Spatial modeling of winter temperature and precipitation in Arizona and New Mexico, USA

David P. Brown*, Andrew C. Comrie

Department of Geography and Regional Development, University of Arizona, Tucson, Arizona 85721, USA

ABSTRACT: The development of a statistical modeling technique suitable for producing mean and interannual gridded climate datasets for a topographically varying domain is undertaken. Stepwise regression models at 1 × 1 km resolution are generated to estimate mean winter temperature and precipitation for the Southwest United States for the years 1961–1990. Topographic predictor variables are used to explain spatial variance in the datasets. Kriging and inverse distance weighting interpolation algorithms are utilized to account for model residuals. The final regression models show a high degree of explained variance for temperature (R^2 = 0.98, mean bias error [MBE] = –0.15°C, root-mean-squared error [RMSE] = 0.74°C) and a moderate degree of explained variance for precipitation (R^2 = 0.63, MBE = –1.4 mm, RMSE = 27.0 mm). Several smaller-scale precipitation regression models are developed for comparison to the domain-wide model, but do not show marked accuracy improvements. Observed values of winter temperature and precipitation from the years 1961–1999 are compared to the 30 yr modeled means, and the differences are interpolated using kriging (temperature) and inverse distance weighting (precipitation). The result is a 39 yr time series of maps and datasets of winter temperature and precipitation at 1 × 1 km resolution for the Southwest United States.

KEY WORDS: Temperature · Precipitation · Southwest US · Regression · Interpolation · GIS

1. INTRODUCTION

1.1. Background

This paper addresses the issue of improving spatial coverage of climate data in the SW United States, both for climatic analysis and for use by public and private users. As is the case across much of the world, the distribution of meteorological stations in this region is not uniform. Large swaths of land, particularly federal- and state-owned territory and tribal lands, have sparse data coverage, while in contrast a clustering of reporting stations can be noted in several of the major metropolitan areas. These gaps in the data must be accounted for in the construction of any spatially continuous datasets.

Because the topography of an area influences spatial climate patterns, a portion of the variability of temper-ature and precipitation may be explained using a terrain-based model (Ffolliott et al. 1989). This type of model permits the use of both regression modeling and interpolation methodologies for climatic dataset creation. Regression techniques can be utilized to develop relationships between topographic variables and temperature and precipitation at locations of measured climate data; these relationships can then be applied to areas possessing appropriate topographic data but lacking measured climate information. Interpolation algorithms may then be employed to account for variance in measured climate data unexplained by topographic variables. Because of the varied topography of the SW United States, several topographic predictor variables may exert influences on spatial patterns of temperature and precipitation.

Numerous attempts to model precipitation fields by way of statistical prediction have been made (e.g. de Montmollin et al. 1980, Bigg 1991, Hevesi et al. 1992a,b, Kilsby et al. 1998, Prudhomme & Reed 1999).

---

*E-mail: dpbrown@u.arizona.edu

Studies at various time scales have been undertaken, such as in the analysis of daily (Beek et al. 1992) and monthly (Goodale et al. 1998) precipitation patterns. Karnieli (1990) and Comrie & Broyles (2002) provide regional examples of how kriging can be used to map precipitation patterns in Arizona. Likewise, a variety of statistical techniques have been applied to temperature data, at timescales ranging from annual (Willmott & Matsuura 1995) to daily (Dodson & Marks 1997). As was the case in the modeling of precipitation, the use of elevation to predict temperature has been explored (e.g. Ishida & Kawashima 1993). Various interpolation algorithms, such as spline and inverse distance weighting, have also been used in temperature analyses, such as by Kurtzman & Kadmon (1999). Frequently, these analyses focused on mountainous areas similar to those found in the SW United States.

The question of choosing an appropriate resolution for statistical climate modeling has in the past been limited by input data availability and computing power. In recent years, however, the ability to create robust statistical models of temperature and precipitation on the order of 10 km has become more feasible, due to the advent of powerful desktop computers and the availability of high resolution digital elevation models (DEMs). The Precipitation-Elevation Regressions on Independent Slopes Model (PRISM) was one of the first such attempts to model climate data at this resolution, using terrain predictor variables at an approximate resolution of 6 km × 9 km (Daly et al. 1994). Given that more recent studies have demonstrated the ability to conduct analyses ranging to 1 × 1 km resolution (e.g. Goodale et al. 1998, Agnew & Palutikof 2000), and the appropriateness of that resolution for many stakeholder uses of climate data, we have elected to conduct our modeling at 1 × 1 km resolution.

Climatic datasets created at 1 × 1 km resolution for the winter season in the SW United States are particularly desirable for both scientific pursuits and more general application. Datasets at this resolution permit investigation of climate patterns at local and subregional levels, as well as the study of the strong correlation between precipitation and snowpack depth, water storage, and runoff (Serreze et al. 1999). Often, these data can be employed in other modeling exercises, such as hydrological and geomorphological impact models (Agnew & Palutikof 2000), and in ecologically based studies including land cover change and fire history. In particular, the seasonal timescale used in this study correlates with the work of tree-ring researchers in the reconstruction of long-term climate variability at high spatial resolution. Finally, these datasets are considered to possess sufficiently high-resolution information for many public uses in the SW

United States by ranchers, farmers, fire managers, and others. This research therefore shows one example of how products resulting from this 'proof of concept' methodology can be applied to other environmental studies.

1.2. Aims and objectives

The aims of this paper are 2-fold: first, the development of a statistical modeling technique suitable for constructing gridded climate datasets for a topographically varying domain; and second, the actual production of mean and interannual winter temperature and precipitation datasets for the SW United States. Within the first aim, we address several subquestions. We identify the amount of variance in temperature and precipitation data that is spatial in nature, and we attempt to account for this variance by way of topography. We note topographic variables that serve best as predictors of temperature and precipitation from a set of modeling approaches. Finally, we determine the interpolation algorithms most useful in mapping regression model residual values.

2. STUDY AREA

The study area for this analysis is Arizona and New Mexico, neighboring states located at approximately 105 to 115° W and 32 to 38° N (Fig. 1). While the focus of the research was these 2 states, the network of meteorological stations used in the analysis also covers neighboring portions of California, Nevada, Utah, Colorado, Oklahoma, Texas, and Mexico, so that possible edge effects arising from interpolation were minimized. The entire domain was thus defined as the area from approximately 30 to 39° N and 100 to 116° W.

The study region is characterized by extremely varied physiography, ranging from isolated and extended mountain ranges to low-elevation deserts. Arizona and New Mexico are 2 of the warmest and most arid states in the United States, with the mountains receiving more moisture annually than the deserts. Several medium- to large-sized cities are included within the study area, but the region also includes many national forests, wilderness areas, tribal lands, farms, and ranches.

3. DATA

Monthly mean temperature (°C) and total precipitation (mm) data for the winter season, defined as December through March, were obtained from the

partial or no data (e.g. a missing monthly value), they were omitted from the analysis of the winter season in question (Fig. 2).

