

## SPATIAL EVALUATION OF PRECIPITATION IN TWO LARGE WATERSHEDS IN NORTH-CENTRAL ARIZONA

Boris Poff,<sup>1,2</sup> Assefa S. Desta,<sup>1</sup> and Aregai Tecle<sup>1</sup>

The USDA Forest Service established the Beaver Creek Experimental Watershed Pilot Project in north-central Arizona in 1957 and operated it until 1982. After the Forest Service discontinued the project in 1982, Northern Arizona University's School of Forestry continued to monitor and do research in two of the largest watersheds, known as Woods Canyon and Bar M. From 1982 to 1995 only streamflow data are available for these watersheds. A weak correlation between streamflow and precipitation for both watersheds was determined by analyzing the USDA data from 1961 to 1982, using spatial interpolation methods in ArcGIS. However, this study is a work in progress and the authors are in the process of using different spatial interpolation methods in ArcGIS to determine better regression functions between streamflow and precipitation. The objective of this study is to evaluate existing historical data using newly available techniques and technology to obtain new results. Methods used so far, in ArcGIS, are four different spatial interpolation methods, namely Inverse Distance Weighting, Radial Basis Function, and Global and Local Interpolation.

### BACKGROUND AND STUDY SITE

Originally, between 1957 and 1962, the USDA Forest Service set up 20 pilot watersheds in the Beaver Creek Experimental Watershed in north-central Arizona. Of the 20 watersheds, 18 ranged from 27 to 824 ha. The remaining two watersheds, known as Bar M (Watershed 20) and Woods Canyon (Watershed 19), are much larger with areas of 6620 and 4893 ha, respectively (Baker and Ffolliott 1999; see Figure 1). Seventeen of the watersheds, including the largest two, are located in the ponderosa pine forest ecosystem. This study uses data from 39 of the 72 rain gauges the USDA

Forest Service had in place to collect rainfall data in the experimental project (see Figure 1). The objective of the Beaver Creek Experimental Watershed Pilot Project was to determine the impacts of forest manipulation on water yield. The ultimate desire was to increase the amount of water available from the forest regions of Arizona for the growing population in the state.

### METHODOLOGY

The data used in the evaluation process for Woods Canyon (WS19) consisted of precipitation data from 39 rain gauges in the Beaver Creek Experimental Watershed for the time period from 1961 to 1981. Streamflow data from the Woods Canyon stream gauging station for the same time period were used as well. The data are available at a Web site maintained by the Rocky Mountain Research Station and the University of Arizona in Tucson (<http://ag.arizona.edu/OALS/watershed/index.html>). The data used in the evaluation process for Bar M (WS20) consisted of the same precipitation data used for Woods Canyon from 39 rain gauges in the Beaver Creek Experimental Watershed from 1961 to 1981, and other 1961–1990 precipitation data from seven additional weather stations east and south of the Beaver Creek Watershed. This became necessary because the boundaries of Bar M extend farther east than the location of any of the 39 Beaver Creek rain gauging stations. The streamflow data from the Bar M stream gauging station were used for the time period of 1961–1981. The streamflow and precipitation data from the 39 Beaver Creek Experimental Watershed rain gauging stations were downloaded from the watershed management Web site given above. Rain data from the additional seven weather stations were taken from the Global Historic Climate Network data, National Climate Data Center, National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

A Trimble GeoExplorer3 handheld GPS unit was used to determine the boundary for the two

