A comparison of three automated precipitation simulation models : ANUSPLIN, MTCLIM-3D, and PRISM
by Sara Teresa Stillman

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Earth Sciences
Montana State University

Abstract:
A comparison of the ANUSPLIN, MTCLIM-3D, and PRISM model performance is needed to assist users with appropriate model selection and elucidate potential differences. The models employ different techniques to develop gridded precipitation surfaces from published climate station (point) data and digital elevation models (DEMs) using the same monthly and annual data sets to determine whether the predicted precipitation surfaces are hydrologically reasonable over a region which contains a diverse physiography and produces a wide range of precipitation regimes. Mean monthly and annual precipitation estimates were prepared for the Bozeman, Billings, Ashton, and White Sulphur Springs 1 x, 2° topographic quadrangles in southwestern Montana and the Cody quadrangle in Wyoming for the 1961-90 data period. Input data included monthly precipitation data from 258 weather stations and a 0.5 km square-grid DEM derived from the appropriate 3 arc-second USGS DEMs with ANUDEM.

The models generated statistically similar results. The mean annual precipitation predictions for the 20 (randomly selected) withheld stations were accepted as statistically similar to the observed data at the 0.05 significance level. ANUSPLIN produced slightly higher mean annual estimates (5.6% and 4.0% higher than MTCLIM-3D and PRISM, respectively), and tended to overestimate precipitation at the 20 withheld stations. This model also generated slightly higher mean absolute errors (MAE) compared to the other two models which tended to underestimate precipitation at the 20 withheld stations. The largest differences between the model predictions occur in high elevation areas (e.g. Absarokas, Tetons) where a lack of climate stations and highly variable precipitation patterns complicate the interpolation process. Similar results suggest model selection should be based on ease of use and efficiency.

The MTCLIM-3D mean values are higher in winter and early spring while PRISM predictions are greater in the late spring-summer months and ANUSPLIN generally has the lowest predictions overall although the differences are minimal. The predictions fell in the same 25 mm classes in over 80% of the DEM cells in 28 out of 36 monthly surface comparisons. The largest differences occurred in the late fall and winter. The largest MAE and bias estimates were generated for all three models in these months. In addition, predictions were different from the observed data at the 20 withheld stations at the 0.05 significance level in November for ANUSPLIN and MTCLIM-3D, December for ANUSPLIN and PRISM, and February for ANUSPLIN. Relatively low agreement between the summed monthly surface and annual surface values for all three models demonstrate the importance of incorporating snow course data.' The models require less climate knowledge compared to the hand-contouring method and provide error estimates as well as precipitation estimates that can be accessed directly by modem geographic information systems. Increased numbers of climate stations in high elevations and more precise measurements of station locations and elevations would improve model predictions in the northern Rocky Mountains.

A COMPARISON OF THREE AUTOMATED
PRECIPITATION SIMULATION MODELS:
ANUSPLIN, MTCLIM-3D, AND PRISM

by

Sara Teresa Stillman

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Earth Sciences

MONTANA STATE UNIVERSITY-BOZEMAN
Bozeman, Montana

April 1996

APPROVAL

of a thesis submitted by

Sara T. Stillman

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

4/1/96	John P. Wil
Date	Chairperson, Graduate Committee

Approved for the Major Department

4/24/96	W. A. Marcus
Date	Head, Major Department

Approved for the College of Graduate Studies

6/10/96	R. Brown
Date	Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature Sara T. Stillman

Date 5/18/96

ACKNOWLEDGMENTS

I am extremely grateful to Dr. John Wilson for his invaluable assistance and innovative ideas and providing both the opportunity and funding not only for my graduate research, but for my attendance at the 1996 International Conference on Integrating Environmental Modeling and GIS and other conferences. I would also like to thank Drs. Stephan Custer and Andrew Marcus for serving on my committee. Mr. Phillip Farnes provided insightful contributions into climate data quality and collection and the hand-drawn precipitation maps. My appreciation is extended to Professor Chris Daly who performed the PRISM model runs, Peter Thornton who performed the MTCLIM-3D runs, and to Dr. Michael Hutchinson for his assistance with the ANUSPLIN runs.