4. METHODOLOGY

4.1. Variable development

Both the temperature and precipitation data were averaged temporally to create 30 yr winter means (December–March) from 1961–1990 for each station. Stations with at least 75% temporal coverage during the 30 yr period were preserved for use in model development. A random spatial subset consisting of approximately 25% of the 1961–1990 temperature and precipitation station data was retained so that the predictive regression functions could be tested on independent data. This large independent data pool was needed to verify the temperature and precipitation models over a wide range of topography across the study area.

Using the ARC/INFO geographic information system (GIS), an initial set of terrain variables including elevation, slope, and aspect were calculated at 1×1 km resolution for each grid cell. The slope for each grid cell was calculated from a 3×3 neighborhood using the average maximum technique, and the direction of the plane fitted to the z-values of these cells provided measurements of aspect (Burrough 1986). Once derived in the GIS, dummy terms were used to define the 4 compass directions of aspect (north, east, south, and west). Latitude and longitude coordinates were also
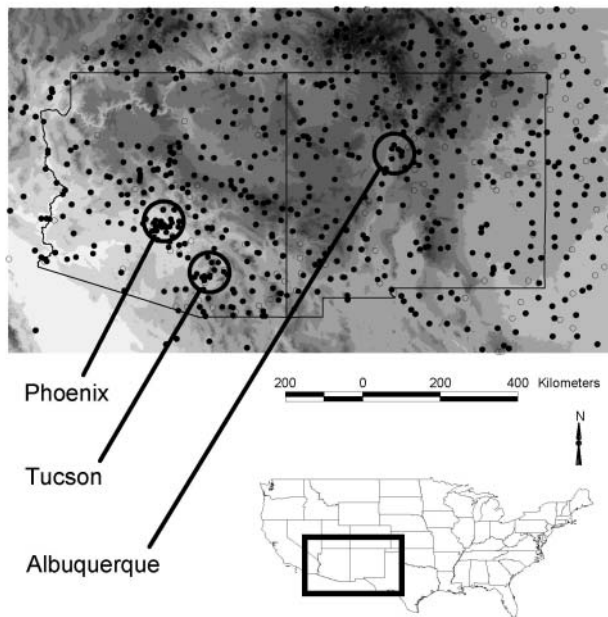

Fig. 1. The study area encompasses the states of Arizona and New Mexico, with data also collected from portions of the surrounding 6 states and Mexico. Locations of National Weather Service and Global Historical Climate Network temperature (solid) and precipitation (open) data stations used in regression model development are also shown

National Climatic Data Center (http://www.ncdc.noaa.gov) for the years 1961–1999. A total of 662 Cooperative and Historical Climate Network stations were used for the temperature analyses, and a total of 572 stations for the precipitation analyses (Fig. 1). A 1×1 km DEM developed by the United States Geological Survey, the gtopo30 DEM, was used to provide topographic data for the study region. Station elevations ranged from 4 m (San Felipe, Mexico) to 3207 m (Trout Lake, Colorado).

The majority of the winter precipitation in Arizona and New Mexico happens during December, January, February, and March; consequently, these months were chosen to represent the climatic winter. The study period 1961–1999 was selected because it was well represented in the station data. Pre-1961 records of temperature and precipitation conditions were less consistent than during the 1961–1999 period. A 30 yr climate 'normal' period with robust spatial coverage of temperature and precipitation data, 1961–1990, was selected for use in the model development portion of the research. In instances where individual reporting stations contained

Fig. 2. Number of temperature (solid) and precipitation (dashed) stations utilized each winter from 1961–1990

assigned to each grid cell and used as regression predictor variables.

Given the exploratory nature of the study, several co-linear predictor variables with the potential to explain spatial variability in the climate data were generated and used in the initial modeling. These included the top 5% and lowest 5% elevation stations, which were assigned dummy variables to identify any variance associated with extremes in elevation. Cross-products involving longitude or latitude and the various elevation terms were also calculated. These regression predictors were used to account for gradients in temperature and precipitation associated both with elevation and geographic location (e.g. more northerly high elevation locations receiving more precipitation than similar elevation sites at more southerly locations). The addition of a final predictor variable, capturing a possible 'oasis effect' associated with the Phoenix metropolitan area in central Arizona, is discussed in Section 4.2. When necessary, variables were transformed via natural logarithms to improve their normality, one of the key assumptions of multivariate regression. Table 1 illustrates the total set of predictor variables generated from the DEM.

4.2. Model development

The methodology utilized to develop the domain-wide temperature and precipitation models is outlined in Fig. 3. Using stepwise multiple linear regression, the predictor variables in Table 1 were used to explain variance in the temperature and precipitation data. Separate models were created for mean winter temperature and mean total winter precipitation. The stepwise regression approach permitted the exclusion of statistically insignificant ($\alpha > 0.05$) predictor variables. Once regression relationships between significant predictors and the climate variables were obtained, estimates of temperature and precipitation were made for each 1×1 km grid cell in the study area based on each grid cell's topographic characteristics. This technique permitted uniform spatial coverage of 1×1 km regression predictions of temperature and precipitation values, including grid cells containing 1 or more meteorological stations.

To identify biases inherent to the regression estimates, station data used in the regression model development were compared to their corresponding predicted grid cell values of temperature and precipitation. Differences between the predicted and observed values, or residuals, of mean winter temperature and mean total winter precipitation were calculated for each station. The residual data were inspected to identify areas of significant model under- or

Table 1. Candidate topographic predictor variables evaluated in the step-wise regressions for temperature and precipitation

Latitude	Longitude × elevation
Longitude	Longitude × High elevation
Elevation	Longitude × Low elevation
Slope	Latitude × Elevation
North aspect	Longitude × Slope
East aspect	Latitude × High elevation
South aspect	'Oasis effect' term
West aspect	Latitude × Low elevation
High elevation (top 5%)	'Oasis effect' term squared
Low elevation (bottom 5%)	Latitude × Slope

over-prediction, as well as for their independence, another basic assumption of regression. This examination revealed a large precipitation residual (under-prediction) to the east of the Phoenix metropolitan area, also corresponding to the location of the Mogollon Rim, an area of high elevation in eastern Arizona.

To account for this residual, a final predictor variable was derived from the DEM and used to capture the effect of the enhanced elevation of the Mogollon Rim. This southeast-to-northwest topographic feature is the first significant high-elevation region encountered by eastward-moving air masses in the Southwest, and the relative effect of orography in this region may therefore be enhanced. In addition, the variable may account for any suspected 'oasis effect' associated with Phoenix, a large urban area (Spronken-Smith et al. 2000). Increased absolute humidity associated with this urban environment may increase the magnitude of precipitation when orographic lifting occurs downwind. To calculate this 'oasis effect' variable, the linear distance from the center of Phoenix to each grid cell was calculated and used as a regression predictor. This term was also squared to account for the possibility that the 'oasis effect' did not exhibit a linear relationship with temperature and/or precipitation; both the squared and non-squared predictors were used as candidate variables in the regression.