<sup>1</sup>USDA Forest Service Rocky Mtn. Research Station, Flagstaff

<sup>2</sup>School of Forestry, Northern Arizona University, Flagstaff

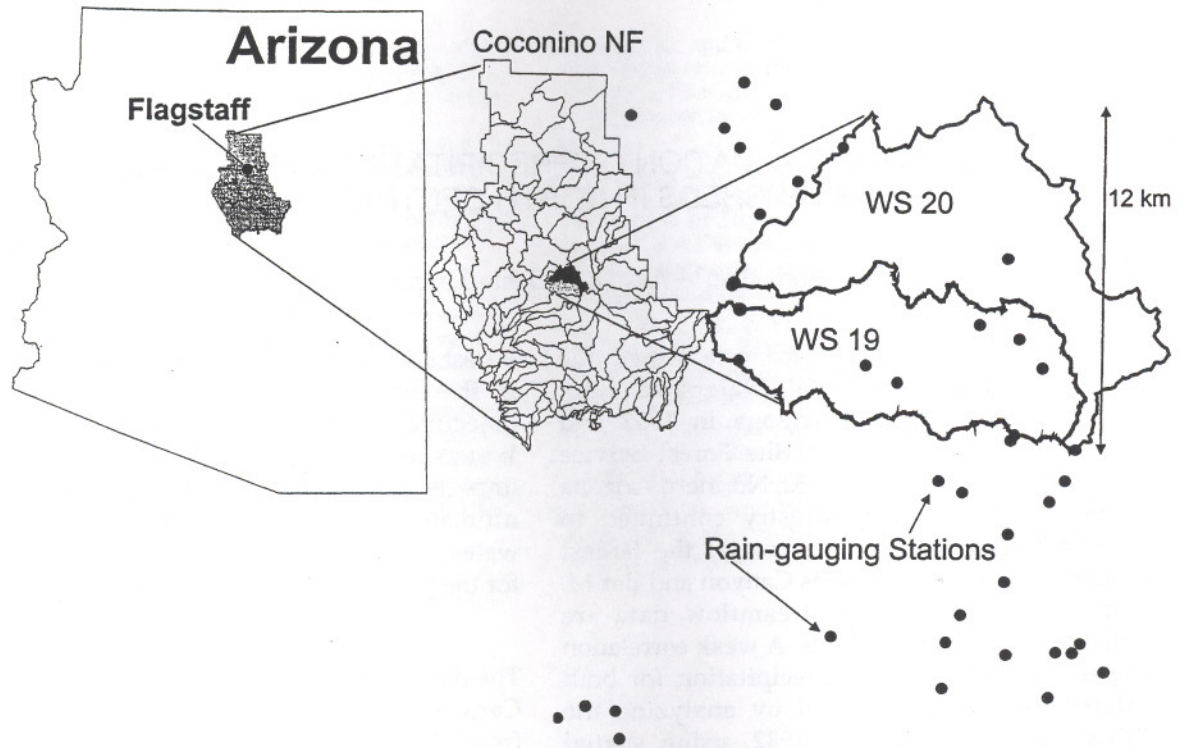


Figure 1. The location of the study sites and of the 39 rain gauges used in this study, in the Beaver Creek Experimental Watershed, Arizona.

watersheds (Poff et al. 2005). The GPS data were recorded in the Universal Transverse Mercator (UTM) system and projected using the North American Datum (NAD) of 1927.

Precipitation data were converted into millimeters where necessary (NOAA data were already in mm) and the rainfall values (1961–1981) were averaged for each calendar month for each rain gauging station. This produced 12 average values for each station—one for each month of the year. These average values were imported into ArcMap, where their UTM coordinates were calculated. After the gauging stations were spatially referenced, Inverse Distance Weighting (IDW), Radial Basis Function (RBF), and Local and Global Interpolation methods were used to create a precipitation interpolation map for each month. IDW, which is a local interpolator, weighs closer points greater than those farther away, applying the fundamental assumption for spatial interpolation (ESRI 2001), which is the first law of geography: Everything is related to everything else, but near things are more related than distant things (Tobler 1970). The second

assumption necessary to using IDW is that this phenomenon is modeled as a spatially continuous dataset. Even though precipitation values are usually collected as discrete point data, it is safe to make this assumption with such a dataset.

Even though the data tend to be normally distributed, there is one outlier, which can be seen in the histogram Voronoi map. Including this outlier in the dataset produced a Kurtosis value of 5.21 and a skewness coefficient of  $-1.00$ . The Voronoi map also indicates clustering of values around several locations, which given the pattern of precipitation is to be expected.

The average rainfall value for August for Woods Canyon was used for a more in-depth investigation after the interpolation output for each month using the IDW method had been reviewed. Because these values were collected during the monsoon season they provide adequate data for statistical analysis.

