Semiempirical Down-Scaling of GCM Output to the Local Scale for Temperature, Precipitation, and Runoff

Frederick Wade Freeman and John A. Dracup

ABSTRACT: An empirically derived multiple linear regression model is used to relate a local-scale dependent variable (either temperature, precipitation, or surface runoff) measured at individual gauging stations to six large-scale independent variables (temperature, precipitation, surface runoff, height to the 500-mbar pressure surface, and the zonal and meridional gradient across this surface). Regression equations are calibrated and verified for each dependent variable, for each station, and for each calendar month. The regression equations are then inverted and operated in a semiempirical mode by substituting in GCM produced large-scale output from a CO_2 doubled simulation for the independent variables. The resulting equations are used to predict local values for the three dependent variables; this is the down-scaling process. Down-scaled values for each dependent variable are plotted and contoured to reveal local-scale features. Model performance statistics (the R^2 and F test statistics) are plotted to indicate the spatial variability of the model's performance.

The area investigated is the western United States. The variance explained by the regression model, as indicated by the R^2 test statistic, displays spatial and temporal differences. The explained variance for domains centered over the Sacramento River Basin, in northern California, are 72 to 90 percent for temperature, 42 to 78 percent for precipitation, and 60 to 87 percent for runoff. Results for the area around the Upper Colorado River Basin are slightly lower. The calibration data set is from 1948 through 1988 and includes data from 268 joint temperature and precipitation stations, 152 streamflow stations (which are converted to runoff data), and 24 gridded 500-mbar pressure height nodes.

This study focuses on changes in the temporal and spatial pattern of the local-scale climate and hydrology in the western United States as characterized by altered temperature, precipitation, and surface runoff values for a simulated global climate with a doubled concentration of atmospheric CO_2, a common "greenhouse" gas. Detailed attention is given to the response of two river systems important to the economies of the western United States: the Sacramento (in northern California) and the Upper Colorado (split mainly between Utah and Colorado). Runoff values were determined from streamflow volumes and drainage basin size data for unimpaired basins only.

This research augments in a unique way the existing set of investigations into the effect of climate change on these two basins and on the western United States in general by using a new climatic and hydrologic data set (Wallis *et al* 1991) with an investigative technique new to the field of hydrology: the semiempirical down-scaling regression model. The specific question addressed is how to increase the resolution of large-scale climatic and hydrologic data generated from climate-simulating computer models. The answer we present is a regression model that correlates

---

In: KT Redmond and VL Tharp, Editors. 1994. Proceedings of the Tenth Annual Pacific Climate (PACLIM) Workshop, April 4-7, 1993. California Department of Water Resources, Interagency Ecological Studies Program, Technical Report 36.

local-scale values of temperature, precipitation, and runoff to large-scale averages of the same three variables plus three atmospheric pressure terms (refer to equations 1.1 through 1.3).

Methodology

Our methodology is derived from the well-documented semiempirical modeling approach first employed by Kim *et al* (1984) and later expanded on by Wilks (1989), Wigley *et al* (1990), Karl *et al* (1990), and Storch *et al* (1993). This modeling approach uses an empirically derived multiple linear regression model in conjunction with computer simulated data to translate large-scale, coarse-resolution information into more useful local-scale, high-resolution information. This union between empirical relations and computer simulations inspires the term "semiempirical".

Semiempirical techniques have been employed to model the climatic variables of surface air temperature and precipitation. In this research effort we investigate, along with these two climatic variables, the applicability of this modeling technique to a third parameter: the hydrologic variable of surface runoff.

Model development and use follow the two-step process of:

- Calibration and validation with observed data as the independent and dependent variables, and then
- Down-scaling with computer-simulated data for a changed climate scenario in place of the observed independent variables.