After removing all reasonably explainable patterns and spatial structure due to topographic variance, the final temperature residuals were interpolated across the study region at 1×1 km resolution. This was permissible as the residuals from the models were random in the sense of being free of terrain-based correlation, but not necessarily other spatial structure (i.e. atmospheric variability). For the residuals, a spherical semivariogram model was generated, with the assumption that variations in the residual data were free of terrain-based structural components (McBratney & Webster 1986). Ordinary kriging was then utilized for the interpolation of the temperature residuals, as this technique provides for exact interpolation (Burrough & McDon-

Method

Obtain gridded topographic variables from DEM

Obtain temperature (T) and precipitation (P) station data from NCDC

Retain ~25% for independent model verification

Utilize ~75% for model development

Assign grid cell topographic variables to each T and P station

Conduct stepwise regression of T and P station data with corresponding topographic variables

Estimate T and P for each grid cell using regression models

Calculate estimated grid cell values minus observed station values (residuals) for each station

Interpolate residuals from station locations to 1x1 km resolution across domain

Subtract residuals from regression predictions to honor original T and P station data and create final T and P models

Compare independent station data to corresponding grid cell predicted values for accuracy assessment

Product

Regression models of T and P

Maps of estimated T and P

Station Residuals

Maps of T and P model residuals

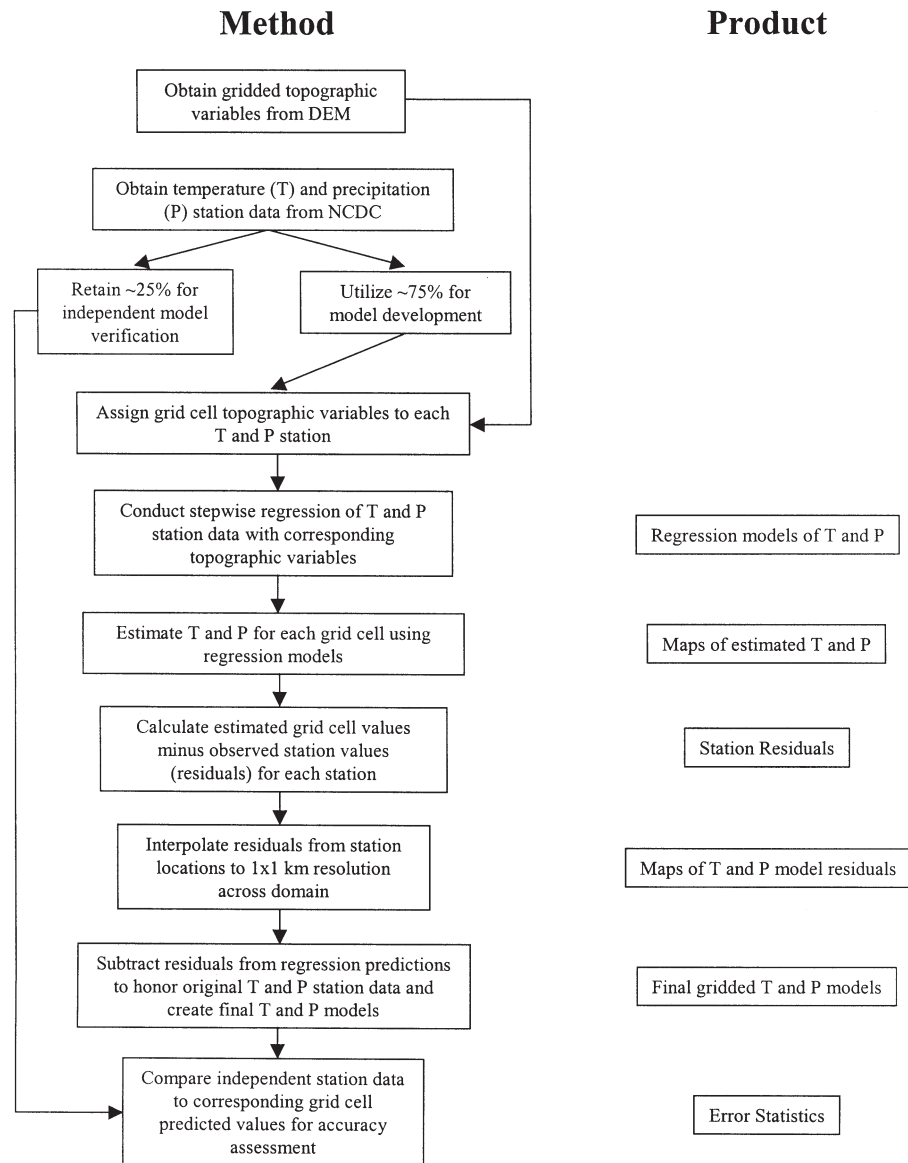Final gridded T and P models

Error Statistics

Fig. 3. Flow chart illustrating methodology used to create regression-based models of winter temperature and precipitation

nell 1998) using the Levenberg-Marquardt method (Press et al. 1988). In the same fashion as for temperature, the precipitation residuals were interpolated at 1 × 1 km resolution using ordinary kriging and a linear semi-variogram model. However, poor initial results generated from the kriging of the precipitation residuals necessitated further analyses using an alternative algorithm, inverse distance weighting (IDW). This interpolation method is useful because of its ability to weight the control points based on their distance from the location of estimation (Davis 1986). The 12 nearest-neighbor stations were used in the predictive function, with slightly greater weight (power of 2) placed on those stations closest to the cell for which an estimated value was being produced. For both kriging and IDW, spherical coordinates were used in a geographic projection.

To create the final modeled maps of winter temperature and precipitation, the interpolated grid cell residual values of temperature and precipitation were added to their corresponding regression estimates, thereby honoring the original station data used for model development. The calculations were made using the GIS software. The region was then 'clipped' to include only Arizona and New Mexico, thus eliminating highly anomalous edge effects that may have resulted from the interpolation along the boundaries of the region.

The independent (cross-validation) station data retained before beginning the regression analysis were compared to the modeled temperature and precipitation data to verify the robustness of the models. Standard measures of model performance, including Pearson's correlation coefficient (R^2), mean bias error (MBE), and root mean square error (RMSE), were calculated. A second measure of the degree of agreement between the observed and modeled data, Willmott's *D* (Willmott 1981), was utilized. The scope of the equation and its development prevent it from being reproduced here; however, it is similar to Pearson's coefficient in that its measure ranges from 0 (no agreement) to 1 (perfect agreement) while being based on actual data units, thereby incorporating actual rather than relative similarity between datasets. A final accuracy assessment was made by computing the proportion of systematic error (PSE), which highlights model-oriented error as opposed to error resulting from data 'noise' (Willmott et al. 1985). Again, the scale of measurement ranges from 0 to 1; in this instance, a smaller PSE indicated a model with fewer inherent errors.