Global and local trends such as latitude (global), as well as elevation and aspect (local), can influence the spatial distribution of precipitation. It

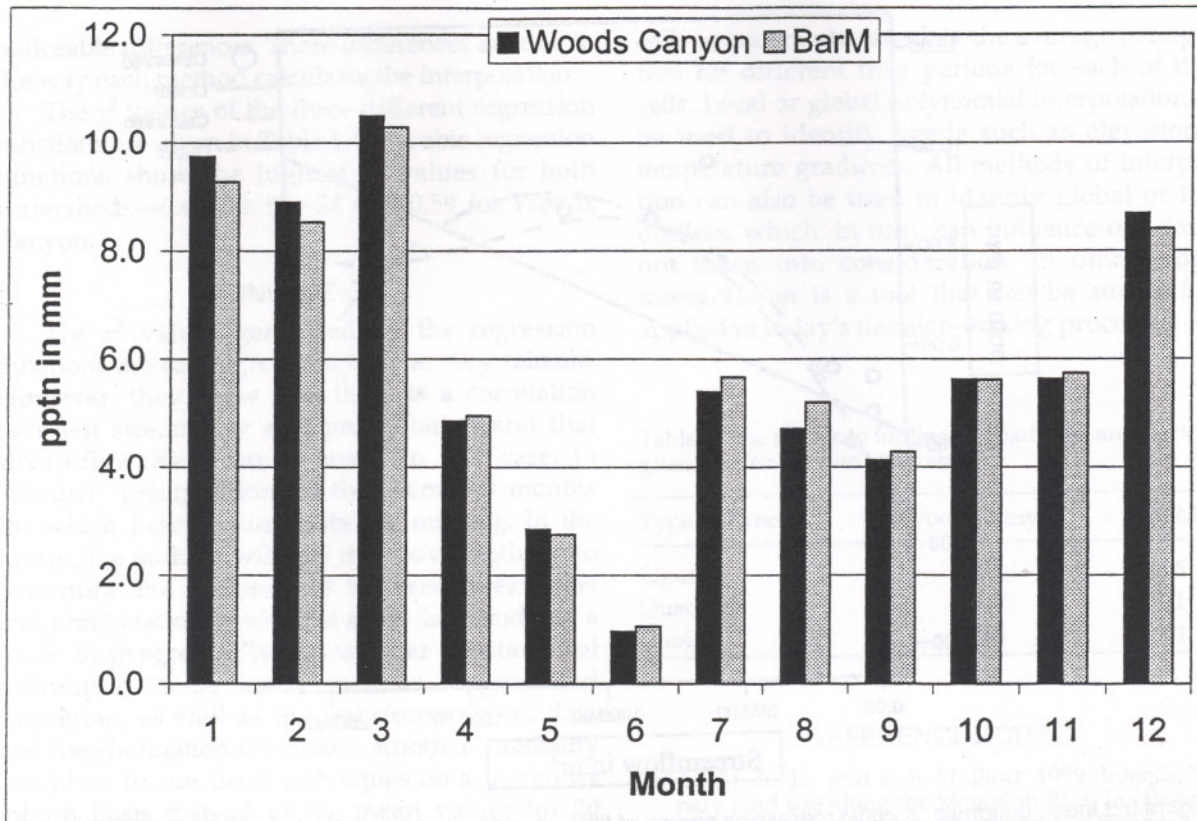


Figure 2. A bar graph showing the mean volume of precipitation for each month for the two watersheds (converted from cu m back to average mm distributed over the entire watershed).

is safe to assume that elevation and consequently temperature are factors causing a trend, given that the gauging stations used to collect this information are distributed along the elevation gradient of the Mogollon rim.

IDW, local, and global polynomial interpolations were run with a power of 1 for reasons of simplicity. The neighborhood search was set to include at least 10 neighbors for all interpolations, whereas we included 15 for the Radial Basis Function (RBF), the IDW, and the local polynomial interpolation. We selected 39 neighbors (or all gauging stations) for the global polynomial interpolation. These helped to compare the different methods to each other at the same (default) settings whenever possible.

The 24 IDW output maps (12 each for Woods Canyon and Bar M) as well as the two watershed polygons were converted into raster layers with a resolution of 50 x 50 m cells. This allowed the use of 3D spatial arithmetic to calculate precipitation volume in cubic meters for each month for each

watershed resulting in a total of 12 values for each watershed (see Figure 2). Next these values were used to predict streamflow based on precipitation volume calculated by IDW for each watershed. The results of the linear, quadratic, and cubic regression functions are shown in Figures 3 and 4.

