with ARS, El Reno, OK, for many profitable exchanges that deepened our insight into the problem and its solution.

References
Atmospheric Radiation Measurement Program, 2001. *Pyranometer Information*. URL: http://www.arm.gov/docs/instruments/static/pyra.html 04/12/2001.

Arnold, J. G., and J. R. Williams, 1994. SWRRB; a watershed scale model for soil and water resource management, Grassland Soil and Water research Laboratory, U. S. Department of Agriculture-Agricultural Research Service, Temple, TX.

Arnold, J. G., J. R. Williams, R. Srinivasan, and K. W. King, 1995. SWAT: soil and water assessment tool, U. S. Department of Agriculture-Agricultural Research Service Grassland Soil and Water research Laboratory, Temple, TX.

Flanagan, D. C. and M. A. Nearing (eds.), 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 47907, 292 pp.

Griffiths, J. F., and D. M. Driscoll, 1982. *Survey of Climatology*, Charles E. Merrill Publishing Co., Columbus, OH. 358 pp.

Hanson, C.L., K.A. Cumming, D.A. Woolhiser and C.W. Richardson, 1994. Microcomputer program for daily weather simulation. U.S. Dept. Agric., Agric. Res. Svc. Pub. No. ARS-114, 38 pp.

Johnson, G. L., C. L. Hanson, S. P. Hardegree, and E. B. Ballard, 1996. Stochastic Weather Simulation: Overview and Analysis of Two Commonly Used Models. *Journal of Applied Meteorology*, 35, 1878 - 1896.

Johnson, G. L., 2001. GEM webpage. URL: http://www.wcc.nrcs.usda.gov/water/climate/gem/gem.html

King, D. L, and D. R. Myers, 1997. *Silicon-Photodiode Pyranometers: Operational Characteristics, Historical Experiences, and New Calibration Procedures*, presented at the 26th IEEE Photovoltaic Specialists Conference, Anaheim California, 5pp.

Knisel, W.G., 1980. CREAMS, a field scale model for chemicals, runoff and erosion from agricultural management systems. USDA Conservation Research Report. No. 26.

Knisel, W. G., ed., 1993. GLEAMS: groundwater loading effects of agricultural management systems, U. S. Department of Agriculture-Agricultural Research Service Southeast Watershed Research Laboratory, Tifton, GA.

Meyer, C. R., 2001. *CLIGEN Weather Generator, Expanded and Improved*.URL: http://horizon.nserl.purdue.edu/Cligen/. 01/11/2001.

Nicks, A. D., L. J. Lane and G. A. Gander, 1995. Chapter 2. Weather generator. *USDA-Water Erosion Prediction Project: Hillslope Profile and Watershed Model Documentation, NSERL Report # 10*, D.C. Flanagan and M.A. Nearing, Eds., NSERL West Lafayette, IN: USDA-ARS National Soil Erosion Research Laboratory, 2.1 – 2.22.

Numerical Recipes, 2000. In: *Fortran the Art of Scientific Computing, 2nd ed.*, Cambridge University Press, 973 pp. URL: http://lib-www.lanl.gov/numerical/bookfpdf.html 02/26/2001.

Richardson, C.W. and D.A. Wright, 1984. WGEN: A model for generating daily weather variables. USDA Agricultural Research Service, ARS-8, 83 p.

Ross, S., 1993. *Introduction to probability models, 5th ed.*, Academic Press, 556 pp.

Sharpley, A. N. and J. R. Williams, eds., 1990. EPIC -- Erosion/Productivity Impact Calculator, USDA-ARS Technical Bulletin 1768, USDA-ARS Grassland, Soil, and Water Research Lab, Temple, TX. URL: http://www.brc.tamus.edu/epic/index.html 05/09/2001.

Wagner L. E., 1999. Wind Erosion Prediction System (WEPS): Overview. URL: http://www.weru.ksu.edu/symposium/proceedings/wagner2.pdf, 6/26/2001.