4.3. Alternative modeling methodologies for precipitation

In addition to the 1 × 1 km domain-wide regression model constructed for winter precipitation, sub-domain regressions were also developed from the precipitation data. This was done with the suspicion that the domain-wide model might not be sensitive enough to local orographic variations in winter precipitation. Fifteen 3° latitude × 3° longitude cells were superimposed on the study area in the form of a 3 × 5 matrix. These larger grid cells were constructed to be as small as possible still while providing for a minimum number of reporting stations within each cell. Stepwise linear regression was applied as before to each of these 15 sub-domains at 1 × 1 resolution, using the same predictor variables utilized in the domain-wide regressions. The pooled error statistics from the 15 sub-domain regions were calculated for the same subset of independent station data and compared to those of the domain-wide regression model.

An additional method of enhancing these sub-domain precipitation regression models was tested. The unstandardized regression coefficient values for the predictor variables were taken from each of the sub-domain models. These coefficients, including constants, were placed at the centroid of each 3° latitude × 3° longitude sub-domain region, and interpolated from the 15 centroids across the domain at 1 × 1 km resolution. For this interpolation, splines were utilized, as they provide a smooth connection between a minimum

number of data points (Davis 1986). The interpolated maps of regression coefficients were used to estimate temperature and precipitation in a fashion analogous to those of the domain-wide and sub-domain regression models. In this instance, however, the interpolation scheme permitted the values of the predictor coefficients to change from grid cell to grid cell. Again, the same subset of independent station data was compared to their corresponding predicted values, and the resulting error statistics compared to those of the other precipitation modeling approaches.

4.4. Creation of interannual winter temperature and precipitation maps

The methodology utilized to develop interannual maps of winter temperature and precipitation is outlined in Fig. 4. Following construction of the winter temperature and precipitation models based on 1961–1990 data, interannual winter maps of the 2 climate variables from 1961–1999 were generated. Individual temperature and precipitation station data for each winter season were compared with their corresponding grid cell means, which remained constant. Temperature and precipitation anomalies were calculated for each winter by subtracting the stations' observed values from their corresponding 30 yr modeled means. These anomalies were interpolated across the domain at 1 × 1 km resolution from the station locations using kriging (temperature) or IDW (precipitation). Anomaly maps were then added to the modeled mean temperature or precipitation maps (e.g. 1961 winter temperature anomalies + 30 yr modeled mean temperatures) to produce maps of actual temperature or precipitation data. This was done for each winter season, resulting in a time series of climate maps wherein observed station data had been honored. The graphical quality of this time series of climate maps was examined by contrasting it with the simple interpolation of observed temperature and precipitation station data from each winter season at 1 × 1 km resolution.

5. RESULTS

5.1. Model development

The significant ($\alpha = 0.05$) predictor variables chosen for the final temperature model were identified as elevation, latitude, longitude, and slope. For precipitation, significant predictors included longitude, elevation, and slope, as well as interaction terms including longitude × elevation, longitude × top 5% elevation stations, and longitude × bottom 5% elevation stations. An additional
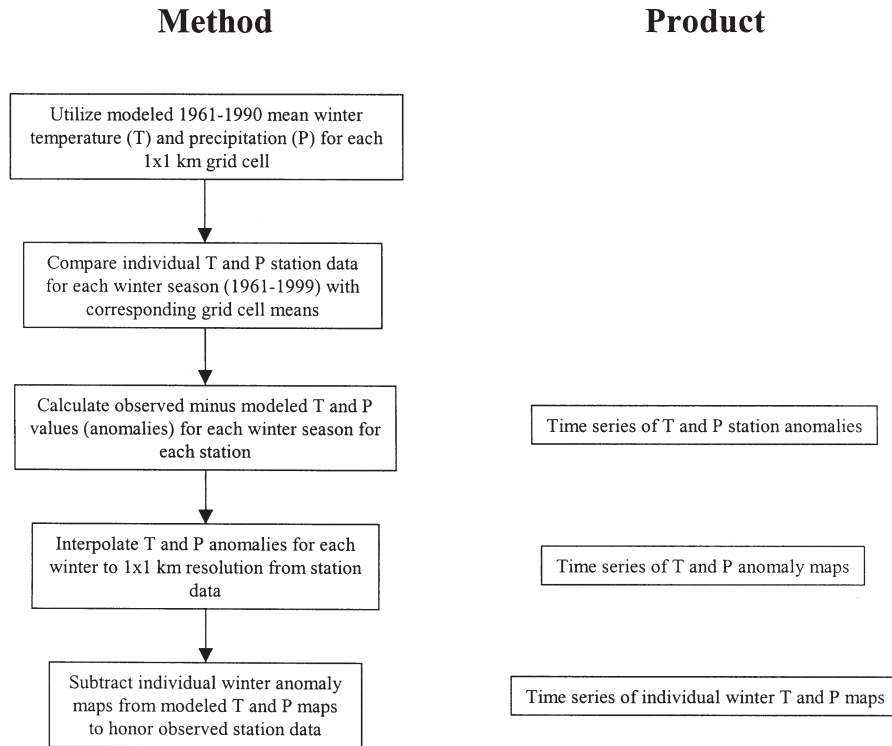
Method   **Product**

Utilize modeled 1961-1990 mean winter temperature (T) and precipitation (P) for each 1x1 km grid cell

Compare individual T and P station data for each winter season (1961-1999) with corresponding grid cell means

Calculate observed minus modeled T and P values (anomalies) for each winter season for each station

Time series of T and P station anomalies

Interpolate T and P anomalies for each winter to 1x1 km resolution from station data

Time series of T and P anomaly maps

Subtract individual winter anomaly maps from modeled T and P maps to honor observed station data

Time series of individual winter T and P maps

Fig. 4. Flow chart illustrating methodology used to create interannual maps of winter temperature and precipitation and associated anomalies

precipitation variable was the squared term representing the Phoenix-related 'oasis effect'. The relative importance of the predictors for both models is noted in Table 2 by their standardized (beta) coefficients. All of the other variables, including the 4 dummy variables representing aspect, were rejected in the stepwise regressions due to lack of statistical significance.

Table 2. Significant standardized beta values from the stepwise regressions for temperature and precipitation. Variables with larger beta coefficients explain a greater percentage of variance in the data

Variable	Standarized (beta) coefficient
Temperature	
Elevation	−0.737
Latitude	−0.345
Slope	0.122
Longitude	−0.093
Precipitation	
Longitude × High elevation	−6.878
High elevation	−6.519
Longitude × Elevation	−5.800
Elevation	−5.620
'Oasis effect'	−0.513
Longitude	0.335
Slope × Slope	0.217
Longitude × Low elevation	0.097

Verification of the performance of the domain-wide regression models on independent (cross-validation) data is provided in Table 3. The temperature model performed exceedingly well, with a coefficient of determination (R^2) of 0.98 after interpolating residual values using kriging and combining them with the regression predictions. The high level of total explained variance in the temperature model is primarily the result of the influence of elevation, as its effect on temperature is generally consistent across the domain. The root-mean-square error (RMSE) is under 1°C, mean bias error (MBE) is small, and proportion of systematic (model-oriented) error (PSE) is only 7%. In addition, the Willmott's D was high, thus lending a high degree of confidence to the model. As an additional assessment, the performance of the temperature model was judged against the results of a smart interpolation technique (Willmott & Matsuura 1995) using the same sets of training and independent data. The regression approach is slightly superior to the smart interpolation method for grid-based temperature prediction in the Southwest ($R^2 = 0.98$ vs $R^2 = 0.95$), as expected, with minor additional skill likely coming from the additional terrain variables included in the regression approach.

