STOCHASTIC PRECIPITATION SIMULATION FOR FUTURE CLIMATE CONDITIONS

David A. Woolhiser

Hydrologic Consultant

2833 Sunstone Drive, Fort Collins, CO 80525, USA

(970) 482-7810

E-mail: woolhiserd@aol.com

STOCHASTIC PRECIPITATION SIMULATION FOR FUTURE CLIMATE CONDITIONS

David A. Woolhiser

Abstract. A procedure is described to adjust Fourier series coefficients of the parameters for the second order Markov chain-mixed exponential daily precipitation model as estimated for a specific daily precipitation station for changed climate scenarios. The technique develops Fourier coefficients to describe seasonal variation of transition probability logits and mixed exponential parameters for target values of mean annual number of wet days and mean annual precipitation with seasonal variations consistent with an analog station. Adjustments of the means, amplitudes and phase angles of the four parameters of the Markov chain and the mean of the mixed exponential distribution of daily precipitation are achieved by numerical optimization techniques. The procedure is tested for present Yucca Mountain, Nevada climate, a southern Arizona monsoon-type climate, and an eastern Washington, USA climate. Daily values of maximum temperature, minimum temperature, and solar radiation with lag-one and cross correlations retained can also be simulated conditioned on the daily occurrence of precipitation. The simulated daily climate sequences can be used with distributed hydrological models to estimate hydrologic states and fluxes for small watersheds. The structure of the stochastic weather simulation model enables easy adjustment of mean annual number of wet days, precipitation and temperature as well as seasonality so that the sensitivities of hydrologic model outputs can be evaluated, and model uncertainties can be quantified.

Key Words: Precipitation, Stochastic, Simulation, Climate change

1. Introduction and Background

Hydrologists are often requested to make estimates of hydrologic fluxes such as runoff, evapotranspiration and groundwater recharge or system conditions such as soil moisture or ground water levels for watersheds where there are few or no precipitation data. With the present concern regarding climate change, they are now also asked to estimate the effects of changes of temperature and precipitation on hydrologic inputs, states and outputs. Physically-based hydrologic models are most commonly used to make these estimations and require inputs such as precipitation, temperature and radiation on at least a daily basis. In the absence of data or for hypothesized future climates, stochastic weather simulation models can provide these inputs. The models WGEN (Richardson and Wright, 1984), CLIGEN (Nicks and Gander, 1994) and USCLIMATE (Hanson et al. 1994) are examples of stochastic weather simulation models. The latter two models have been thoroughly compared by Johnson et al. (1996). Shortcomings were identified for both models and recommendations for improvements were made. However Johnson et al. (1996) concluded "…the desired application for a climate simulation model should dictate model choice. For linkage to natural systems modeling, USCLIMATE seems superior in most respects, especially because of its daily parameterization, realistic simulations of variability and preservation of correlation among elements." USCLIMATE utilizes a first order Markov chain (MC) precipitation occurrence process and a mixed exponential (ME) distribution for daily precipitation depth on wet days. This model will be referred to as the MCME model. Although a first order model is parameter efficient, a second order model is usually superior based on such measures as the Akaike information criterion (AIC) (Akaike, 1974).

Recently some investigators have developed procedures to modify parameters of stochastic weather models, specifically WGEN, to identify climate change sensitivities and impacts (Wilks, 1992; Katz, 1996; Mearns et al., 1997). WGEN uses a monthly parameterization for the precipitation model. This stepwise parameterization was identified as a weakness of CLIGEN by Johnson et al. (1996). Because USCLIMATE utilizes the same algorithms for the generation of daily temperatures and solar radiation as WGEN, yet has daily parameterization by means of Fourier series, it seems worthwhile to investigate methods to adjust parameters of USCLIMATE to create weather sequences for locations without data and for hypothesized future climates.

This paper has two objectives: (1) to describe the MC2ME model, an improvement of the MCME model as used in USCLIMATE by replacing the first order Markov chain daily occurrence process with a second order Markov chain, and (2) to present methods to adjust the Fourier coefficients that specify the seasonal variation of model parameters to achieve target mean annual precipitation (MAP) and numbers of wet days while closely matching specified seasonal variations in both the occurrence (i.e., MC2) process and the function for mean daily precipitation amount.

2. The Second Order Markov Chain-Mixed Exponential Model (MC2ME)

2.1 Markov Chain

Precipitation occurrence or nonoccurrence on day n of year τ is represented by the random variable $X_\tau(n)$; $\tau = 1,2, \ldots,M$; $n = 1,2,3,\ldots,365$; where M = number of years and:

$$
\begin{aligned}
&X_\tau(n) = 0 \quad \textit{if precipitation did not occur on day } n \\
&X_\tau(n) = 1 \quad \textit{if precipitation occurred on day } n
\end{aligned} \tag{1}
$$

The dependence between wet and dry occurrences on successive days is modeled as a daily-varying Markov chain with transition probabilities:

$$
\begin{aligned}
&P_{ij0}(n) = P\{X_\tau(n) = 0 \mid X_\tau(n-1) = j,\ X_\tau(n-2) = i\}; \quad n = 3,4,\ldots 365;\ i,j = 0,1 \\
&P_{ij0}(n) = P\{X_\tau(n) = 0 \mid X_{\tau-1}(365) = j,\ X_{\tau-1}(364) = i\}; \quad n = 1 \\
&P_{ij0}(n) = P\{X_\tau(n) = 0 \mid X_\tau(1) = j,\ X_{\tau-1}(365) = i\}; \quad n = 2
\end{aligned} \tag{2}
$$

Because $P_{ij1}(n) = 1 - P_{ij0}(n)$; $i, j = 0,1$, four parameters (P_{000}, P_{100}, P_{010}, P_{110}) are required for each day. The logit transform (Zucchini and Adamson, 1984) is used to convert probabilities bounded by 0 and 1 to variables bounded by $-\infty$ to $+\infty$.

$$
G_{ij0}(n) = \ln\left[\frac{P_{ij0}(n)}{1 - P_{ij0}(n)}\right] \tag{3}
$$

To reduce the number of parameters that must be estimated and to insure smooth seasonal variations, the daily values of the four transition probability logits are described by finite Fourier series.

$$
G_{ij0}(n) = \Gamma_{ij0} + \sum_{k=1}^{m_{ij0}} c_{ij0k} \sin(2\pi n k/365 + \Phi_{ij0k});\ i,j = 0,1; \quad n = 1,2,\ldots 365 \tag{4}
$$

where m_{ij0} is the minimum number of harmonics required to describe the seasonal variability of the transition probability, Γ_{ij0} is the annual mean, c_{ij0k} is the amplitude, and ϕ_{ij0k} is the phase angle in radians for the kth harmonic. The means, amplitudes and phase angles are estimated by numerical maximum likelihood techniques as described by Woolhiser and Pegram (1979), Roldan and Woolhiser (1982), Zucchini and Adamson (1984), and Woolhiser et al., (1993). In this study a maximum of five harmonics was considered but for each parameter m_{ij0} represents the number of harmonics with the minimum AIC. Because P_{000} has a larger sample size than the other three MC2 parameters, the minimum AIC was always determined with a larger the number of harmonics for P_{000} than the other MC2 parameters.

