GEM MODEL TEMPERATURE AND PRECIPITATION PARAMETER VARIABILITY, AND DISTRIBUTION USING PRISM

Gregory L. Johnson
USDA-NRCS, National Water and Climate Center, Portland, Oregon

Christopher Daly and George H. Taylor
Oregon State University, Corvallis, Oregon

Clayton L. Hanson and Yunyun Lu
USDA-ARS, Northwest Watershed Research Center, Boise, Idaho

1. INTRODUCTION

Multivariate stochastic models of weather and climate have traditionally been point models. Model parameters are derived from available climatological data at a location, and the generated time series from the model is typically applied to some small region around this station where the climate is assumed to be essentially the same.

In practice, climate is rarely spatially homogeneous, rendering the generated time series inappropriate for application to places even a few km distant (in an extreme case). The usefulness of the stochastic model, and its applicability, are thus greatly reduced. In fact, one of the reasons chiefly touted for using generated time series is that they can be produced for locations with no available climate data (Richardson and Wright, 1984; Woolhiser et al., 1988; Nicks and Gander, 1994; Hanson et al., 1994). However, as climate increases in spatial complexity it becomes less likely that parameters for a specific location will be applicable to any surrounding location.

There is thus motivation for the development of some type of parameter interpolation method which is sensitive to the major forcing mechanisms producing spatial variability; namely, topography (elevation, aspect, slope), scale, and proximity to larger water bodies. With such an interpolation procedure established it then becomes possible to generate time series at any number of locations, regardless of the availability of climate data.

2. MODELS USED

To develop and test these ideas a stochastic weather generator model and an "intelligent" interpolation model were needed. The weather generator model chosen was GEM (Generation of weather Elements for Multiple applications), developed by researchers with the USDA-Agricultural Research Service, and previously known as WGEN (Richardson and Wright, 1984) and USCLIMATE (Hanson et al., 1994). The interpolation model used was PRISM (Parameter-elevation Regressions on Independent Slopes Model), developed by researchers at Oregon State University (Daly et al., 1994).

GEM was chosen because of its proven ability to successfully replicate most aspects of the real climate, including the preservation of serial and cross-element correlations (Johnson et al., 1996). GEM produces a daily time series of precipitation, maximum and minimum air temperature, solar radiation, mean daily dewpoint and wind speed. It utilizes a two-state Markov chain of first order for the representation of precipitation occurrence, and all other generated quantities are dependent on whether a given day is wet or dry. Precipitation amounts are drawn from a mixed-exponential distribution. Temperature and other elements are generated based on daily mean and standard deviation values, the previous day's value and the correlation with other elements.

PRISM has been successfully used to map mean monthly and annual precipitation, and, more recently, temperature, frost dates, growing season lengths and snow water equivalent (Daly et al., 1997). It is an expert system that uses point data and a digital elevation model (DEM) to generate gridded estimates of climate elements. PRISM was developed to successfully estimate climate in regions where topography and other factors produce significant variability in climate, such as mountainous regions. The effects of terrain on climate play a central role in the model's conceptual framework.

3. METHODOLOGY

The concept proposed and tested here seeks to examine the spatial variability of various GEM parameters and then use PRISM to produce gridded fields of these parameters. The methodology also seeks to link these PRISM-produced layers so that a user can select a location, extract all necessary GEM parameters for that location and ingest them into the model, and then generate a desired-length time series

* *Corresponding author address:* Greg Johnson, USDA-NRCS, 101 SW Main St., Suite 1600, Portland OR 97204-3224.
Email: gjohnson@storm.wcc.nrcs.usda.gov

of weather, all with a simple point-and-click configuration.

The methodology was tested with a group of 80 climate stations in southern Idaho and southeastern Oregon (Figure 1). Approximately 60 were NOAA Cooperative stations with generally complete 1961-1993 data. An additional 17 USDA-Natural Resources Conservation Service (NRCS) SNOTEL stations and three USDA-Agricultural Reserach Service (ARS) stations at the Reynolds Creek Experimental Watershed were used to provide data from high elevation regions. The SNOTEL record lengths were much shorter, and the stations utilize a coarser increment for reporting precipitation (2.5 mm); however, procedures were developed to make the SNOTEL parameters compatible and consistent with the NOAA and ARS stations.