As expected, terrain influences did not account for the spatial patterns of precipitation as well as they did

Table 3. Error statistics resulting from testing of temperature model on independent data, and precipitation models on independent and model development data. IDW: inverse distance weighting; RMSE: root-mean-square error; MBE: mean bias error; *D*: Willmott's *D* (Willmott 1981); PSE: proportion of systematic error

	Temperature: independent data Regression plus kriging	Precipitation: independent data Regression plus IDW	Precipitation: independent data Regression plus kriging	Precipitation: model data Regression plus IDW	Precipitation: model data Regression plus kriging
N	129	113	113	425	425
Average observation (°C or mm)	5.7	76.28	76.28	86.39	86.39
Average prediction (°C or mm)	5.5	80.07	78.81	83.39	84.4
R^2	0.98	0.63	0.60	0.90	0.73
RMSE (°C or mm)	0.74	27.0	28.0	13.7	22.7
MBE (°C or mm)	−0.15	−1.4	−1.0	−3.0	−2.0
D	0.94	0.68	0.67	0.91	0.74
PSE	0.07	0.46	0.48	0.12	0.42

for temperature. The precipitation-elevation relationship is not as strong as that of temperature and elevation, thus making the creation of a single domain-wide predictive model more difficult. It was therefore necessary to compare several precipitation estimation techniques to identify the most effective domain-wide methodology. Table 3 shows the results of testing 2 domain-wide precipitation models on independent data, one using IDW interpolation and one using kriging. In both cases, the models explain less variance in the precipitation data than did the temperature model. A comparison of the 2 precipitation models' performances on the same independent dataset reveals only minor differences. Both models explain about 60% of the variance in the data, and the RMSE and Willmott's *D* statistics were almost identical. The MBE values were about 1 mm, with the PSE approaching 50%.

When the 2 models' regression predictions of precipitation were compared to the original station data used for model development, however, clear differences became apparent. In terms of agreement between the observed data used in the regression modeling and the corresponding grid cell predictions of precipitation, the IDW interpolation algorithm clearly outperformed kriging. There are notable disparities in the R^2 (0.90 vs 0.73), RMSE (13.71 vs 22.67), and Willmott's *D* (0.91 vs 0.74) statistics that favor the inverse distance weighting approach, and the PSE is markedly reduced (12% model-oriented error vs 42%). Only a marginal discrepancy in the MBEs tends to support the kriging approach. Based on this analysis, therefore, the use of regression plus inverse distance weighting to develop the final domain-wide precipitation model was chosen. The use of kriging for interpolating the precipitation residuals actually introduced errors by way of induced 'smoothing' of the residual data. This was partially related to software limitations that did not permit the specification of a non-zero nugget value. An improved

topographic regression model for precipitation would alleviate much of this error; in particular, kriging of the temperature residuals did not yield large errors, because the initial terrain-climate relationship was so strong. While the IDW technique does result in somewhat enhanced local signals (i.e. 'bull's-eyes'), the fact that it is less 'smooth' than kriging, and the more appropriate choice given the GIS software being used, made it the superior interpolation algorithm for this portion of the study.

Fig. 5 shows the regression-based models of mean winter temperature for Arizona and New Mexico (Fig. 5a), the observed 30 yr mean winter temperature (station data interpolated using ordinary kriging) (Fig. 5b), and the residual map of predicted minus observed temperatures (Fig. 5c). The benefits of the regression modeling are clearly evident in Fig. 5a. Localized temperature features stand out in the model, such as the Grand Canyon and Colorado River, while such features are obscured in the interpolation of solely observed data in Fig. 5b. Regional features are clear as well; for example, the warmer conditions in central Arizona and the cooler climate of the north-central New Mexico mountains are captured with striking detail. The 30 yr residuals map in Fig. 5c identifies areas where the temperature model has slightly over-predicted (e.g. the high elevations of SE Arizona) or under-predicted (e.g. the low elevations of SW Arizona). This is a common feature of regression wherein extremes in the data are not well reproduced.

Fig. 6 shows modeled 30 yr mean total winter precipitation (Fig. 6a), the 30 yr observed values of precipitation derived from a simple IDW interpolation of the station data (Fig. 6b), and the predicted minus observed residual precipitation values (Fig. 6c). As was the case for temperature, localized spatial features in the precipitation model in Fig. 6a are noticeably improved over the map of observed data only in Fig. 6b. The spa-

tial patterns on the Mogollon Rim in Arizona and along the Rio Grande River in New Mexico stand out in particular, as do the ridges of higher precipitation in the Cabeza Prieta National Wildlife Refuge of SW Arizona

(Comrie & Broyles 2002). Local features, such as precipitation in the Santa Catalina and Rincon Mountains bordering the Tucson metropolitan area, are also captured well by the model. Model predictions at the high-
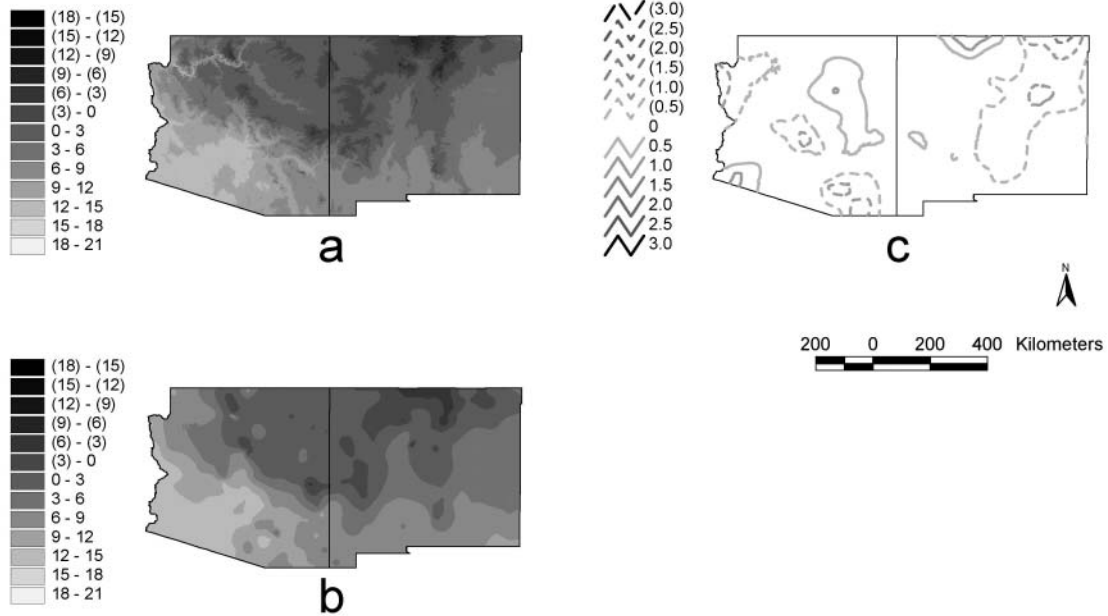