This project could not have been completed without the continued support of the GIAC staff, my family, and fellow graduate students, Skip Repetto, Andrea Wright, and Mandy Lineback.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

A comparison of the ANUSPLIN, MTCLIM-3D, and PRISM model performance is needed to assist users with appropriate model selection and elucidate potential differences. The models employ different techniques to develop gridded precipitation surfaces from published climate station (point) data and digital elevation models (DEMs) using the same monthly and annual data sets to determine whether the predicted precipitation surfaces are hydrologically reasonable over a region which contains a diverse physiography and produces a wide range of precipitation regimes. Mean monthly and annual precipitation estimates were prepared for the Bozeman, Billings, Ashton, and White Sulphur Springs 1 x 2° topographic quadrangles in southwestern Montana and the Cody quadrangle in Wyoming for the 1961-90 data period. Input data included monthly precipitation data from 258 weather stations and a 0.5 km square-grid DEM derived from the appropriate 3 arc-second USGS DEMs with ANUDEM.

The models generated statistically similar results. The mean annual precipitation predictions for the 20 (randomly selected) withheld stations were accepted as statistically similar to the observed data at the 0.05 significance level. ANUSPLIN produced slightly higher mean annual estimates (5.6% and 4.0% higher than MTCLIM-3D and PRISM, respectively), and tended to overestimate precipitation at the 20 withheld stations. This model also generated slightly higher mean absolute errors (MAE) compared to the other two models which tended to underestimate precipitation at the 20 withheld stations. The largest differences between the model predictions occur in high elevation areas (e.g. Absarokas, Tetons) where a lack of climate stations and highly variable precipitation patterns complicate the interpolation process. Similar results suggest model selection should be based on ease of use and efficiency.

The MTCLIM-3D mean values are higher in winter and early spring while PRISM predictions are greater in the late spring-summer months and ANUSPLIN generally has the lowest predictions overall although the differences are minimal. The predictions fell in the same 25 mm classes in over 80% of the DEM cells in 28 out of 36 monthly surface comparisons. The largest differences occurred in the late fall and winter. The largest MAE and bias estimates were generated for all three models in these months. In addition, predictions were different from the observed data at the 20 withheld stations at the 0.05 significance level in November for ANUSPLIN and MTCLIM-3D, December for ANUSPLIN and PRISM, and February for ANUSPLIN. Relatively low agreement between the summed monthly surface and annual surface values for all three models demonstrate the importance of incorporating snow course data.

The models require less climate knowledge compared to the hand-contouring method and provide error estimates as well as precipitation estimates that can be accessed directly by modern geographic information systems. Increased numbers of climate stations in high elevations and more precise measurements of station locations and elevations would improve model predictions in the northern Rocky Mountains.

CHAPTER 1

INTRODUCTION

Scope and Purpose

Estimates of the amount and spatial distribution of rainfall data provide critical inputs for regional resource assessments and environmental modeling applications. Average monthly and annual precipitation data are required to evaluate potential land uses, water supply, drought hazards, and fire risk. Many silviculture, insect and disease, and hydrological models also need spatially variable precipitation estimates as inputs (eg: Lebel et al. 1987; Running et al. 1987; Hungerford et al. 1989; Caprio et al. 1990; Nielsen et al. 1990; Phillips et al. 1992; Wilson et al. 1993). Precipitation variability is a critical factor for determining the water budget of a region, particularly in the western United States where a large percentage of the total available water may be contained in the snowpack at high elevations and spatial variability can be extremely large over very small distances (<5 km) (Johnson and Hanson 1995). Knowledge of precipitation distribution can also assist in reservoir placement (Giorgi et al. 1992), irrigation, and overall watershed management (Basist et al. 1994).