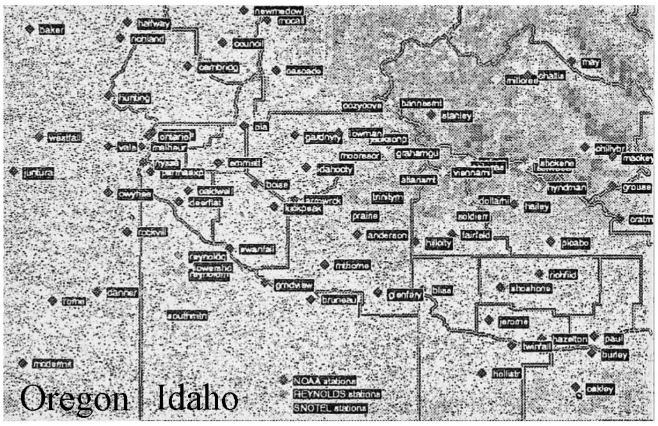

Figure 1. Location of 80 stations used in study over S. Idaho and S.E. Oregon.

Precipitation and maximum and minimum air temperature (Tmax, Tmin) were the elements chosen for testing, since these three elements were recorded at all of these sites. A total of 49 precipitation and 9 temperature parameters were necessary for proper generation of GEM time series; thus, a total of 58 PRISM-produced parameter layers were developed using these procedures. Precipitation occurrence parameters used in GEM were p00 and p10, representing the transition probability (between 0 and 1) of a dry day following a dry day and a wet day, respectively. Parameters used for generating the amount of precipitation on any given day, taken from the mixed-exponential distribution, were μ, the mean amount of precipitation on a wet day; β, the mean value of the smaller exponential distribution; and α, a weighting parameter used to define the relative contribution of each of the smaller and larger exponential distributions, with values between 0 and 1. Previous versions of the model utilized an optimization routine to select only significant annual harmonics and Fourier coefficients, but spatial variability of these was found to be complex. Instead, discrete, monthly parameter values were derived, analyzed and interpolated using PRISM for all parameters except α, which showed no significant seasonal changes.

Parameters used for generating temperature values were the annual mean and amplitude values of Tmax and Tmin, and their associated coefficients of variation. Earlier studies have shown that it is important to parameterize the annual mean Tmax differently for wet and dry days ($Tmax_w$, $Tmax_d$) and this convention was followed here.

In order to provide guidance for PRISM interpolation of parameters, a thorough analysis of the spatial variability of GEM parameters and their temporal characteristics was first conducted. Graphs of domainwide (all 80 stations) monthly and annual parameters versus elevation were prepared. Parameters were plotted on maps to analyze spatial patterns. Then, PRISM was used to distribute these parameter values to grid points. Fifty-eight parameter maps were derived from the PRISM interpolation of the 80 station values. Each map consisted of 7957 pixel values. Statistics summarizing PRISM performance over the entire domain were then calculated and compared to the domain-wide, average statistics derived from the raw station values.

4. REGIONAL TEST RESULTS

Region-wide parameter/elevation graphs revealed that elevational gradients were large for many parameters, small for others, and typically had seasonal dependencies. Mean regression slopes from PRISM were consistent with these region-wide elevational gradients. The study domain is dominated by strong westerly flow with frequent intrusions of Pacific moisture during the winter. Summers are typically warm and very dry, with only occasional storms. More than 80% of total annual precipitation falls during the six month winter period at highest elevations, where most of it comes in the form of snow. At lower elevations precipitation is more evenly distributed throughout the year.

It was hypothesized that GEM parameters describing the occurrence and amount processes of precipitation would show a greater elevational dependence in the winter. This was true for precipitation amounts (μ), but not so for occurrence parameters (p00 and p10). Thus, over this study domain a much greater amount of precipitation falls on a wet day at higher elevations in the winter. This is the dominating factor in enhancing mean total precipitation at higher elevation locations, and not the occurrence process. However, there seem to be significant local dependencies which are not accounted for in this domainwide analysis. Thus, a more detailed study of these local effects, and parameter spatial variability, was warranted. These analyses also confirmed the need for a more sophisticated interpolation procedure, such as PRISM, rather than relying on a single elevational relationship, even for a domain as relatively small as this one.