2.2. DISTRIBUTION OF WET DAY PRECIPITATION

A mixed exponential (ME) distribution is used to describe the precipitation depth above a threshold, *T*:

$$f_n(y') = \frac{\alpha(n)\exp[-y'/\beta(n)]}{\beta(n)} + \frac{[1-\alpha(n)]\exp[-y'/\delta(n)]}{\delta(n)} \tag{5}$$

where *y' = Y-T*; *Y* is the daily precipitation, $\alpha(n)$ is a weighting parameter having values between 0 and 1; and $\beta(n)$ and $\delta(n)$ are the means of the smaller and larger exponential distributions respectively. The threshold, *T*, is introduced as a practical matter because the smallest amount typically recorded in the U.S. is 0.25 mm (0.01 inch). A larger threshold could be used for special studies. Let $\mu(n)$ be the mean of the mixed exponential distribution on day *n*, specified by the following:

$$\mu(n) = \alpha(n)\beta(n) + [1-\alpha(n)]\delta(n) \tag{6}$$

The seasonal values of the parameters $\mu(n)$, $\beta(n)$ and $\alpha(n)$ are also described by Fourier series and the means, amplitudes and phase angles are obtained by numerical optimization of the log likelihood function as described by Woolhiser and Roldan. (1986) and Woolhiser et al. (1993). Significant harmonics are determined by the Akaike information criterion (AIC) (Akaike, 1974).

2.3 Expected Accumulated Precipitation

Todorovic and Woolhiser (1975) developed a mathematical expression for the distribution function of the total precipitation during an *m* day period, *S(m)*, for a first order Markov chain model with an exponential distribution of precipitation depths. However, the Markov chain was homogeneous and the precipitation depths were assumed to be independent and identically distributed random variables. The MC2ME model used in this study is much more complicated and an analytical expression for the distribution of *S(m)* is not possible.
The Markov chain and the distribution of amounts are assumed independent and the precipitation amounts are assumed to be serially independent, so the expected total precipitation for *m* days can be approximated as:

$$E\{S(m)\} = \sum_{n=1}^{m} E\{X(n)Y(n)\} = \sum_{n=1}^{m} E\{X(n)\}E\{Y(n)\} \tag{7}$$

where *Y(n)* is the precipitation amount on day *n* and the expected value of the number of wet days on day n is equal to the equilibrium probability of a wet day:

$$E\{X(n)\} = P\{X(n) = 1\} = P_W(n) \tag{8}$$

and the equilibrium probability of a wet day is defined as:

$$\begin{aligned} P_W(n) = {} & P[X(n-2)=0, X(n-1)=0][1-P_{000}(n)] \\ & + P[X(n-2)=1, X(n-1)=0][1-P_{100}(n)] \\ & + P[X(n-2)=0, X(n-1)=1][1-P_{010}(n)] \\ & + P[X(n-2)=1, X(n-1)=1][1-P_{110}(n)] \end{aligned} \tag{9}$$

Although there is no simple approximation for $P_w(n)$ for the non-homogeneous, second order Markov chain, it can be obtained numerically. For a second order non-homogeneous Markov chain, the state space can be defined as (c. f. Cox and Miller, 1977):

State 1: $X(n\text{-}1) = 0; X(n) = 0.$
State 2: $X(n\text{-}1) = 1; X(n) = 0.$ (10)
State 3: $X(n\text{-}1) = 0; X(n) = 1.$
State 4: $X(n\text{-}1) = 1; X(n) = 1.$

The transition probability matrix then becomes:

$$\begin{bmatrix} p_{000}(n) & 0 & 1-p_{000}(n) & 0 \\ p_{100}(n) & 0 & 1-p_{100}(n) & 0 \\ 0 & p_{010}(n) & 0 & 1-p_{010}(n) \\ 0 & p_{110}(n) & 0 & 1-p_{110}(n) \end{bmatrix} = M(n) \tag{11}$$

If $P_i(n)$ denotes the equilibrium probability of being in state i on day n, then the probability of a wet day, $P_w(n)$, is the sum of the probabilities of the states in which $X(n)=1$:

$$P_w(n) = P_3(n) + P_4(n) \tag{12}$$

where $P_3(n)$ and $P_4(n)$ are the steady state probabilities of being in states 3 and 4 as defined by equations (10).

The equilibrium probabilities can be obtained by sequentially multiplying the vector, $P_i(n\text{-}1)$ by the matrix, $\boldsymbol{M}(n)$, i.e.:

$$\begin{Bmatrix} P_1(n-1) \\ P_2(n-1) \\ P_3(n-1) \\ P_4(n-1) \end{Bmatrix} \times M(n) = \begin{Bmatrix} P_1(n) \\ P_2(n) \\ P_3(n) \\ P_4(n) \end{Bmatrix} \tag{13}$$

The equilibrium probabilities of a wet day as defined by equations (12) can be calculated sequentially for days 1 through 365 by using the first-order transition probabilities for days

363 and 364 as initial estimates for state probabilities $P_i(365)$ and then using equation (13). Although the effect of the initial values became very small after about two weeks, the calculations were carried out for 3 years and the probabilities for the third year were used for subsequent calculations.

The expected accumulated number of wet days function is:

$$E\{N(m)\} = \sum_{1}^{m} P_w(n) \tag{14}$$

The mean annual number of wet days (MAN) occurs at m =365. Leap years are not explicitly accounted for in this model although leap days can be included in the simulation mode.

The expected accumulated precipitation on day m can be written as:

$$E\{S(m)\} = \sum_{n=1}^{m} [P_W(n)][\mu(n) + T] \tag{15}$$

where $P_w(n)$ is the equilibrium probability of a wet day on day n and $[\mu(n) + T]$ is the expected value of Y. Expected annual precipitation is obtained when m = 365.