## RESULTS

The global polynomial interpolation is the least exact interpolation, because it uses the entire dataset to make its predictions and shows values that are different from the measured values. However, it does show a trend in the level of precipitation. It indicates that precipitation increases from southeast to northwest. This trend is also indicated by the local polynomial interpolation, whose predicted values more closely resemble the measured ones. The results obtained using both the IDW and RBF, better represent the measured data than those obtained using the polynomial interpolations. Even though the results in the former two methods are relatively similar to each other, there are

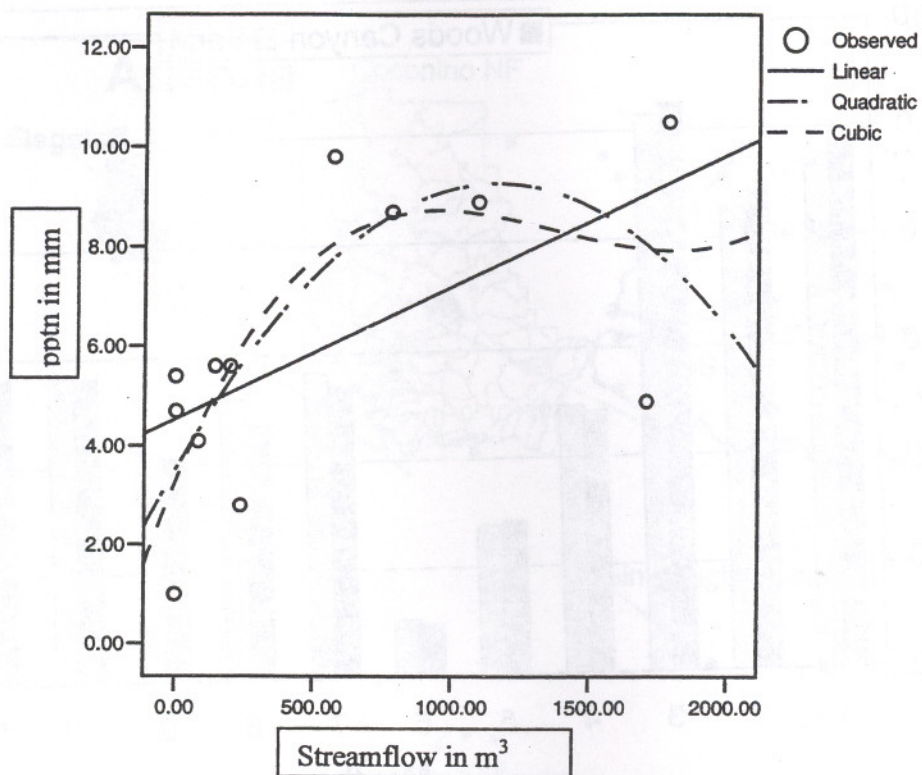


Figure 3. Linear, quadratic, & cubic regression curves of fitting the Woods Canyon data; the cubic line has the best fit.