Fig. 5. Comparison of regression-based winter temperature model (a) with simple kriging of observed 30-year mean winter temperature from station locations (b). Predicted minus observed model residuals, interpolated for the study area, are shown in (c), representing variance in the data not captured by the temperature model. Model under-predictions are shown as negative (dashed) values and over-predictions (solid) as positive values. All values are in °C

Fig. 6. Comparison of regression-based winter precipitation model (a) with simple IDW of observed 30-year mean winter precipitation from station locations (b). Predicted minus observed model residuals, interpolated for the study area, are shown in (c), representing variance in the data not captured by the precipitation model. Model under-predictions are shown as negative (dashed) values and over-predictions (solid) as positive values. All values are in mm
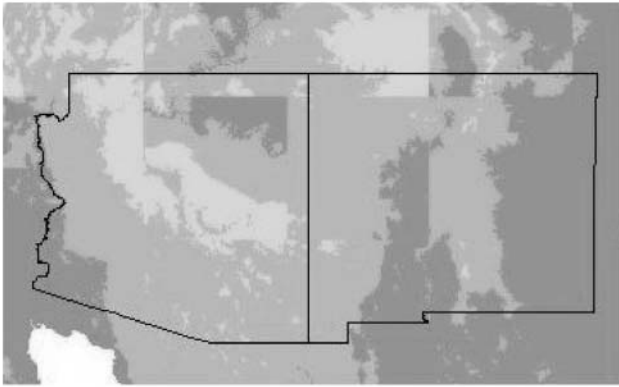
Fig. 7. Difficulties encountered in 'edge-matching' sub-domain regression models for winter precipitation are evident in the sharp lines along sub-domain model boundaries

est elevations in Arizona and New Mexico are not altogether reliable, and they reflect the limitations of both the regression methodology and the reporting station distribution; an inflated regression approach may permit better climate estimates at these extreme elevations. This is less of a problem for the temperature models, for which elevation is the dominant factor in explaining temperature variances. Like the temperature model, the precipitation residual map in Fig. 6c shows areas of slight over-prediction (SW New Mexico) and under-prediction (Mogollon Rim in central Arizona), although these residuals do not reveal a consistent single cause that might be added to the model.

5.2. Alternative modeling methodologies for precipitation

It was thought that reducing the areal extent of the region being modeled, then aggregating smaller-scale regressions to the domain scale, could result in improvements to the 1 × 1 km precipitation model. It was desirable to maintain a consistent shape throughout the methodology (i.e. square-shaped cells), particularly for ease of use within a GIS. It was also necessary to retain a minimum number of meteorological stations in each sub-domain region to ensure the development of statistically significant regression models. Thus, a 3° latitude × 3° longitude grid in the form of a 3 × 5 matrix was chosen to delineate the sub-domain regions, with 20 to 45 stations per sub-domain. This grid is visible in Fig. 7.

A comparison of the error statistics from both the domain-wide and sub-domain regression models produced mixed results (Table 4). The coefficient of determination (R^2) is higher for the domain-wide model, while for the sub-domain models the Willmott's D statistic increases and the PSE is lessened. Arguments could be made for the use of both models; however, the difficulty of edge-matching the various sub-domain regions in a GIS rendered the sub-domain models impractical (Fig. 7). Stark boundaries between the sub-domain regions are clearly evident and represent an unrealistic application of the sub-domain models on a domain-wide scale.

As a potential mechanism to overcome the edge-matching issues presented by the sub-domain regressions approach, the interpolation of sub-domain regression coefficients from each 3° latitude × 3° longitude grid cell centroid was thought to be beneficial in 2 important ways. Smooth interpolations of each predictor coefficient may provide better regression approximations for stations lying near the boundaries of sub-domain regions, due to the influence of topographic forcings from multiple regions. A second benefit was the ability to examine the spatial patterns of the regression coefficients themselves, as they demonstrate the changing influences of the various predictor variables over space.

As seen in Table 4, however, this method of improving upon the domain-wide and sub-domain regression models was not successful. While the R^2 value pooled from the various sub-domain regions did increase, model errors as indicated by the RMSE, MBE, Will-

Table 4. Error statistics from testing 3 different regression models for winter precipitation: (a) domain-wide model; (b) 15 sub-domain models; (c) interpolation of regression coefficients from sub-domains in (b)

	(a) Domain-wide regression model	(b) Sub-domain regression models (pooled)	(c) Interpolated regression coefficients (pooled)
N	113	113	113
Average observation (mm)	108.71	107.95	80.52
Average prediction (mm)	108.46	108.46	108.46
R^2	0.44	0.37	0.58
RMSE (mm)	10.92	11.68	49.28
MBE (mm)	0.25	−0.51	−27.94
D	0.51	0.62	0.35
PSE	0.54	0.35	0.60

mott's *D*, and PSE statistics also increased noticeably. This can likely be attributed to the fact that not all of the regression coefficients were significant at the $\alpha = 0.05$ level in each sub-domain model. It was impractical to omit less significant coefficients from the various interpolations in order to achieve continuous spatial coverage across the domain.

Based on a comparison of the 3 approaches to modeling winter precipitation, therefore, the earlier domain-wide winter precipitation regression model with IDW was selected for use in this analysis. Its statistical accuracy was similar to the later methodologies, while being far simpler to develop and apply. In addition, it allowed for easy comparison of modeled temperature and precipitation datasets due to their identical domain size.

5.3. Interannual winter temperature and precipitation maps

Figs. 8 & 9 illustrate the results of the methodology used to create interannual time series of winter temperature and precipitation maps. Space constraints prevent us from showing the total range of results here; however, the full set of winter temperature and precipitation maps from 1961–1999, accompanied by maps of anomalies and the comparative interpolations of ob-

served data, can be found at a supporting website (http://geog.arizona.edu/~comrie/climas/anim.htm). The winters of 1964 and 1983 were selected as examples to be shown in this paper because of their marked difference in precipitation regimes (1964 being dry, 1983 being wet); for consistency, the temperature maps from these winters are shown as well.

The predicted temperature maps in Fig. 8a,b possess a high degree of detail. The corresponding anomaly maps identify deviations from the 30 yr temperature means for each of the 2 winters. The winter of 1964 proved to be an anomalously cool winter across nearly all of the 2-state area, as indicated both by the anomaly map in Fig. 8c and the modeled temperature values. The 1983 winter, considered to have occurred under the influence of an El Niño event, did not exhibit major sub-regional temperature anomalies, as shown in Fig. 8d.