3. Adjustment of MC2ME Parameters for Climates

As climate changes, it is anticipated that there will be changes in mean annual precipitation, mean annual number of wet days, and the seasonal distribution of precipitation. In the MC2ME model the seasonal distribution of precipitation is controlled by the amplitudes and phase angles of the significant harmonics for each parameter. The term "seasonality" is used in this context.

Four cases can be identified:

A. Data exist at multiple locations near the site of interest and the range of MAN and mean annual precipitation (MAP) at these sites covers existing and target MAN and MAP values. Seasonality will not change.

B. Data exist for a single station at or near the site of interest but the MAN and MAP are different from estimated existing values and target values. Seasonality will not change.

C. It is hypothesized that the seasonality as well as MAN and MAP will change and data exist at multiple locations with desired seasonal patterns and the range of MAN and MAP at these sites covers existing and target values.

D. It is hypothesized that the seasonality as well as MAN and MAP will change and data are available at a site that has the desired seasonality.

For cases A and C, a regression approach similar to that utilized by Hanson et al. (1989) could be used. These investigators identified MCME parameters using data from a network of raingages on the Reynolds Creek Experimental Watershed in southwestern Idaho, USA to evaluate the effect of mean annual precipitation (MAP) on Fourier coefficients (means, amplitudes and phase angles) of the parameters of a first order Markov chain-mixed exponential (MCME) model. The significant regression relationships were then utilized to estimate MCME coefficients for a changed MAP. Regional averages were used for coefficients with no significant relation to MAP. This technique does not include MAN explicitly, but usually MAN is highly correlated with MAP.

A different approach is required for cases B and D or where there are an insufficient number of stations to obtain reliable regression relations. The term “analog station” will refer to the station whose MC2ME parameters are to be adjusted to provide values for a “target station”. The parameters for the analog station must be estimated from measured daily precipitation data. If it is assumed that the relative seasonal variations of the occurrence process and the mean wet-day precipitation will remain the same as the analog station with increased (or decreased) mean annual precipitation, it is hypothesized that the following criteria should be met.

1.) The mean annual number of wet days (MAN) should match a specified target.
2.) The normalized expected number of wet days as a function of time for the target station should closely follow that for the analog station.
3.) The mean annual precipitation (MAP) should match a specified target.
4.) The normalized expected accumulated precipitation as a function of time for the target station should closely follow that for the analog station.

Techniques designed to achieve the above criteria are described in the following section.

3.1 Adjusting Fourier Coefficients of the MC2ME Model to Obtain Target Values of MAN, MAP, and Desired Seasonality

3.1.1 Adjusting MC2 ME Parameters to Adjust Mean Annual Number of Wet Days (MAN) and Mean Annual Precipitation (MAP) to match specified targets.

The target MAN (MANT) will be achieved by adjusting the mean logits Γ_{000}, Γ_{100}, Γ_{010} and Γ_{110} to reach a zero crossing of the function:

$$F(\Gamma_{000}, \Gamma_{100}, \Gamma_{010}, \Gamma_{110}) = E\{N(365)\} - MANT \tag{16}$$

Given trial values, a Newton-Raphson iterative scheme is utilized until the absolute value of the error function is smaller than an ϵ equal to (0.005)MANT. The corrections are apportioned by the relations:

$$
\begin{aligned}
w_1 F + \frac{\partial F}{\partial \Gamma_{000}} \Delta \Gamma_{000} &= 0 \\
w_2 F + \frac{\partial F}{\partial \Gamma_{100}} \Delta \Gamma_{100} &= 0 \\
w_3 F + \frac{\partial F}{\partial \Gamma_{010}} \Delta \Gamma_{010} &= 0 \\
w_4 F + \frac{\partial F}{\partial \Gamma_{110}} \Delta \Gamma_{110} &= 0
\end{aligned}
\tag{17}
$$

where $\Delta \Gamma_{ij0}$ are correction terms and w_i are weighting factors such that $\sum_1^4 w_i = 1$.

The approximate derivatives were obtained by finite differences. The choice of the weighting factors, w_i, is not well established. In this study the ratio of the log likelihoods of each transition probability function as identified for the analog station to the sum of the log likelihoods for all four transition probability functions was used with satisfactory results. The target MAP (MAPT) is achieved by adjusting the Fourier series constant of the mean precipitation on a wet day, $\bar{\mu}$ to find the zero crossing of the function:

$$F(\bar{\mu}) = E\{S(365)\} - MAPT \tag{18}$$

We wish to adjust $\bar{\mu}$ such that *F(.)* = 0. Given *E{S(365)}* as described above with a trial value of $\bar{\mu}$, we get $F(\bar{\mu})$ and then use a Newton-Raphson iterative scheme until the absolute value of the error function is smaller than some ϵ.

3.1.2 Adjusting MC2 ME Parameters to Fit the Normalized Functions of Expected Accumulated Number of Wet Days, $E^*\{N(n)\}$, and Expected Accumulated Precipitation $E^*\{S(n)\}$.

The normalized expected number of wet days function can be written as:

$$E^*\{N(m)\} = \frac{1}{E\{N(365)\}} \sum_{n=1}^{m} P_w(n); \quad m = 1, 2, \ldots, 365 \tag{19}$$

where $N(m)$ is the accumulated number of wet days on day m and $P_w(n)$ is the equilibrium probability of a wet day.

The normalized expected accumulated precipitation function, $E^*\{S(m)\}$, is obtained by dividing equation (8) by the expected annual precipitation, i.e.:

$$E^*\{S(m)\} = E\{S(m)\}/E\{S(365)\} \tag{20}$$

The parameters of the functions $E^*\{N(n)\}$ and $E^*\{S(n)\}$ for the target station were adjusted by a simplex algorithm to minimize the sum of the squared errors between the corresponding functions for the analog station. A sequential procedure was followed for the logits of each transition probability, $G_{ij0}(n)$:

1) Optimize phase angles for the target station.
2) Optimize the amplitudes for the target station.

This procedure was embedded in a loop which included matching the target MAN and was exited when both the MAN criterion and the minimum squared error criterion were met. At this point the parameters of the second order Markov chain for the target station meet the criteria so the parameters of the mixed exponential distribution can be optimized to match MAP and to minimize the sum of squared deviations between $E^*\{S(n)\}$ for the target station and the corresponding function for the analog station.