The first step, called the "calibration step", is as follows. Observed data from an individual hydrometeorological station — either temperature, precipitation, or runoff — are used to describe the local-scale characteristics and serve as the dependent variable, which are regressed on a set of five or six independent variables. The independent variables are large-scale values determined by averaging together observed data from all stations located within a specified large-scale domain, which includes the single station used as the dependent variable. It is convenient to set the large-scale domains equal to the established grid elements of a well-tested GCM (general circulation model). We have chosen to use the Goddard Institute for Space Studies (GISS) GCM, which is described by Hansen *et al* (1983) and Hansen *et al* (1988). GISS grid domains are shown in Figure 1.

The six large-scale area average values are: temperature, precipitation, and runoff plus the three atmospheric variables of height to the 500-millibar pressure surface, and the zonal (east-west) and meridional (north-south) 500-millibar pressure gradients. Inclusion of atmospheric pressure data has been shown to increase model performance (*eg*, Wigley *et al* (1990) use three pressure terms; Karl *et al* (1990) use six pressure terms). Calibration equations are established for each calendar month and for

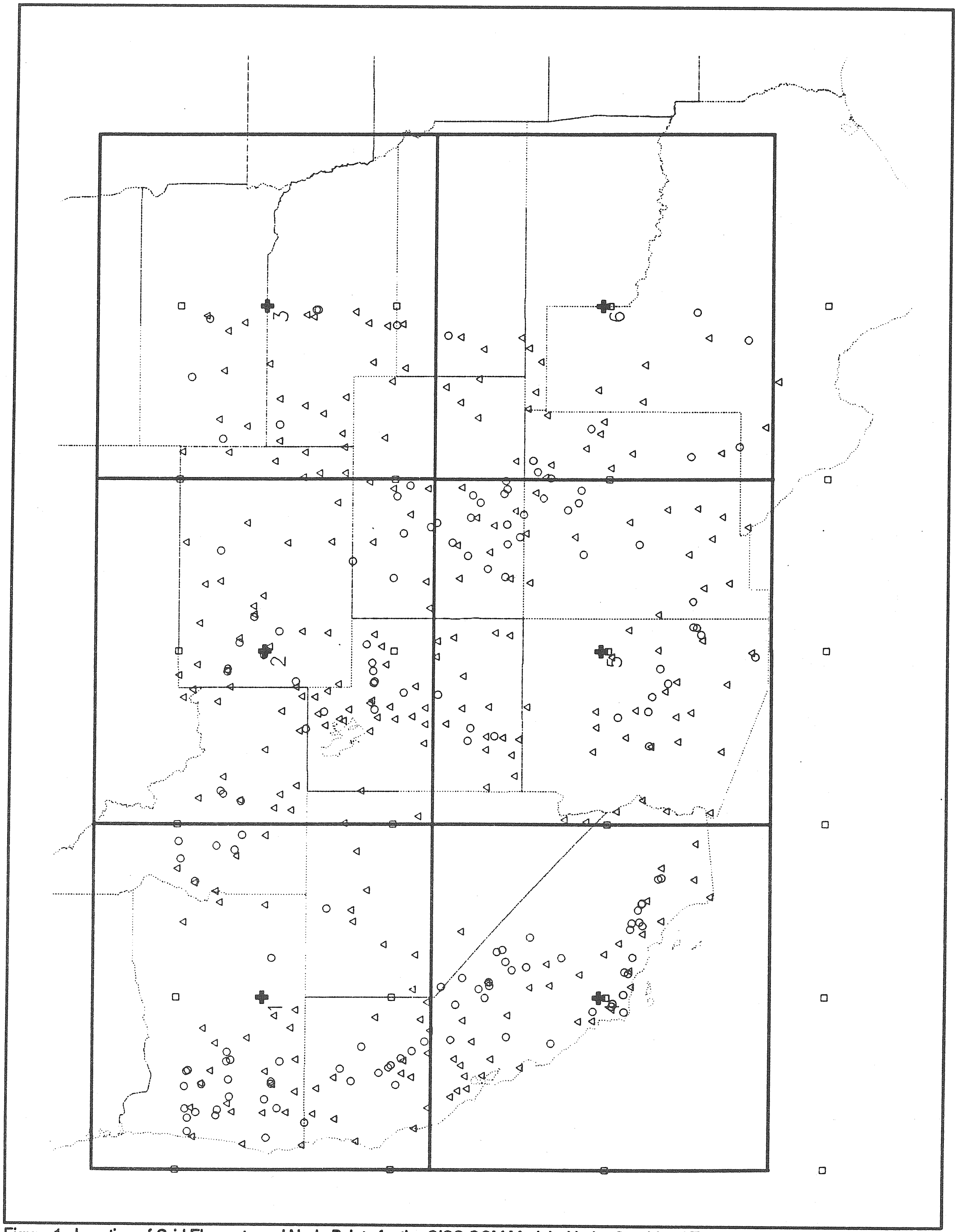

Figure 1. Location of Grid Elements and Node Points for the GISS GCM Model with the Combined Temperature and Precipitation Stations (triangles), Runoff Stations (circles), and Pressure Stations (squares).
The map is shown on a latitude by longitude coordinate system.

each individual temperature, precipitation, and runoff station from 41 years of observed data (1948 through 1988).

The calibration equations take one of the three following forms:

$$TEMP_{myi} = \beta_{0mi} + \beta_{1mi}\overline{TEMP}_{my} + \beta_{2mi}\overline{PRCP}_{my} + \beta_{3mi}\overline{PRESSZ}_{my} + \beta_{4mi}\overline{PRESSU}_{my} + \beta_{5mi}\overline{PRESSV}_{my} + \varepsilon_{mi} \tag{1.1}$$

$$PRCP_{myi} = \beta_{0mi} + \beta_{1mi}\overline{TEMP}_{my} + \beta_{2mi}\overline{PRCP}_{my} + \beta_{3mi}\overline{PRESSZ}_{my} + \beta_{4mi}\overline{PRESSU}_{my} + \beta_{5mi}\overline{PRESSV}_{my} + \varepsilon_{mi} \tag{1.2}$$

$$FLOW_{myi} = \beta_{0mi} + \beta_{1mi}\overline{FLOW}_{my} + \beta_{2mi}\overline{TEMP}_{my} + \beta_{3mi}\overline{PRCP}_{my} + \beta_{4mi}\overline{PRESSZ}_{my} + \beta_{5mi}\overline{PRESSU}_{my} + \beta_{6mi}\overline{PRESSV}_{my} + \varepsilon_{mi} \tag{1.3}$$

where *TEMP*, *PRCP*, and *FLOW* are observed local-scale dependent variables for temperature, precipitation, and runoff; $\overline{TEMP}$, $\overline{PRCP}$, and $\overline{FLOW}$ are observed large-scale area average independent variables for these same three quantities; $\overline{PRESSZ}$, $\overline{PRESSU}$, and $\overline{PRESSV}$ are observed large-scale area average independent variables for the three pressure terms: *Z* is for the 500-millibar surface height, *U* is for the zonal gradient across this surface, and *V* is for the meridional gradient across this surface; the βs are regression coefficients; the εs are the error terms; and *m*, *y*, and *i* are the calendar months, the specific year from the time series, and a specific gauging station, respectively ($m = 1, 2, \ldots 12$; $y = 1, 2, \ldots 41$; and $i = 1, 2, \ldots 268$ for temperature and precipitation, and $i = 1, 2, \ldots 152$ for runoff). $\overline{FLOW}$ is not included among the independent variables when regressing for *TEMP* or *PRCP* because there is no clear physical linkage in this direction. During the calibration step, the equations are solved to determine the value of the βs.