Similar to temperature, the modeled precipitation maps shown in Fig. 9a,b possess excellent spatial clarity, and the anomaly maps facilitate the discernment of areas of above- and below-normal precipitation during each of the 2 winter seasons. The winter of 1964 was anomalously dry across the majority of the study area, as illustrated in Fig. 9c, while Fig. 9d shows that 1983 was much wetter than normal, likely due to the influence of that winter's El Niño conditions.
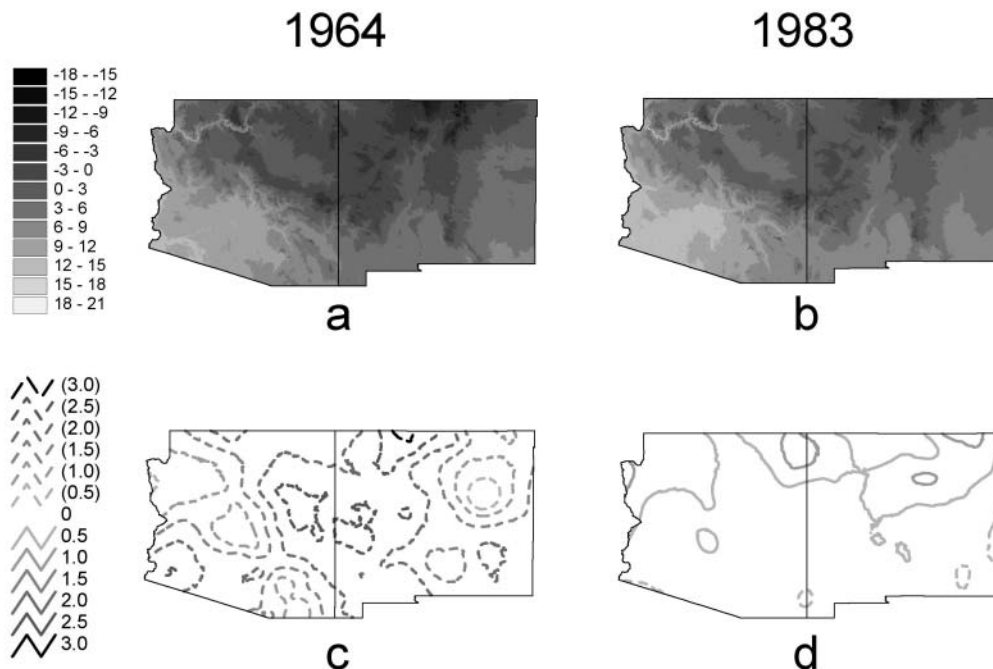

Fig. 8. Examples of individual winter temperature maps, (a) 1964 and (b) 1983, generated by comparing the long-term temperature model to individual winter deviations interpolated from station locations (c and d). Negative (dashed) anomalies indicate regions of lower than average temperatures, while positive (solid) anomalies indicate higher than average temperatures. All units are in °C
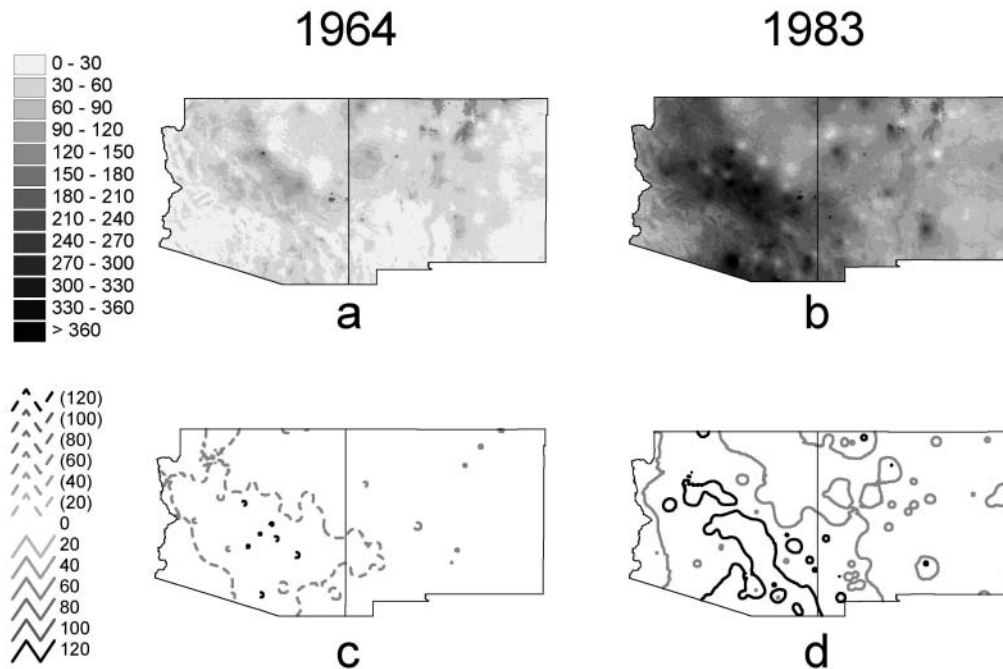
Fig. 9. Examples of individual winter precipitation maps, (a) 1964 and (b) 1983, generated by comparing the long-term precipitation model to individual winter deviations interpolated from station locations (c and d). Negative (dashed) anomalies indicate regions of drier than average precipitation, while positive (solid) anomalies indicate above-average precipitation. All units are in mm

6. CONCLUDING DISCUSSION

The 2 broad objectives of this paper were the development of a statistical modeling technique suitable for constructing gridded climate datasets for a topographically varying domain and the production of mean and interannual winter temperature and precipitation datasets for the SW United States. A number of issues central to model development were addressed. These included identifying the amount of spatial variance inherent in temperature and precipitation station data, determining the most effective algorithms for interpolating model residuals, and generating time series of temperature and precipitation maps using interannual station anomalies in conjunction with the long-term modeled means.

A domain-wide winter temperature model, utilizing regression and kriging for model residuals, explained 98% of the spatial variance in an independent sample. As the physical relationship between temperature and elevation is highly consistent, this level of explained temperature variance using topographic predictor variables was expected. Model residuals were interpolated with kriging, as the 'smoothing' effects of this algorithm complement the continuous spatial nature of temperature.

For winter precipitation, a domain-wide topographic model explained 63% of the spatial variance during

cross-validation. Both kriging and inverse distance weighting (IDW) were tested as interpolation schemes for the model residuals, with IDW better capturing the localized spatial variability of precipitation patterns. Two alternative methodologies, both involving the modeling of precipitation at the sub-domain level, were tested in an attempt to improve upon the initial model's degree of explained variance. These alternatives, however, introduced additional errors into the methodology, such as problems incurred when attempting to 'edge match' sub-domain models or interpolate less statistically significant beta values ($\alpha >$ 0.05). In the end, the single domain-wide precipitation model proved as statistically robust as the alternatives while being far easier to implement.

Once the baseline models of temperature and precipitation were constructed, time series of winter temperature and precipitation were calculated from 1961–1999. Observed station values for each winter in the time series were compared to their corresponding modeled values, and the anomalies were interpolated using kriging (temperature) or IDW (precipitation). Both the anomaly maps and resulting maps of actual winter temperature and precipitation are valuable tools for climate analysis and applications. Because the elevation of some regions within the study area fell outside the range of station elevations used in model development, temperature and precipitation predic-

tions for these extreme areas is not altogether reliable; however, these regions represent only a fraction of the total spatial coverage of the study area.