A sequential procedure was followed for the mean of the mixed exponential distribution, $\mu(n)$:

1) Optimize phase angles for the target station.
2) Optimize the amplitudes for the target station.

A simplex algorithm was utilized for both steps. This procedure was also embedded in a loop which included matching the target MAP and was exited when both the MAP criterion and the minimum squared error criterion were met as shown in the flowchart in Figure 1. Although there are nonlinearities in the Fourier series expressions for logits of the transition probabilities and the mixed exponential distribution, the simplex algorithm achieved very good fits to the normalized functions $E^*\{N(n)\}$ and $E^*\{S(m)\}$. However, there is no guarantee of a unique solution.

4. Example for Yucca Mountain, Nevada, USA

Yucca Mountain, Nevada, USA, is located in a semi-arid climate that is currently in an interglacial dry period (Bechtel SAIC LLC, 2004; Sharpe, 2003; Stothoff and Walter, 2007). Past climates have been and future climates are expected to vary from the dryer interglacial climate to wetter glacial transition and monsoonal climates. Integrating climate, surface water redistribution and infiltration, catchment models have been used to help understand the timing and distribution of infiltration expected from event-based storms reflecting the current climate (Woolhiser, et al., 2006). Catchment models reflect water that has percolated below the root zone on upland areas and water that has infiltrated into the beds of ephemeral channels. To predict future infiltration rates at Yucca Mountain, the catchment model will require precipitation inputs reflecting future glacial and monsoonal climatic conditions.

Two approaches can be used to estimate infiltration fluxes under future conditions: (1) find an analog site that has a climate similar to the expected future climate and also has similar geologic characteristics and make physical measurements of fluxes and (2) utilize a hydrologic model that describes the appropriate runoff and infiltration mechanisms and model the Yucca Mountain conditions under a hypothesized future climate. Finding sites with appropriate climates is much more readily accomplished, thus approach (2) is followed here. Because of the strong relationship between precipitation and elevation in southern Nevada, it is possible to find an analog site nearby that has the hypothesized future annual precipitation. However, the seasonal distribution of precipitation may change in the future. The modeling approach requires the development of a stochastic weather simulator to provide such variables as precipitation, maximum temperature, minimum temperature and solar radiation on at least a daily time step.

Stochastic daily precipitation models can be used for simulation of infiltration under current climate conditions and for hypothsized future climate conditions. In each case it is appropriate to estimate hydrologic model sensitivity to errors in target MAN and MAP. With a climate change, there may be a change in the source of air masses. It has been hypothesized that the seasonal pattern of precipitation under a warmer climate with an increase of precipitation might be similar to the monsoonal pattern of southern Arizona. Similarly, if the climate cools, but precipitation increases, the seasonal pattern may be similar to that in eastern Washington state. The changed seasonal patterns would be reflected in both the Markov chain occurrence model and in the mixed exponential distribution of daily precipitation. These changes would be reflected in changes in the phase angles and amplitudes of the Fourier series representations of the model parameters.

4.1 Regression Approach for the Yucca Mountain, Nevada, Region

The regression approach may be an appropriate method for specifying a model for current conditions at Yucca Mountain. Precipitation measurements in the area immediately around Yucca Mountain have rather short records, while there are longer records at nearby stations with wide ranges of MAN and MAP. It is estimated that the current MAP at a reference elevation of 1,400 m on Yucca Mountain is 185 mm (Bechtel SAIC LLC, 2004; Stothoff and Walter, 2007). Increases in precipitation for expected future climates could be a factor of up to approximately 2.6 (Stothoff and Walter, 2007); thus, suggesting an upper bound of approximately 480 mm precipitation.

If one accepts the hypothesis that a future wetter and cooler climate at Yucca Mountain would be similar to the climates presently existing at nearby locations at higher elevations, the relationships developed by Hanson et al. (1989) could be useful. Hanson et al. (1989) utilized data from a network of raingages on the Reynolds Creek Experimental Watershed in southwestern Idaho to evaluate the effect of MAP on parameters of an MCME model. They used procedures described by Woolhiser and Roldan (1986) to estimate the MCME model parameters for each of the 16 gages on the watershed. The annual precipitation at these gages ranged from 241 to 1,144 mm at the Reynolds Creek Watershed, and the MAN of wet days ranged from 84 to 133.9 mm. Because the projected annual precipitation range of 185 mm to 480 mm at Yucca Mountain is on the order of, or smaller, than the range of the Idaho stations, the regression approach may also be reasonable for Yucca Mountain.

It would be unrealistic to assume that the regression equations developed from Idaho data would provide good estimates of the parameters of the MC2ME model for Yucca Mountain under current or future climates. However, the same approach could be used with precipitation data from the Yucca Mountain region. Daily precipitation data for ten stations from the nearby Nevada National Security Site (NNSS) network and the National Oceanic and Atmospheric Administration (NOAA)/National Weather Service (NWS) Las Vegas station were used to develop a regional approach. A list of the stations and relevant information is included in Table 1.

The data were edited to conform to the input requirements for the MC2ME program, i.e. data begin on March 1 and leap days are eliminated. Years with large blocks of missing data were eliminated. Fourier coefficients of the parameters of the MC2ME model were identified for each station. There was a high degree of correlation ($R^2 = 0.94$) between the mean annual number of wet days (MAN) and the mean annual precipitation (MAP), so only MAP was used as an independent variable in a regression analysis. Regression coefficients for the Markov chain are shown in Table 2 and for the mixed exponential distribution in Table 3. As expected, the regression parameters in Tables 2 and 3 differ from those found by Hanson et al. (1989), although the relationships are consistent. The Nevada regressions have smaller R^2 values and there is no significant relation between the ME weighting parameter α and MAP. However there are significant relations between the mean of β and the amplitude of the second harmonic of μ and MAP. MAP has little effect on the seasonality of the Markov chain, but has a weak effect on the seasonality of the ME distribution.

The regression relationships shown in Tables 2 and 3 were used to estimate MC2ME parameters for a modern climate at Yucca Mountain with an estimated MAP of 181 mm. and for a future climate with MAP = 331 mm. Target MAN was determined by regression with MAP. Where there were no significant relations between the parameters and MAP, Fourier coefficients for an analog station were used and were adjusted as described in following sections.