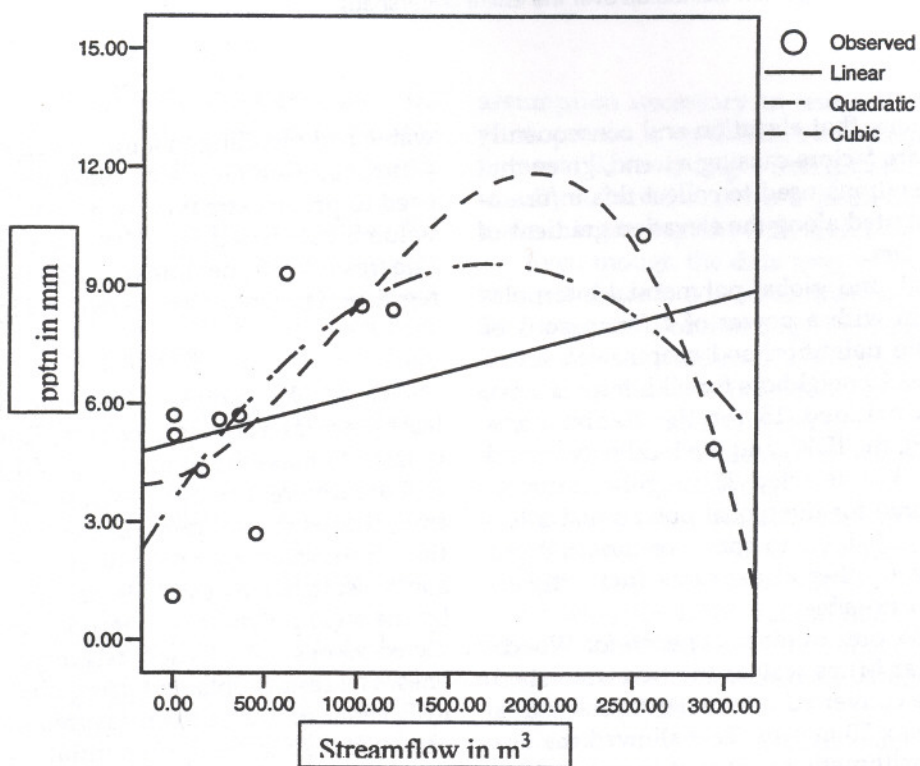


Figure 4. Linear, quadratic, and cubic regression curves of fitting the Bar M data. The cubic line has the best fit.

noticeable differences. These differences are due to the way each method calculates the interpolation.

The  $r^2$  values of the three different regression functions are given in Table 1. The cubic regression functions show the highest  $r^2$  values for both watersheds—0.60 for Bar M and 0.59 for Woods Canyon.

### CONCLUSION

The  $r^2$  values generated by the regression functions are not high enough to be very reliable. However, they show that there is a correlation between streamflow and precipitation and that streamflow data can be used, in this case, to "predict" precipitation for the years and months for which precipitation data are missing. In the future, the authors will use the above methods to determine the relationship between streamflow and precipitation; hence this study is considered a work in progress. There are other geostatistical techniques to be used, such as Kriging and Cokriging, as well as the transformation of data and the elimination of outliers. Another possibility would be to run these techniques on a month by month basis instead of the mean values for 20 years. The goal is to use such techniques to find a more reliable correlation (higher  $r^2$  value) between streamflow and precipitation in order to "predict" precipitation based on streamflow data with a higher accuracy and reliability for the water years 1982–1994.

Interpolation in general, but IDW or RBF in particular, can be very useful geostatistical tools for predicting values in areas where adequate data are not available or when it is difficult to predict values in the past. Precipitation is a prime example. Areas can be divided into relatively small

cells to accurately calculate the average precipitation for different time periods for each of these cells. Local or global polynomial interpolation can be used to identify trends such as elevation or temperature gradients. All methods of interpolation can also be used to identify global or local outliers, which, in turn, can influence our data if not taken into consideration. In other words, interpolation is a tool that can be successfully applied in today's decision-making processes.

Table 1. The  $r^2$  values for linear, quadratic, and cubic regressions for the two watersheds.

| Type of Function | Woods Canyon | Bar M |
|------------------|--------------|-------|
| Linear           | .359         | .205  |
| Quadratic        | .548         | .521  |
| Cubic            | .558         | .601  |

### REFERENCES CITED

- Baker, M. B., Jr., and P. F. Ffolliott. 1999. Interdisciplinary land use along the Mogollon Rim. In *History of Watershed Research in the Central Arizona Highlands*, compiled by M. B. Baker, Jr. USDA Forest Service General Technical Report RMRS-GTR-29, Ft. Collins, CO.
- ESRI. 2001. Using ArcGIS™ Geostatistical Analyst. ESRI, Redland, CA.
- Poff, B., D. S. Leao, A. Tecle, and D. Neary. 2005. Determining watershed boundaries and area using GPS, DEMs, and traditional methods: A comparison. Proceedings of the 7th Biennial Conference of Research on the Colorado Plateau. USGS Plateau Field Station, Northern Arizona University, Flagstaff, November 2003. University of Arizona Press, Tucson.
- Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46(2): 234–240.