Unfortunately, precipitation is assumed to be uniform for most applications due to low data density and computational and statistical difficulties associated with extrapolation (Anderson 1973; Johnson and Hanson 1995). Interpolation of precipitation data to unmeasured locations is particularly difficult in complex mountainous regions and at the mountain-plains interface where highly variable spatial patterns of precipitation are produced (Peck and Brown 1962; Doesken et al. 1989; Daly and Neilson 1992; Phillips et al. 1992). Common obstacles include acquisition of a complete and continuous data set of rainfall measurements, determination of accurate and appropriate precipitation/elevation lapse rates, incorporation of orographic effects, selection of the appropriate grid resolution, and accounting for the additional data complexity caused by the large temporal and spatial variation in the distribution of rainfall.

The overall goal of this project is to compare the performance of three models using the same monthly and annual data sets to determine whether the predicted precipitation surfaces are hydrologically reasonable over a region which contains a diverse physiography and produces a wide range of precipitation regimes. The models are ANUSPLIN (Hutchinson 1989a, 1995, 1996), MTCLIM-3D (Running and Thornton 1996), and PRISM (Daly et al. 1994; Daly and Taylor 1996). These models employ different techniques to develop gridded precipitation surfaces from point data, and several model runs were performed to:

(1) Determine if the models generate the same mean annual precipitation surfaces using climate station records and 3-arc-second DEMs for the Bozeman, Billings, Ashton, and White Sulphur Springs 1 x 2° quadrangles in southwest Montana and the Cody quadrangle in Wyoming. A randomly selected subset of the climate station data will be withheld from these model runs and the predicted values will be compared with measured values to assess model performance.

(2) Determine if the models generate the same mean monthly precipitation surfaces

using monthly climate station records. The same subset of data withheld in Objective 1 will be withheld from these model runs and used to evaluate model performance.

(3) Determine if the models generate the same mean annual precipitation surfaces as the professionally hand-drawn 1961-1990 precipitation contour maps produced by Phillip Farnes (USDA Soil Conservation Service (SCS) 1977). No data will be withheld for these model runs. Because the "true" average annual precipitation at many locations is not known for Montana (or any other state), the veracity of a computer-generated map is best assessed with respect to a tested and widely used reference map (Custer et al. 1996).

These answers are needed because it is essential that models designed for similar purposes be compared to other models and tested against observed measurements prior to producing baseline data for land and resource management decisions.

Spatial Interpolation Techniques Used in Climatology

Climatological work examining spatial variability of precipitation has taken one of two approaches: (1) a statistical approach where precipitation is distributed by use of spatial interpolation methods; or (2) a physically-based modeling approach in which precipitation is "dynamically" or deterministically simulated (Johnson and Hanson 1995). Dynamic models often lack the finer spatial-scale resolution required for estimating precipitation in mountainous regions and require substantial computer resources for long-term simulations. Statistical methods that use topographic variables as predictors can effectively estimate the spatial distribution of mean annual precipitation in mountainous regions (Spreen 1947; Hutchinson 1973; Danard 1976; Vidal and Varas 1982; Basist et al. 1994). The remainder of the focus in this study will therefore be on statistical models.

A large number of spatial interpolation methods have been proposed for estimation of precipitation in unsampled areas from measured data (Creutin and Obled 1982; Tabios and Salas 1985; Hutchinson 1991c; Phillips et al. 1992; Dingman 1994). These techniques incorporate not only different statistical methods, but also vary in terms of computational complexity, data requirements, determination of lapse rates, and their ability to incorporate additional variables. As a result, very different estimates may be derived for the same location on different computer-generated maps. The commonly used techniques have been classified as local interpolation, moving average interpolation, spline, and kriging methods (Moore and Hutchinson 1991).