Several improvements to the modeling analysis are possible. Other studies have introduced additional predictor variables into stepwise linear regression methodologies (e.g. Goodale et al. 1998, Agnew & Palutikof 2000), such as distance from a coastal area, direction to the nearest large water source, etc. In the arid SW United States these connections were less explicit, and were likely captured in the latitudinal and longitudinal gradients as well as the Mogollon Rim predictor term in the precipitation model. However, a 'land cover' variable, more explicit consideration of snowfall versus rainfall, and possibly other proxy variables might improve the regression models and potentially remove further topographic variance from the data.

The clustering of data points around metropolitan areas and the lack of coverage across large sections of both Arizona and New Mexico present some difficulties for interpolating data from observed stations values, especially precipitation. Yet this difficulty was one of the main motivations for developing gridded datasets with more robust spatial coverage. The choice of a 3° latitude × 3° longitude grid to delineate sub-domain precipitation regions was, in some respects, arbitrary in nature. A principal-components-based regionalization, such as that conducted by Comrie & Glenn (1998), provides an alternative method of dividing a domain into subsets. However, the joining together of non-symmetric regions for the purposes of interpolation is not easily accomplished in a GIS, and proved too cumbersome for the scope of this paper. The prospect of utilizing sub-domain level regression modeling holds some promise, as indicated by the accuracy assessment outlined in Table 4. The difficulty of edge-matching adjoining regions of separate model predictions needs to be overcome, however, possibly through the use of algorithms borrowed from the GIS and remote-sensing literatures.

The question of choosing the 'best' resolution at which to model climate variables represents a much-debated topic in geographic information science, and intersects with research into the modifiable areal unit problem (Openshaw 1981, Fotheringham & Wong 1991, Malanson & Armstrong 1997). Specifically, it is difficult to discern scales of action for temperature and precipitation events without initial high-resolution datasets and extensive analyses of variograms and other statistical tools. A thorough investigation of the modifiable areal unit problem is beyond the scope of this paper, but we are investigating it separately as it pertains to climate modeling in Arizona and New Mexico.

Several types of future investigations may complement this research. From a methodological perspec-

tive, the datasets could be extended temporally to include winters prior to 1961, as long as data coverage is sufficient for the generation of individual winter season maps. Temporally, the regression methodology developed in this research can be expanded to include the spring, summer, and autumn seasons and could also be used to produce monthly maps in addition to seasonal datasets. The use of terrain variables to predict temperature and precipitation patterns can be applied to areas outside Arizona and New Mexico, although the predictor variables themselves may change for different regions of study. A 'windowing' technique, such as that used in the PRISM approach (Daly et al. 1994), could alleviate problems encountered with changing predictor variables over space.

LITERATURE CITED

Agnew MD, Palutikof JP (2000) GIS-based construction of baseline climatologies for the Mediterranean using terrain variables. Clim Res 14:115–127

Beek EG, Stein A, Janssen LLF (1992) Spatial variability and interpolation of daily precipitation amount. Stochast Hydrol Hydraul 6:304–320

Bigg GR (1991) Kriging and intraregional rainfall variability in England. Int J Climatol 11:663–675

Burrough PA (1986) Principles of geographical information systems for land resources assessment. Oxford University Press, New York

Burrough PA, McDonnell R (1998) Principles of geographic information systems. Oxford University Press, New York

Comrie AC, Broyles B (2002) Variability and spatial modeling of fine-scale precipitation data for the Sonoran Desert of Southwest Arizona. J Arid Environ 50:573–592

Comrie AC, Glenn EC (1998) Principal components-based regionalization of precipitation regimes across the southwest United States and northern Mexico, with an application to monsoon precipitation variability. Clim Res 10: 201–215

Daly C, Neilson RP, Phillips DL (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J Appl Meteorol 33:140–158

Davis JC (1986) Statistics and data analysis in geology. John Wiley & Sons, New York

de Montmollin FA, Olivier RJ, Simard RG, Zwahlen F (1980) Evaluation of a precipitation map using a smoothed elevation-precipitation relationship and optimal estimates. Nord Hydrol 11:113–120

Dodson R, Marks D (1997) Daily air temperature interpolated at high spatial resolution over a large mountainous region. Clim Res 8:1–20

Ffolliott PF, Gottfried GJ, Baker MB Jr (1989) Water yield from forest snowpack management: research findings in Arizona and New Mexico. Water Resour Res 25:1999–2007

Fotheringham AS, Wong DWS (1991) The modifiable areal unit problem in multivariate statistical analysis. Environ Plan 23:1025–1044

Goodale CL, Aber JD, Ollinger SV (1998) Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. Clim Res 10:35–49

Hevesi JA, Istok JD, Flint AL (1992a) Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I. Structural analysis. J Appl Meteorol 31:661–676

Hevesi JA, Flint AL, Istok JD (1992b) Precipitation estimation in mountainous terrain using multivariate geostatistics. Part II. Isohyetal maps. J Appl Meteorol 31:677–688

Ishida T, Kawashima S (1993) Use of cokriging to estimate surface air temperature from elevation. Theor Appl Climatol 47:147–157

Karnieli A (1990) Application of kriging technique to areal precipitation mapping in Arizona. GeoJournal 22:391–398

Kilsby CG, Cowpertwait PSP, O'Connell PE, Jones PD (1998) Predicting rainfall statistics in England and Wales using atmospheric circulation variables. Int J Climatol 18:523–539

Kurtzman D, Kadmon R (1999) Mapping of temperature variables in Israel: a comparison of different interpolation methods. Clim Res 13:33–43

Malanson GP, Armstrong MP (1997) Issues in spatial representation: effects of number of cells and between-cell step size on models of environmental processes. Geogr Environ Model 1:47–64

McBratney AB, Webster R (1986) Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. J Soil Sci 37:617–639

Openshaw S (1983) The modifiable areal unit problem. Concepts Tech Modern Geogr 38:41

Press WH (1988) The art of scientific computing. Cambridge University Press, New York

Prudhomme C, Reed DW (1999) Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. Int J Climatol 19:1337–1356

Serreze MC, Pulwarty RS, Clark MP, Armstrong RL, McGinnis DA (1999) Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. Water Resour Res 35:2145–2160

Spronken-Smith RA, Oke TR, Lowry WP (2000) Advection and the surface energy balance across an irrigated urban park. Int J Climatol 20:1033–1047

Willmott CJ (1981) On the validation of models. Phys Geogr 2:184–194

Willmott CJ, Matsuura K (1995) Smart interpolation of annually averaged air temperature in the United States. J Appl Meteorol 34:2577–2586

Willmott CJ, Ackleson SG, Davis RE, Feddema JJ, Klink KM, Legates DR, O'Donnell J, Rowe CM (1985) Statistics for the evaluation and comparison of models. J Geophys Res 90:8995–9005