The second step of the modeling process begins once the regression relationships are calibrated and the regression coefficients, the βs, are known. In this step, the observed large-scale data originally used as the independent variables are directly replaced with large-scale data simulated by the GISS GCM operated under the climate-changed scenario of a doubled concentration of atmospheric CO_2. With the GCM data in place and the βs known, the regression equations are said to be "inverted" and are now operated to generate local-scale values at individual observation stations for the climate condition employed by the GCM. This application of the model is termed the "down-scaling step". Observed values for the three pressure terms were used in the down-scaling step (as well as in the calibration step) because GCM-simulated values were not available.

The modeling procedure is complete after regressions are first calibrated and then down-scaled for each individual observation station, for each of the three dependent variables, and for each month of the year. The modeling philosophy is summarized by Giorgi and Mearns (1991) as follows: The strategy is to treat large-scale forcings explicitly through the use of GCMs, and account for local-scale forcings in an empirical fashion.

When we view the process of down-scaling coarse-resolution data to the local scale, or sub-grid scale, as depicted schematically by Gates (1985) in Figure 2, we see that the semiempirical model presented here allows for the computation of step number two labeled "local climate change". The other steps can be easily accomplished with existing models and statistical procedures. The result is a complete methodology to determine the local-scale effect of climate change on temperature, precipitation, and runoff — which is labeled as step four "local impact". An outcome of having adapted a known meteorological regression technique to basin-scale streamflow and runoff is that a new avenue now exists for the investigation of hydrologic drought resulting from a changed climate. As a result, drought statistics can be generated on a local scale from GCM-simulated data for a CO_2 doubled environment.

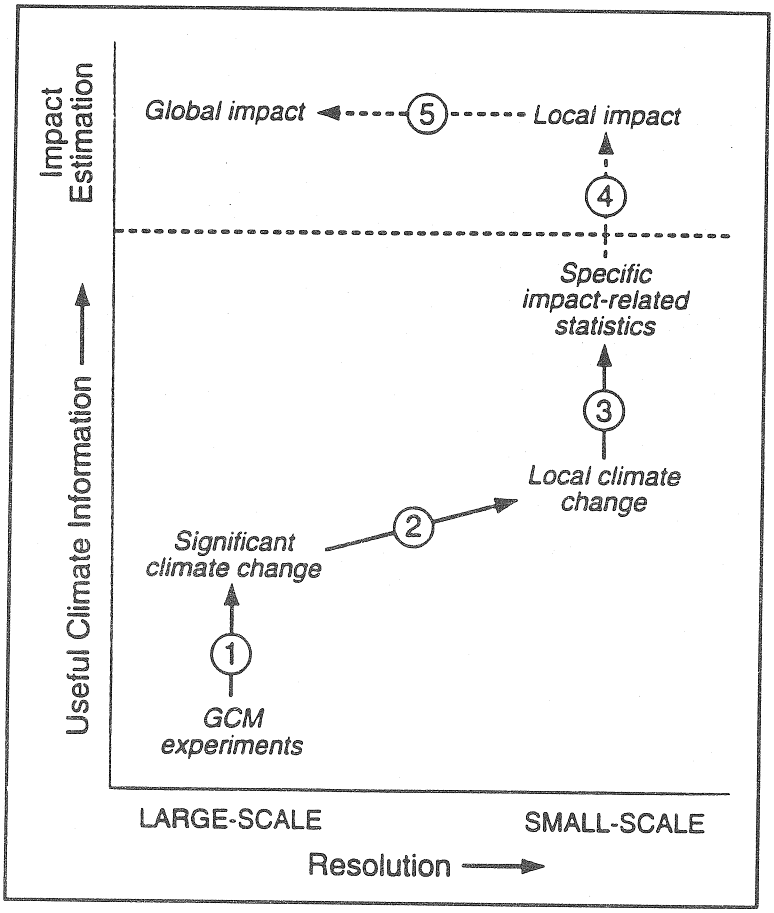

Figure 2. Schematic of the Processes of a Strategy for Down-Scaling to the Local (or Small) Scale the Results from a GCM Experiment Simulating Climate Change
(From Gates 1991)

The output from the regression model serves two significant purposes.