Local Interpolation Methods

Local interpolation methods include the use of triangulation, simple bivariate analysis, trend surface analysis, and similar methods that fit a polynomial equation or another simple function to a subregion to create a surface, with its complexity adjusted by changing the order of the polynomial (Shaw and Lynn 1972; Akima 1978). The resulting surfaces can suffer from somewhat arbitrary restrictions on their form and can be sensitive to the position of the data points because irregularly spaced and spurious effects may be generated away from the data points (Hutchinson and Bischof 1983). These techniques are quite complex to implement in more than two dimensions and do not lend themselves easily to the smoothing of noisy data (Moore and Hutchinson 1991). These methods can also be implemented and used to fit the data very closely, whether or not the fit is justified in terms of the amount of noise associated with the data (Hutchinson and Bischof 1983). In addition, these techniques show persistent patterns in residuals from trend surfaces of different degrees, especially in regions of extrapolation (Edwards 1972; Shulze 1976; Hughes 1982; Hutchinson and Bischof 1983;).

Moving Average Interpolation Methods

Weighted interpolation or moving average methods use a moving window technique which requires a subjective choice of a weighting function defined in terms of a user-specified radius of influence beyond which data points are ignored (Goodin et al. 1979; Lancaster 1979). Linear regression is employed to develop estimates for each grid cell. The degree of data smoothing depends on the choice of weighting function. These methods also tend to fit the data very closely (Hutchinson and Bischof 1983). In addition, the choice of an optimum radius of influence can present a problem when the density of data points varies greatly across the data network. The selection of the appropriate digital elevation model (DEM) resolution is also very important (Hutchinson 1989b). The modeling purpose should determine the grid spacing of the DEM data which directly affects the degree of topographic generalization. Studies have shown that regional means do not change with generalization; however, the regional variances change significantly for different types of terrain (Dubayah et al. 1989; Dubayah 1990; Dubayah and van Katwijk 1992).

The original MT-CLIM model (Running et al. 1987; Hungerford et al. 1989) applied an inverse distance weighting technique to extrapolate meteorological variables from a point of measurement to the site of interest, making corrections for differences in elevation, slope and aspect, based on a user-specified domain-wide lapse rate for the precipitation-elevation (P/E) relationship. Inverse distance weighting assigns each grid cell a value by summing the product of the nearest data values (within a radius of influence) and the inverse of some power of the distance between the grid cell and the nearest base station. MT-CLIM was developed for forest ecosystem modeling applications and used daily observations. Studies ranging in spatial scale from point simulations (Running and Coughlan 1988; Running 1994) to single watershed simulations (Band et al. 1991; Band et al. 1993; White and Running 1994) and regional simulations over areas of 1-2000 km^2 (Running and Nemani 1991; Nemani et al. 1993) have demonstrated the successful application of the basic MT-CLIM logic.

A modified form of this logic is used in MTCLIM-3D to generate long-term average climate surfaces. A slightly more sophisticated spatial smoothing algorithm than the algorithm used in the version of MTCLIM-3D described by Thornton et al. (1996) is used to select the appropriate degree of smoothing for the DEM required for final map production (Thornton 1996, pers. comm.).

MTCLIM-3D uses the spatial convolution of a truncated Gaussian filter with a DEM and an unlimited number of stations as the interpolation framework to generate surfaces for temperature, precipitation, humidity, and incoming shortwave radiation over large regions. The convolution of the filter with the DEM results produces a list of weights associated with observations for each grid cell. The truncation distance from the cell center, R_p, is varied as a smooth function of the local station density through the iterative estimation of local station density at each prediction point. The interpolation method for a given set of observations and a given prediction grid is defined by four parameters: I, the observation location; N, the average number of observations to be included at each point; a, a unitless shape parameter; and R_p. Additional inputs required to run MTCLIM-3D include a DEM, precipitation means with large gaps filled with data from nearby stations using linear regression or another method, and station elevations and locations. Output accuracy is estimated by monthly and annual mean absolute errors and standard errors of estimation. MTCLIM-3D was applied to the state of Montana on a 1-km resolution grid and produced mean absolute errors of 11.83 cm yr^{-1} or 20% measured as a proportion of total annual precipitation (Running and Thornton 1996).