4.2 Normalized Function Approach for the Yucca Mountain, Nevada, Region

For a test of the procedure described in section 3, MC2ME parameter estimates for two intermediate target stations were made from two analog stations, one with a higher MAP and one with a lower MAP. Analog stations are:

1) Rainier Mesa. MAN =53.3 days; MAP = 314.2 mm.
2) Jackass Flats. MAN = 32.1 days; MAP = 140.8 mm.

The same adjustment weights, obtained from the log likelihood functions when the MC2ME parameters were identified for the analog stations were used for both stations. The weights were: $w_1 = 0.622$, $w_2 = 0.090$, $w_3 = 0.146$, $w_4 = 0.142$.

The target stations are:

1) Fortymile Canyon. MAN = 43.05 days; MAP = 203.8 mm.
2) Tippipah Springs. MAN = 43.9 days; MAP = 224.3 mm.

Markov chain transition probabilities and the mean of the mixed exponential distribution for Fortymile Canyon as adjusted from the analog stations are compared with sample values for 13 time intervals from the historical record in Figures 2 a - c. Similar results for Tippipah Springs are shown in Figure 3a - c. The fit for the most sensitive parameter, P_{000}, is very good, with the adjustment from the wetter station, Rainier Mesa, slightly better than that for Jackass Flats. There is more variability for the other three parameters, which is not surprising because of the smaller sample sizes. Again the fits obtained by adjustment from Rainier Mesa are better.

4.3 Normalized Function Approach for a Monsoon Climate

This procedure was tested for a monsoon climate by identifying the MC2ME parameters for two stations with similar seasonality but different MAP and MAN (Nogales, AZ and RG4, AZ) and to adjust the parameters for each one to attain the target MAP and MAN of the other station. The relevant target values are:

RG4: MAP = 302.26 mm, MAN = 53.57 days.
Nogales: MAP = 411.72 mm, MAN = 58.84 days.

Adjustment weights: $w_1 = 0.599$, $w_2 = 0.114$, $w_3 = 0.165$, $w_4 = 0.122$.

The procedures described above were utilized to identify MC2ME parameters for the target values of MAN and MAP while maintaining the dimensionless functions, $E^*\{N(n)\}$ and $E^*\{S(n)\}$ for the analog station. The sample transition probabilities and parameter, $\mu(n)$, identified from data for Nogales and the Fourier series fits of the Nogales data are plotted with those obtained by adjusting from the analog station, RG4, to achieve the Nogales MAN, MAP and seasonality in Figure 4 a-c. The corresponding accumulated number of wet days and precipitation functions are shown in Figures 5a and b. An examination of the transition probabilities reveals that both stations have the same number of significant harmonics and that the adjusted values are rather close to those estimated from the station data. Deviations should be expected because the stations are, in fact, different and have different periods of record. Sampling variability, model error and shortcomings in the adjustment procedures play a role and will be reflected in the adjusted $E^*\{N(n)\}$ and $E^*\{S(n)\}$ functions and impose a limit on the accuracy that can be achieved.

4.4 Normalized Function Approach for Northwestern USA

MC2ME parameter estimates for one intermediate target station were made from two analog stations, one with a higher MAP and one with a lower MAP. Analog stations are:

1) Pullman, WA. MAN =113.7 days; MAP = 537.9 mm.
Adjustment weights: $w_1 = 0.454$, $w_2 = 0.158$, $w_3 = 0.180$, $w_4 = 0.207$

2) Yakima, WA. MAN = 69.4 days; MAP = 208. mm.
Adjustment weights: $w_1 = 0.580$, $w_2 = 0.127$, $w_3 = 0.166$, $w_4 = 0.127$

The target station is:
Spokane, WA. MAN = 111.5 days; MAP = 408.8 mm.

Markov chain transition probabilities and the mean of the mixed exponential distribution for Spokane as adjusted from the analog stations are compared with sample values for 26 time intervals from the historical record in Figures 6 a - d. The fit for the most sensitive parameter, P_{000}, is very good, with the adjustment from the wetter station, Pullman, slightly better than that for Yakima. There is more variability for the other three parameters, which is not surprising because of the smaller sample sizes. Again the fits obtained by adjustment from Pullman are better. For the mean of the mixed exponential, the adjusted values from Yakima show less seasonal variability that the data and the estimates from Pullman.

5. Discussion and Conclusions

5.1 Changes in MC2ME Parameters

The adjustments of the Fourier constants, amplitudes and phase angles were examined for the Yucca Mountain examples. When the analog station had smaller MAP and MAN than the target station, the Fourier constants for the transition probability logits decreased while the amplitudes of the harmonics increased slightly. The Fourier constant for $\mu(n)$ increased as did the amplitudes. There was very little change in the phase angles for either the counting process or the mean of the mixed exponential. The parameters $\alpha(n)$ and $\beta(n)$ were not changed by the current procedure. When the target station had lower MAP and MAN, the changes were of opposite signs, as expected.

The regression relationships shown in Tables 2 and 3 were used to estimate MC2ME parameters for a modern climate at Yucca Mountain with an estimated MAP of 181 mm and for a future climate with MAP of 331 mm. MANT was determined by regression with MAP. Where there were no significant relationships between the parameters and MAP, Fourier coefficients for an analog station were used for starting values and were adjusted using the normalized adjustment procedure.

There were virtually no differences between the regression appproach and the normalized adjustment procedure for the calculated functions $E\{N(n)\}$ and $E\{S(n)\}$. There were minor differences between Fourier coefficients for both MC2 and ME, but they appear insignificant. Consequently, the adjustment procedure is superior to the regression approach in this case because it does not have the requirement of a sufficient number of records to obtain reliable regression.

Although these were not exhaustive tests, the techniques are superior to some procedures used previously such as scaling the daily precipitation generated from a stochastic model for a station with a lower or higher precipitation by the ratio of the MAPs, thus ignoring the change in number of wet days.

5.2 Conclusions

The procedure for adjusting MC2ME parameters from an analog station to achieve target station

MAN and MAP along with closely matching the seasonal variation of both the occurrence process and the distribution of wet day precipitation for an analog station performs satisfactorily. Tests of the procedure were performed for three climatic regions in the western U.S. by fitting the MAN and MAP for an existing target station by adjusting parameters for an existing analog station and comparing the adjusted transition probabilities and wet day mean function with sample values from target station data. Accuracy of the results are contingent upon close fits between the normalized expected accumulated number of wet day and normalized accumulated precipitation functions for analog and target stations. Different periods of record as well as sampling variability and model error also contribute to differences. The procedure appears superior to regression techniques because it doesn't require a sufficient number of nearby stations to achieve reliable regression relationships between MC2ME model parameters and MAN.