- Down-scaled values can easily be used to construct contour maps, with local-scale resolution, that display monthly values for temperature, precipitation, and runoff for an environment with a doubled CO_2 concentration. Two statistical parameters are also plotted along with the down-scaled values. The first is the goodness of fit statistic R^2, which measures the proportion of variance in the dependent variable that is explained by the model. The second is the F statistic, a significance test of the model as a whole.

- The output can also be used as a comparison data set for similar quantities produced from GCMs. Comparisons of output from the regression model and GCMs can serve to validate GCM simulations and to help modify the underlying parameterizations used in the hydrologic processor internal to several GCMs (Giorgi and Mearns 1991).

Only the first purpose is explored here.

Analysis

Presented next are the down-scaled results for the western United States for temperature, precipitation, and runoff based on a CO_2 doubled climate simulated by the GISS GCM. Results are determined as differences (CO_2x2–CO_2x1) for temperature and as ratios (CO_2x2/CO_2x1) for precipitation and runoff. Figure 3 shows contoured January output for temperature, and Figure 4 is a contour map of the model performance statistic R^2. Figures 5 and 6 show the same information for precipitation. These maps reveal the local-scale variation of the down-scaled values. Values for the F statistics are not shown. Down-scaled runoff values for July and the corresponding R^2 values are plotted on Figures 7 and 8. Similar plots for the remaining months have been generated but are not included here.

A high degree of spatial variability for the change in down-scaled runoff is observed over regions encompassing the Sacramento and Upper Colorado River Basins, with only limited variability elsewhere. This is most pronounced during winter months for the Sacramento Basin and during summer months for the Upper Colorado Basin. Down-scaled temperature change displays a fairly even amount of variability throughout the year, although the model output does show a gentle preference for greatest variability over the Upper Colorado Basin during winter and over the Sacramento Basin and the southern Sierra Nevada during summer. Down-scaled change in precipitation is more similar to down-scaled runoff than temperature, with greater variability occurring over western Oregon and northern California during winter and spring, and over the Rocky Mountains and the western Great Plains during summer.

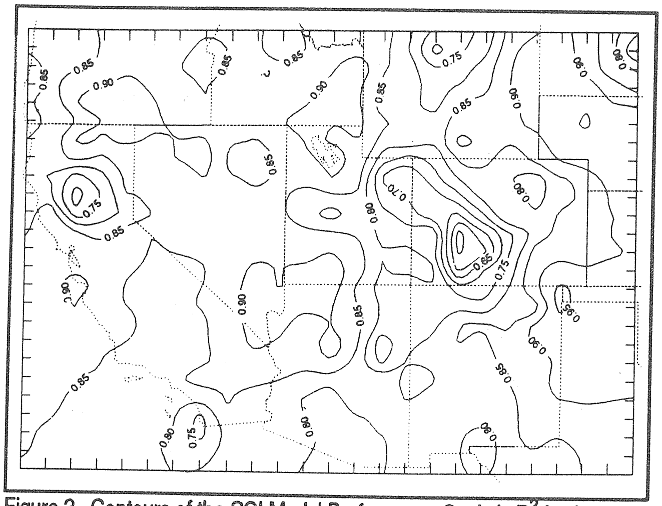

Figure 3. Contours of the SCI Model Performance Statistic R^2 for January for the Down-Scaling of Temperature