Daly et al. (1994) have implemented an alternative moving average method in the PRISM model to generate gridded estimates of monthly and annual precipitation. The model is similar to MTCLIM-3D in that it generates a meteorological database for ecological models that integrate the role of microclimate in key forest processes such as forest evapotranspiration and photosynthesis over large areas. This topoclimatological model automatically computes climate surfaces in order to provide high resolution, terrain-sensitive daily climate data. The three main components of the conceptual framework of PRISM include the evaluation of the effects of elevation on precipitation, the determination of the spatial scales at which orographic effects are observed, and the inclusion of the effects of complex terrain on the spatial patterns of orographic regimes.

Inputs include a DEM, monthly and annual precipitation means, and user-specified minimum and maximum radii of influence, minimum and maximum slopes for the regression function, and the minimum number of stations required for the precipitation/orographic elevation (P/OE) calculation. These values can be determined for each application or left as default values (Daly and Neilson 1992).

PRISM was applied to northern Oregon and the entire western United States and produced a minimal increase in bias (4.5% versus 3.5%) and absolute errors (17% versus 16%) when applied to the larger region (Daly and Neilson 1992). High residual errors from stations in northern Oregon were attributed to either high precipitation variability on the leeward side of major mountain barriers, the altered P/E lapse rate below the crest, or the poor spatial resolution of the 5-minute DEM. Furthermore, P/OE regression functions developed from stations in relatively dry valley bottoms for regions spanning hundreds of meters of elevation may have lead to an underestimation of precipitation at the mountain crests (Daly and Neilson 1992). The Oregon Department of Water Resources uses PRISM to develop water supply forecasts and the model is currently being used by the USDA-Natural Resources Conservation Service (NRCS) (formerly the Soil Conservation Service (SCS)) to develop mean annual precipitation maps for the conterminous United States (Daly and Taylor 1996).

Splines

Splining has been developed primarily by Wahba (1980), Wahba and Wendelberger (1980) and implemented by Hutchinson (1991a, 1995). The method is related to certain forms of optimum objective analysis proposed by Gandin (1965) and described in Goodin et al. (1979) and Wahba (1990). A summary of the basic methodology of thin plate splines, focusing on climate interpolation, can be found in Hutchinson (1991a, 1995).

The ANUSPLIN suite of programs (Hutchinson and Bischof 1983; Hutchinson 1989a, 1991a, 1991b) employ a multi-dimensional Laplacian partial thin plate smoothing spline technique and exemplify this type of approach. Tri-variate thin plate splines allow for spatially variable dependence on elevation which is suitable for applications across large heterogeneous areas (Hutchinson 1991c, Hutchinson et al. 1993). ANUSPLIN is a contouring routine which fits spline surfaces to spatial data with the degree of smoothing determined by minimizing the predictive error of the surface with generalized cross validation (GCV). The method is self-validating so that an optimal smoothing parameter is derived as each data point is removed. The degree of smoothing represents a trade-off between data infidelity, as measured by the mean square residual from the data points weighted according to variance estimates, and surface roughness, as measured by the total curvature of the fitted spline. This approach has been implemented in several Australian applications (Hutchinson and Bischof 1983; Hutchinson and Johnson 1991; Hutchinson 1995), and both the prestandardized and non-diagonal error covariance models of ANUSPLIN were recently applied to 34 years of annual rainfall data in southeastern Australia. The predicted mean rainfalls were 907 mm and 914 mm with estimated standard errors of 38 mm (4% of the areal mean) and 26 mm (3% of the areal mean) respectively (Hutchinson 1995).