For illustration purposes, the month of December was examined. There was significantly less spatial variability in p00 than most of the other parameters, with a range between highest and lowest values across the domain of just 0.17. Highest values were found in the driest locations, as expected. Only a moderate dependency on elevation was noted, even in local areas like the ARS Reynolds Creek Watershed where the three stations are close to one another (within 15 km), yet a significant change in elevation (1200 to 2100 m) and mean annual precipitation (300 to 900 mm) is observed.

December p10, μ and β had significantly more spatial variability and the local elevation relationship was stronger than for p00. The range of observed p10 values over the domain was nearly 0.50. At Reynolds Creek, p10 decreased by 0.14 from lowest to highest elevations, and in the eastern section of the study area, p10 decreased more than 0.30 over distances of less than 40 km and elevation changes of between 1200 and 1800 m. Significant, local elevation dependencies were noted domainwide, with p10 decreasing in all cases (increasing persistence of wet days with elevation).

Mu, the average precipitation on a wet day, had similar local elevation relationships. At Reynolds Creek, average December μ increased from 2.8 to 7.4 mm/day with a 900 m increase in elevation, or approximately 5 mm/km. In eastern sections of the domain the μ/elevation gradient was even more significant, especially on some leeward slopes. Beta was a much more difficult parameter to interpret because of the complex and less-than-straightforward manner in which it changed spatially. In general, β increased with local elevation, but the change was quite different from one region to another. At Reynolds Creek, December β increased just 1 mm in 900 m of elevation increase. In eastern sections of the domain, β increased at rates between 10 and 30 mm/km. Alpha had statistically insignificant change through the year at each location. Thus, a single, annual value of α was calculated and used at each location. The map of α (not shown) revealed very little spatial variability and no significant elevation relationship.

PRISM regression slopes of precipitation occurrence parameters p00 and p10 were generally smallest in the summer (Table 1). P00 had steepest lapse rates in the winter and spring, while p10 had two maxima corresponding to the transition months (February through April and September and October). Greatest predictability, as denoted by r^2, was in the summer. The largest individual station residuals (predicted minus observed value) were in the late fall through the early spring. In general, the PRISM-derived regressions correctly identified the important local elevational gradients which were maximized in the winter, and were consistent with the domainwide predictability of p10 from elevation alone.

TABLE 1. Summarized PRISM interpolation statistics of monthly and annual GEM precipitation and temperature parameters.

Month	P00 Slope (/km)	P00 r-square	P10 Slope (/km)	P10 r-square
Jan	-.041	.15	-.124	.17
Feb	-.043	.18	-.150	.22
Mar	-.047	.18	-.163	.21
Apr	-.049	.21	-.143	.24
May	-.049	.30	-.121	.28
Jun	-.047	.28	-.119	.28
Jul	-.039	.26	-.111	.22
Aug	-.032	.23	-.128	.25
Sep	-.028	.20	-.149	.30
Oct	-.035	.30	-.146	.23
Nov	-.043	.17	-.136	.16
Dec	-.042	.13	-.126	.15

Month	Mu Slope (/km)	Mu r-square	Beta Slope (/km)	Beta r-square
Jan	.125	.25	.100	.18
Feb	.114	.26	.096	.20
Mar	.092	.27	.071	.18
Apr	.071	.28	.063	.20
May	.069	.25	.065	.20
Jun	.061	.22	.058	.20
Jul	.061	.18	.048	.16
Aug	.062	.15	.052	.16
Sep	.069	.22	.064	.18
Oct	.090	.27	.086	.18
Nov	.113	.25	.096	.19
Dec	.122	.23	.099	.18

Parameter	Vertical Layer #	Mean Ann. P/E Slope (/km)	r-square
Mean Tmaxd	--	-5.33 C	.62
Amp. Tmax	--	-2.45 C	.44
Mean CV Tmax	_	.001	.19
Amp. CV Tmax	--	-.00095	.16
Mean Tmaxw	--	-6.34 C	.70
Mean Tmin	1	-4.09 C	.29
	2	-5.14 C	.35
Amp. Tmin	1	-.31	.11
	2	-.71	.27
Mean CV Tmin	1	.00015	.11
	2	.00022	.15
Amp. CV Tmin	--	-.00012	.05

Mean local regression slopes from PRISM-interpolated monthly values of μ and β were also greatest (most negative) in the winter, and smallest in the summer. In contrast to p00 and p10, predictability was least in the summer and nearly the same in all other months.