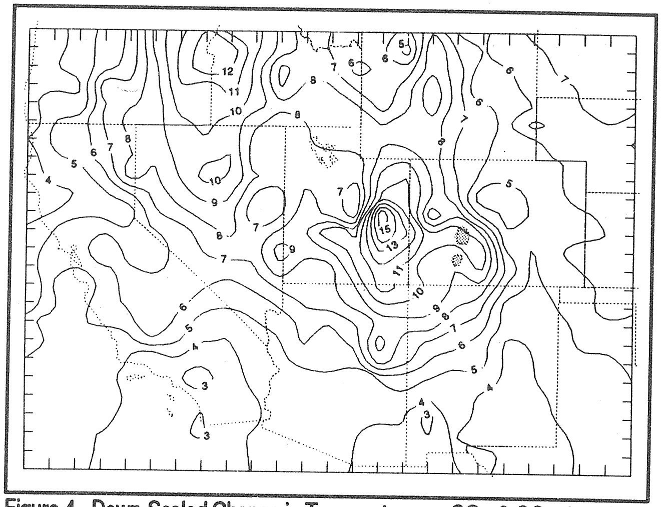

Figure 4. Down-Scaled Change in Temperature, as CO_2x2-CO_2x1 and Measured in °C, for January Using GISS GCM Output Data.
Shaded areas indicate regions where the R^2 value is below 0.5.

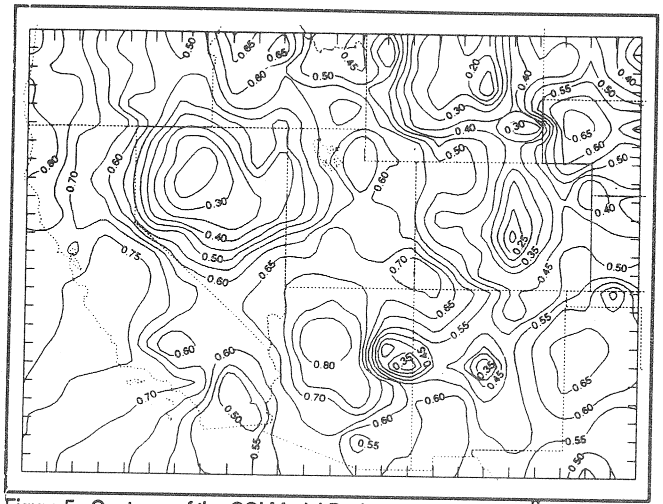

Figure 5. Contours of the SCI Model Performance Statistic R^2 for January for the Down-Scaling of Precipitation

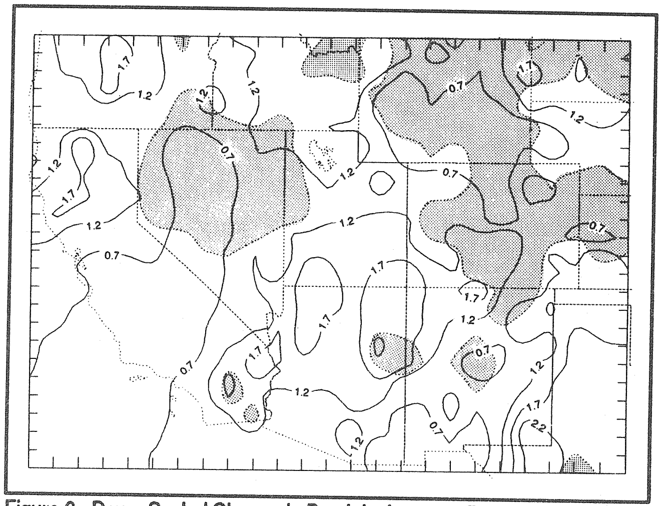

Figure 6. Down-Scaled Change in Precipitation, as a Percent of CO_2x1 Precipitation and Measured in Millimeters/Day, for January Using GISS GCM Output Data.
Shaded areas indicate regions where the R^2 value is below 0.5.