Daily radiation and maximum and minimum temperatures can subsequently be generated conditioned on the precipitation occurrence while maintaining cross and lag-one serial correlation using procedures described by Richardson (1981, 1982) and included in the program USCLIMATE (Hanson et al. 1994). The second order Markov chain model described here should be superior to a first order process to provide input to simulation models to estimate hydrologic fluxes and states under conditions of climate change. Because of the ease of creating new daily sequences with different MAN, MAP and mean temperature values, and the fact that both temperatures and radiation are generally conditioned on precipitation occurrence, the sensitivity of runoff and infiltration to these factors can be readily evaluated.

6. Acknowledgements

This paper was prepared to document work performed by the Center for Nuclear Waste Regulatory Analyses (CNWRA) for the U.S. Nuclear Regulatory Commission (NRC) under Contract No. NRC–02–07–006. The activities reported here were performed on behalf of the NRC Office of Nuclear Material Safety and Safeguards, Division of High-Level Waste Repository Safety. This paper is an independent product of CNWRA and does not necessarily reflect the view or regulatory position of NRC. Precipitation data were obtained from the Western Regional Climate Center and the Southwestern Rangeland Watershed Research Center, ARS-USDA. Data near Yucca Mountain were obtained from the U.S. Geological Survey (USGS) and Science Applications International Corporation (SAIC) and are available from the YM Project, U.S. Department of Energy.

7. References

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control*, 6, 716-723.

Bechtel SAIC Company, LLC. 2004. "Future Climate Analysis." Rev. 01. Las Vegas, Nevada: Bechtel SAIC Company, LLC.

Cox, D. R., and Miller, D. R. 1977. *The Theory of Stochastic Processes*. Chapman and Hall, London, 398 pp.

Hanson, C. L., K. A. Cumming, D. A. Woolhiser and C. W. Richardson. 1994. Microcomputer Program for Daily Weather Simulation in the Contiguous United States. U. S. Dept. of Agriculture, Agricultural Research Service, ARS-114, 38 p.

Hanson, C.L., Osborn, H.B. and Woolhiser, D.A. 1989. Daily Precipitation Model for Mountainous Areas. *Trans. ASAE* 32:3. 865-873.

Johnson, G. L., C. L. Hanson, S. P. Hardegree and E. B. Ballard. 1996. Stochastic weather simulation: overview and analysis of two commonly used models. *J. Am. Meteor. Soc*. 35(10):1878-1896.

Katz, R.W. 1996. Use of stochastic models to generate climate change scenarios. *Climatic Change* 32, 237-255.

Mearns, L.O., C. Rosenzweig and R. Goldberg, 1997. Mean and variance change in climate scenarios: Methods, agricultural applications, and measures of uncertainty. *Climatic Change* 35, 367-396.

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

Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature and solar radiation. *Water Resour. Res*., 17, 182-190.

Richardson, C.W., 1982. Dependence structure of daily temperature and solar radiation. *Trans. ASAE*, 25, 735-739.

Richardson, C.W. and D.A. Wright, 1984. WGEN: A model for generating daily weather variables. US Dept. Agr., Agricultural Research Service, Pub. No. ARS-8, 83pp.

Roldan, J., and Woolhiser, D. A. 1982. Stochastic Daily Precipitation Models. 1. A Comparison of Occurrence Processes. *Water Resour. Res*. 18(5):1451-1459.

Sharpe, S. 2003. “Future Climate Analysis—10,000 Years to 1,000,000 Years After Present.” MOD–01–001. Rev. 00. Reno, Nevada: Desert Research Institute.

Stothoff, S., and G. Walter. 2007. Long-Term Average Infiltration at Yucca Mountain, Nevada: Million-Year Estimates. CNWRA-2007-003. Center for Nuclear Waste Regulatory Analyses, San Antonio, TX.

Todorovic, P. And Woolhiser, D.A. 1975. A Stochastic Model of n-Day Precipitation. *J. Applied Meteor*. 14(1); 17-24.

Wilks, D.S., 1992. Adapting stochastic weather generation algorithms for climate change studies. *Climate Change*, 22, 67-84.

Woolhiser, D. A., and J. Roldan, 1986. Seasonal and Regional Variability of Parameters in Stochastic Daily Precipitation Models - South Dakota, U.S.A. *Water Resour. Res*. 22(6):965-978.

Woolhiser, D. A., and G. G. S. Pegram, 1979. Maximum Likelihood Estimation of Fourier Coefficients to Describe Seasonal Variations of Parameters in Stochastic Daily Precipitation Models. *J. Applied Meteor*. 18 (1):34-42.

Woolhiser, D.A., T.O. Keefer and K.T. Redmond. 1993. Effects of the Southern Oscillation Index on Daily Rainfall in the Southwestern U.S. *Water Resour. Res*. 29(4), 1287-1295.

Woolhiser, D. A., Fedors, R. W., Smith, R. E. And Stothoff, S. A. 2006. ‘Estimating Infiltration in the Upper Split Wash Watershed, Yucca Mountain, Nevada’, *J. Hydrologic Engr.* 11(2) 123-133.

Zucchini, W. and Adamson, P. T. 1984. The Occurrence and Severity of Drought in South Africa., Dept. Of Civil Engr., Univ. of Stellenbosch and Dept. Of Water affairs, WRC Report No. 91/1/84, 198 pp.

8. Figures

6c. Sample Markov chain parameter $P_{11\,0}$ for Spokane, WA and as Estimated from Pullman and Yakima.

6d. Sample Mixed Exponential parameter μ for Spokane, WA and as Estimated from Pullman and Yakima.