Over the entire domain some of the nine temperature parameters were clearly related to elevation, while others were poorly related. Local elevational relationships were, in general, stronger for temperature parameters than for precipitation parameters. Mean maximum temperature, separately examined for dry and wet days, had the most linear elevation dependence, with domainwide r^2 values near 0.9. PRISM-interpolated average r^2 values for $Tmax_d$ and $Tmax_w$ were 0.62 and 0.70, respectively. These values are very high considering that they represent the mean r^2 of thousands of regressions (one for each grid cell), each of which uses approximately 10 to 20 weighted station values.

Mean maximum temperatures were cooler on wet days than on dry days, and this difference was more pronounced with increasing elevation. Average lapse rates were approximately 4.8 C/km, domainwide, on dry days, and approximately 6.8 C/km on wet days. Average local PRISM lapse rates were similar to these. The largest $Tmax_d$ and $Tmax_w$ station residuals were around 3 C. The seasonal variation of mean Tmax (Amp. Tmax) decreased with elevation domainwide (r^2=0.55), and even more so on a local basis. Amp. Tmax was largest in locally-low elevation locations and across the Snake River Plain, and smallest at exposed locations near or on ridges. Overall, the decrease in Amp. Tmax with elevation was approximately 2 C/km, but locally ranged from 1 to 4 C/km. The mean slope of all PRISM regressions was 2.45 C/km. This means that on a local basis the average difference in the mean maximum temperature between winter and summer decreased at a rate of nearly 2.5 C every km. Thus, there was a relatively greater seasonal change in the maximum temperature of valleys versus mountaintop locations. For all resolvable cells (73.6%), the average r^2 was 0.44, and the largest station residual was 2.4 C.

The seasonality of interdaily variability (Amp. CV Tmax) was more weakly related to elevation (r^2=0.10 for the whole domain), and local elevation dependence (from PRISM averages) was only slightly stronger. Amp. CV Tmax also had a fairly complex spatial pattern. The mean coefficient of variation of maximum temperature had very little spatial variability and was weakly linked to elevation, both domainwide (r^2=0.24) and locally. Smallest mean CV Tmax values were at lowest elevation and westernmost locations.

Microclimatic differences were more evident in mean minimum temperatures with a domainwide elevational regression r^2 of 0.62. In general, minimum temperature decreased with elevation domainwide, but locally the relationship was anything but straightforward. For instance, at Reynolds Creek, the middle elevation site had the warmest mean minimum, at 3.4 C, due to frequent downslope warming, while the lowest elevation site was nearly as cold as the mountaintop site (1.6 C vs. 0.5 C), despite being 900 m lower in elevation. The coldest location in the domain was Stanley (-7.6 C), a dry, high elevation valley site, which was 1 to 5 C colder than all locations in the immediate vicinity, many of which were 200 to 800 m higher in elevation.

The interpolation of Tmin was more difficult due to the greater microclimatic effects which are typically reflected in parameters associated with Tmin. One of the larger-scale phenomenon known to affect Tmin was a persistent temperature inversion which dominated valley locations in winter. To simulate this, a utility in PRISM was used that allows climate stations to be divided into two vertical layers, with regressions done on each separately. Layer 1 represented the boundary layer, and layer 2 the free atmosphere. The thickness of the boundary layer was prescribed to reflect the height of the mean wintertime inversion height over Boise, ID. The elevation of the top of the boundary layer was spatially distributed to a grid by using the elevation of the lowest DEM pixels in the vicinity as a base, and adding the inversion height to this elevation. As a result, large valleys tended to fall within this layer, while local ridgetops and other elevated terrain jutted into the free atmosphere.

To accommodate the spatially and temporally varying strength of the inversion, PRISM was designed to allow varying amounts of "crosstalk" (sharing of data points) between the vertical layer regressions depending on the similarity of the regression functions. Under strong inversion conditions in winter, the regression functions would be very different, and crosstalk would be minimized. During summer and in well-mixed locations, the regression functions would show similar characeristics, and stations would be shared more freely across the layer 1/layer 2 boundary.