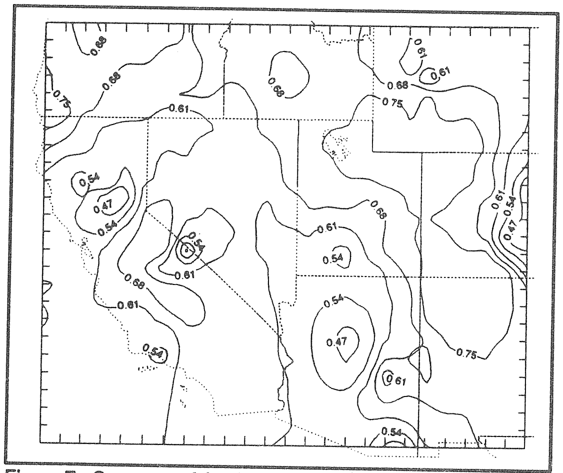

Figure 7. Contours of the SCI Model Performance Statistic R^2 for July for the Down-Scaling of Runoff

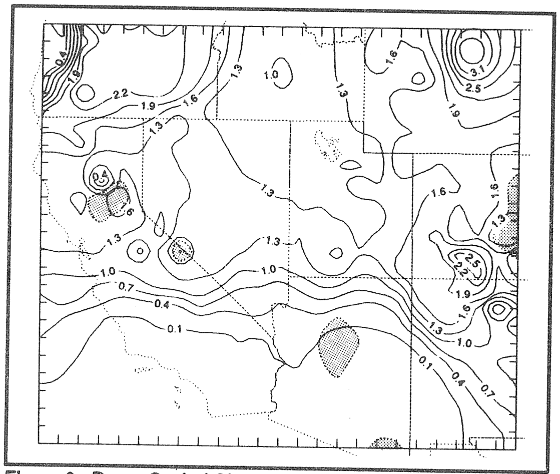

Figure 8. Down-Scaled Change in Runoff, as a Percent of CO_2x1 Runoff and Measured in Millimeters/Day, for July Using GISS GCM Output Data.
Shaded areas indicate regions where the R^2 value is below 0.5.

References

Gates, WL. 1985. The Use of General Circulation Models in the Analysis of the Ecosystem Impacts of Climate Change. *Climatic Change.* 7:267-284.

Giorgi, F, and LO Mearns. 1991. Approaches to the Simulation of Regional Climate Change: A Review. *Reviews of Geophysics.* 29:191-216.

Hansen, J, G Russell, D Rind, P Stone, A Lacis, S Lebedeff, R Ruedy, L Travis. 1983. Efficient Three-Dimensional Global Models for Climate Studies: Models I and II. *Monthly Weather Review.* 3:609-662.

Hansen, J, I Fung, A Lacis, D Rind, S Lebedeff, R Ruedy, G Russell, P Stone. 1988. Global Climate Changes as Forecast by Goddard Institute for Space Studies Three-Dimensional Model. *Journal of Geophysical Research.* 93(D8):9341-9364.

Karl, TR, W-C Wang, ME Schlesinger, RW Knight, D Portman. 1990. A Method of Relating General Circulation Model Simulated Climate to the Observed Local Climate, Part I, Seasonal Statistics. *Journal of Climate.* 3:1053-1079.

Kim, J-W, J-T Chang, NL Baker, DS Wilks, WL Gates. 1984. The Statistical Problem of Climate Inversion: Determination of the Relationship between Local and Large-Scale Climate. *Monthly Weather Review.* 112:2069-2077.

Storch, HV, E Zorita, U Cubasch. 1993. Downscaling of Global Climate Change Estimates to Regional Scales: An Application to Iberian Rainfall in Wintertime. *Journal of Climate.* 6:1161-1171.

Wallis, JR, DP Lettenmaier, EF Wood. 1991. A Daily Hydroclimatological Data Set for the Continental United States. *Water Resources Research.* 27:1657-1663.

Wigley, TML, PD Jones, KR Briffa, G Smith. 1990. Obtaining Sub-Grid-Scale Information from Coarse-Resolution General Circulation Model Output. *Journal of Geophysical Research.* 95:1,943-1,953.

Wilks, DS. 1989. Statistical Specification of Local Surface Weather Elements from Large-Scale Information. *Theoretical and Applied Climatology.* 40:119-134.