9. Tables

1. Precipitation station data used in regression study for Yucca Mountain, Nevada.

2. Regression relationships for second order Markov chain (MC2), Yucca Mountain, NV Region.

3. Regression relationships for Mixed Exponential Distribution (ME), Yucca Mountain, NV Region.

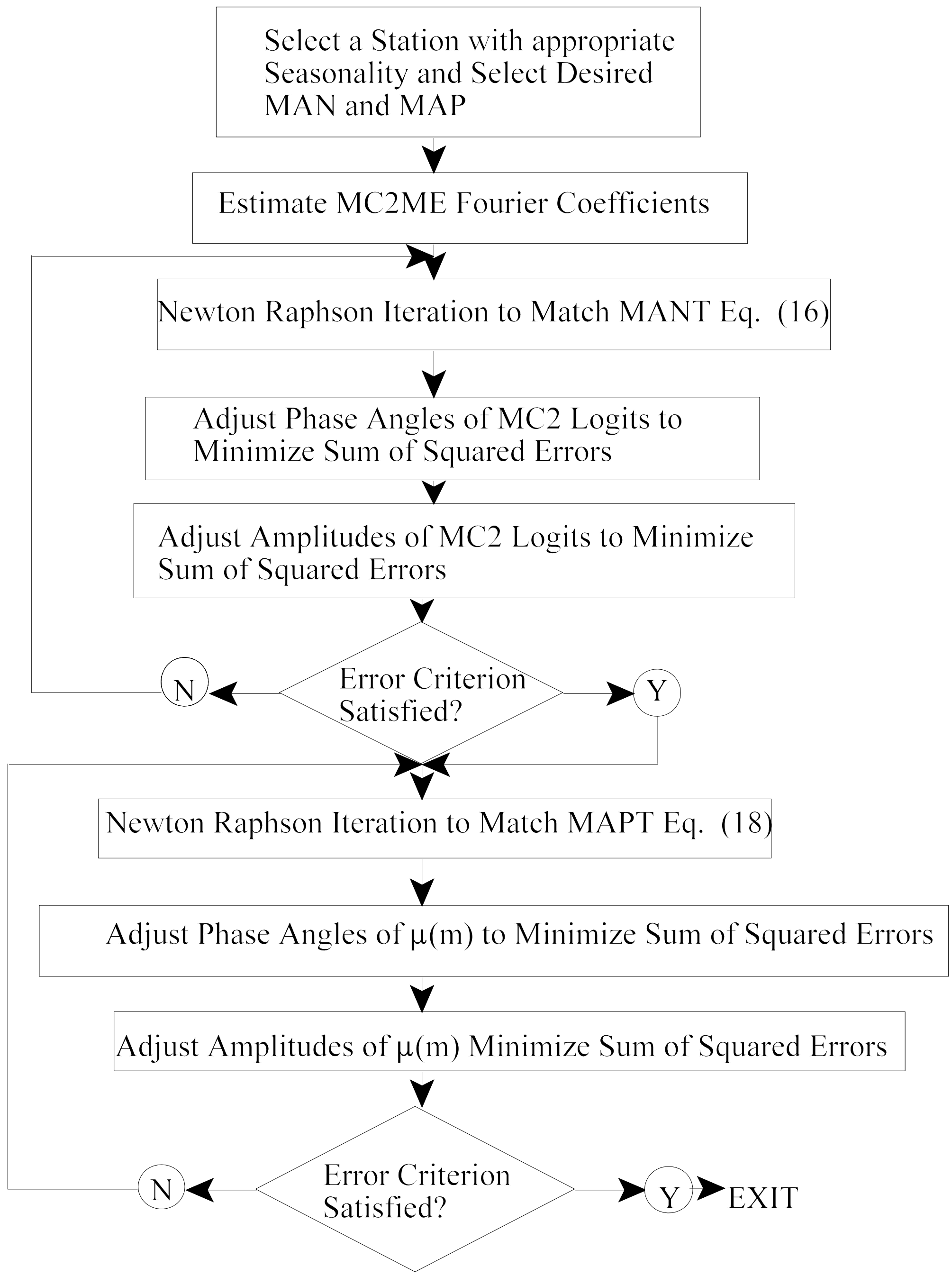

Figure 1. Flowchart of Algorithm to Match MAN and MAP with Minimum Sum of Squared Errors for Normalized E{N(m)} and E{S(M)} Functions

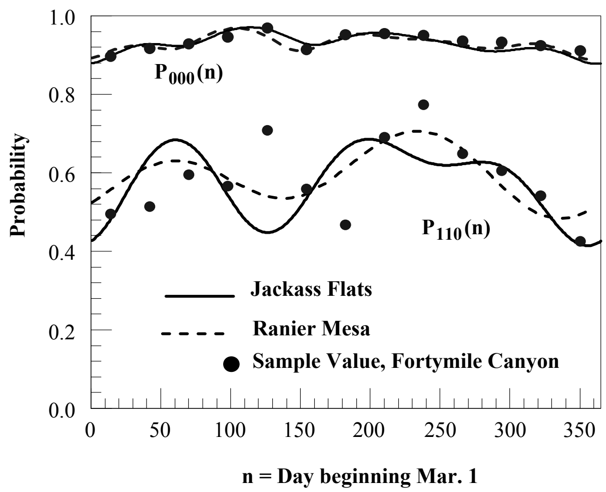

Figure 2a. Markov chain parameters P_{000} and P_{110} for Fortymile Canyon, NV and as Estimated from Jackass Flats and Rainier Mesa.

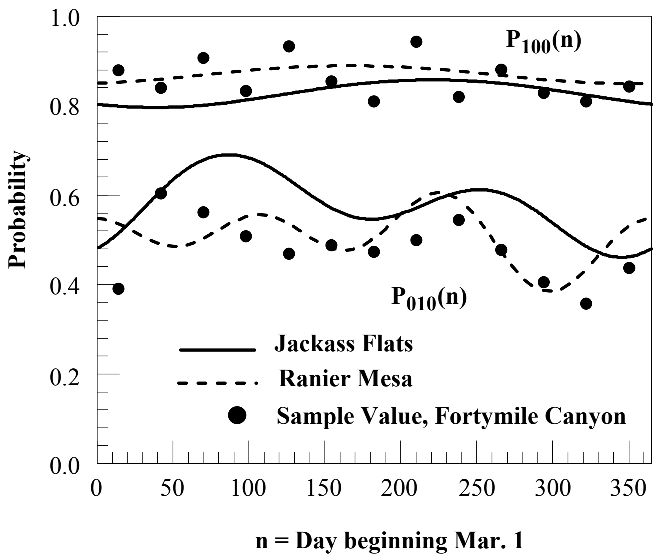

Figure 2b. Markov chain parameters P_{100} and P_{010} for Fortymile Canyon, NV and as Estimated from Jackass Flats and Rainier Mesa.