Kriging

Kriging and a number of more sophisticated geostatistical interpolation techniques that incorporate various dependencies on topography have been developed (Chua and Bras 1982; Hevesi et al. 1992; Phillips et al. 1992; Daly et al. 1994). Kriging is a geostatistical method in which a semi-variogram model that best fits the data is developed to arrive at optimum station weights for interpolation (Daly et al. 1994). Whereas thin plate splines are defined by minimizing the roughness of the interpolated surface with a prescribed residual from the data, kriged surfaces are defined by minimizing the variance of the error of estimation which normally depends on the preliminary semi-variogram analysis (Hutchinson and Gessler 1994). Due to the dependence on the accuracy of the fitted semi-variogram model to determine the minimum error properties of kriging, it is not immediately apparent whether kriging is a more accurate interpolator than splines (Hutchinson and Gessler 1994). The derivation of the semi-variogram in kriging is discussed in Armstrong (1984), Davis (1987), Russo and Jury (1987), Laslett and McBratney (1990), and Cressie (1991), and recent comparisons with splines are presented in Hutchinson et al. (1993), Hutchinson and Gessler (1994), Laslett (1994), and Hutchinson (1996). The formal equivalence with splines is discussed in Matheron (1981), Dubrule (1983, 1984), Watson (1984), and Wahba (1990).

Although kriging extends easily to larger data sets, its "main limitation is that it depends critically on first estimating a spatial covariance function or variogram. The method is hampered by ad hoc assumptions about the form that the variogram should take and the computational difficulties in assessing the merit of different functional forms" (Hutchinson 1991b, p. 106-7). The main advantage of splines is the lack of a requirement for prior estimation of a spatial autocorrelated covariance structure which can be difficult to estimate and validate (Hutchinson 1995). Daly et al. (1994) point out that kriging implicitly relies on the data to directly represent the spatial variability of the actual precipitation field. If the variability is not representative (which will often be the case in complex terrain), the accuracy of the resulting interpolated field will be questionable. Although the modified kriging techniques such as elevationally detrended kriging or cokriging show more topographically-related spatial patterns in complex terrain, these methods can only be applied to areas characterized by a strong overall P/E relationship. Furthermore, multiple semi-variograms may be needed to estimate precipitation at various time periods because of variations in the relative importance of different precipitation sources.

Study Area Description

A robust test of a model's predictive capabilities is provided by the generation of precipitation surfaces for a region which incorporates the Rocky Mountains and northern Great Plains which display a large variety of precipitation regimes. The Bozeman, Billings, Ashton, and White Sulphur Springs U.S.G.S. 3-arc-second quadrangles in southwest Montana and the Cody Wyoming quadrangle cover 47,328 km^2 and contain numerous national forests (NF) and parks, including all of Yellowstone National Park (YNP), as well as several wilderness areas. Split by the Continental Divide, the region includes multiple watersheds and mountain ranges, rangeland, cropland, and intermontane plains and valleys (Figure 1). The elevations (as recorded on the DEMs) range from 2840 m above sea level northeast of Billings to 4197 m above sea level (Grand Teton, WY).

The amount of precipitation that is received depends on the orientation of nearby mountain ranges, elevation, rainshadows, the storm direction, intensity, and time of year. Annual precipitation across the study area is as high as 152.1 cm (Black Bear, 2484 m) on the Yellowstone Plateau, 130.8 cm in the Madison Range (Carrot Basin, 2743 m) and 148.9 cm at Fisher Creek (2774 m) in the Beartooth/Absaroka mountains, and as low as 17.0 cm (Basin, 1170 m) in the plains region near Greybull WY and 30.0 cm at Canyon Ferry Dam (1119 m) in the northeast corner.

The disparate precipitation distribution between valley and mountainous areas is exemplified in the Gallatin Range where the valley stations, Ennis (1510 m) and Jack Creek (2027 m), receive 32.4 cm and 36.8 cm, respectively, and the Sentinel Creek (2530 m) and Bear Basin (2484 m) mountain stations receive 91.4 cm and 116.8 cm, respectively. The Bridger Bowl summit station (2210 m) receives 136.6 cm annually whereas the Bozeman 12NE station at the base (1814 m) receives 89.2 cm and stations on the eastern side of the Crazy Mountains, Melville (1635 m) and Wilsall (1539 m), receive only 42.4 cm and 39.2 cm, respectively due to the rainshadow effect. The rainshadow effect, in which the intensity of a storm dissipates as it rises over a mountain range, is further illustrated when the snowfall in the Madison Drainage (60-67 cm) is compared to that of the Gallatin Drainage (44-49 cm).