The mean elevational regression slope (Table 1) of Tmin in the lower layer was -4.09 C/km and, as expected, was more negative in layer 2 (-5.14 C/km). The largest station residual was 4.4 C. Thus, the inclusion of a two-layered approach to Tmin interpolation certainly improved the results over approaches without consideration of a mean inversion top, but other factors beyond the scope of this work would have to be included in order to more accurately replicate Tmin in all locations.

The mean value of the amplitude of the first harmonic of Tmin had a significantly smaller lapse rate for both PRISM layers 1 (-0.31 C/km) and 2 (-0.71 C/km) than the average lapse rate of the amplitude of Tmax, and was consistent with domainwide values. This was interpreted to mean that locally throughout the domain there was comparatively little elevational gradient in the seasonal change of Tmin.

Mean CV of Tmin was quite uniform, with no appreciable elevational dependence (domainwide r^2=0.09; mean of PRISM layers r^2=0.13). Slopes of domainwide and PRISM regressions also were negligible. Largest values were at moderately high elevation, dry valleys like the Stanley Basin which normally have the greatest variability about the mean minimum temperature. The seasonal variation of this parameter, represented by Amp. CV Tmin, was closely tied to the mean CV, with largest absolute values also at Stanley and other moderately high valleys and plateau locations. Domainwide, elevational dependence was near zero; however, on a local basis, there was a slightly greater relationship.

Bias and mean absolute error (MAE) jackknife cross-validation statistics for all 58 interpolated parameters were computed and given as both actual values and as percentages (not shown). These statistics confirmed many of the hypotheses discussed above. Best overall model performance based on these four criteria was for p00, mean $Tmax_d$ and $Tmax_w$, mean Tmin, and Amp. Tmax. Largest error statistics were noted for those parameters that were discussed above as being difficult to interpolate, often with no clear elevational or geographic dependencies. These included β, Amp. CV Tmax and Amp. CV Tmin. PRISM-interpolated p00 and p10 values were consistently negatively-biased; i.e., the model tended to underpredict these values, but the amount of underprediction was very small.

4. CONCLUSIONS

In an effort to generate climate time series at locations with no recorded climate history, and in regions with spatially non-homogeneous climate, a study was undertaken to examine the spatial variability of the parameters of a stochastic weather generator model and to interpolate these to a high resolution (2 km) grid. Most precipitation and temperature parameters of the GEM weather generator model were found to be topographically-dependent, and were successfully interpolated to grid points using the PRISM model. A prototype system has been developed whereby users select a location, GEM parameters are extracted, and a daily time series for a specified length of time using GEM is generated. This prototype is for the region of southern Idaho and southeastern Oregon, described above, and can be viewed at the site:

http://ars-boi.ars.pn.usbr.gov/nwrc/climate/gem.html

Future efforts will focus on refining the methodology, testing it in other regions, and expanding the elements generated to include dewpoint, solar radiation and wind speed.

5. REFERENCES

Daly, C., R.P. Neilson and D.L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.*, 33: 140-158.

Daly, C., G.H. Taylor and W. Gibson, 1997: The PRISM approach to mapping precipitation and temperature. *In*: Preprints, 10th Conf. on Applied Climatology, Amer. Meteorological Soc., Reno NV, Oct. 20-24. 5pp.

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. *J. Appl. Meteor.,* 35, 1878-1896.

Nicks, A.D. and G.A. Gander, 1994: CLIGEN: A weather generator for climate inputs to water resource and other models. *Proc. Fifth Int. Conf. Computer in Agric.*, Orlando FL, Am. Soc. Agric. Eng., 903-909.

Richardson, C.W. and D.A. Wright. 1984. WGEN: A model for generating daily weather variables. U.S. Dept. Agric., Agric. Res. Svc. Pub. No. ARS-8, 83 pp.

Woolhiser, D.A., Hanson C.L. and Richardson, C.W. 1988. Microcomputer program for daily weather simulation. US Dept. Agr., Agricultural Research Service, ARS-75, 49 pp.