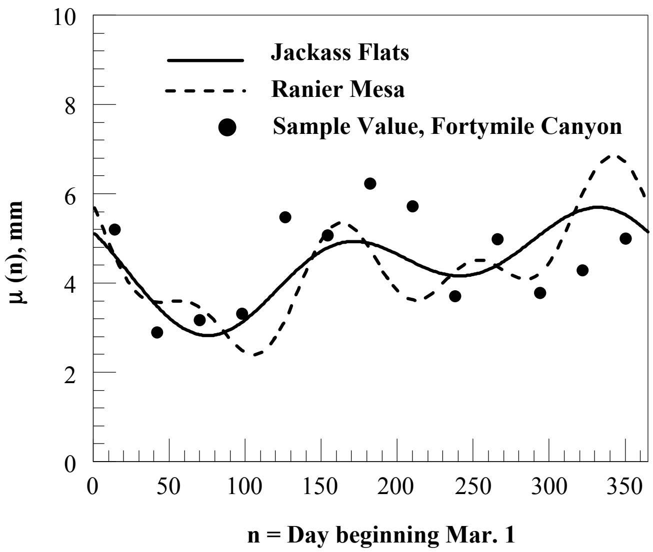

Figure 2c. Mixed Exponential parameter μ for Fortymile Canyon, NV and as Estimated from Jackass Flats and Rainier Mesa.

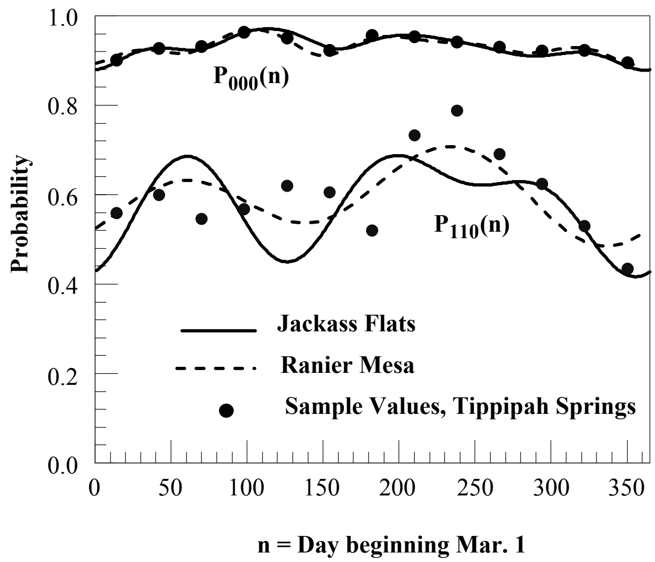

Figure 3a. Markov chain parameters P_{000} and P_{110} for Tippipah Springs, NV and as Estimated from Jackass Flats and Rainier Mesa.

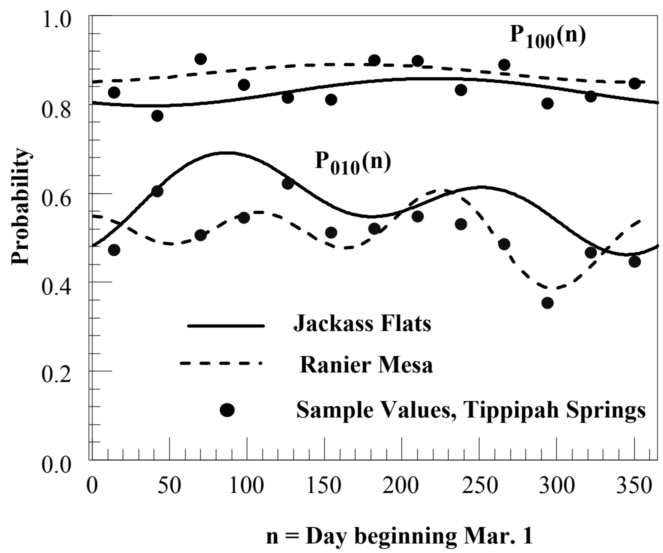

Figure 3b. Markov chain parameters P_{100} and P_{010} for Tippipah Springs, NV and as Estimated from Jackass Flats and Rainier Mesa.

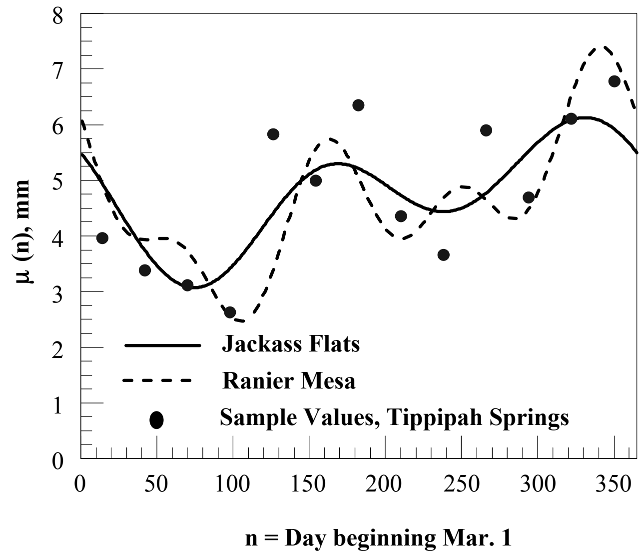

Figure 3c. Mixed Exponential parameter μ for Tippipah, NV and as Estimated from Jackass Flats and Rainier Mesa.

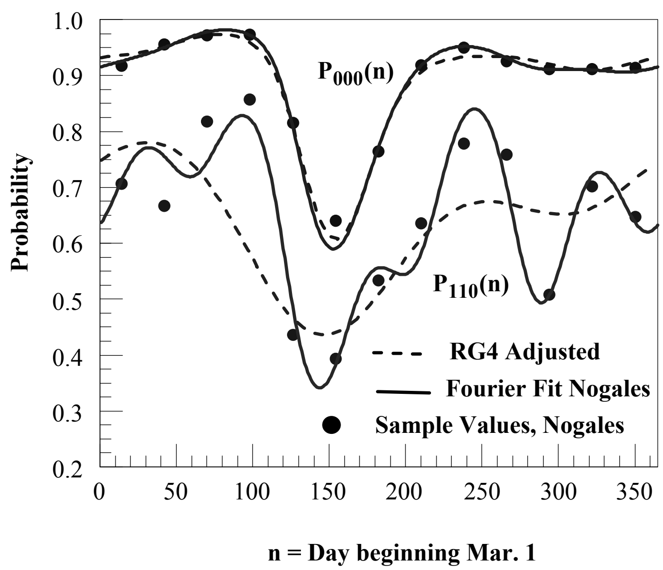

Figure 4a. Markov chain parameters P_{000} and P_{110} for Nogales, AZ and as Estimated from Walnut Gulch RG4.

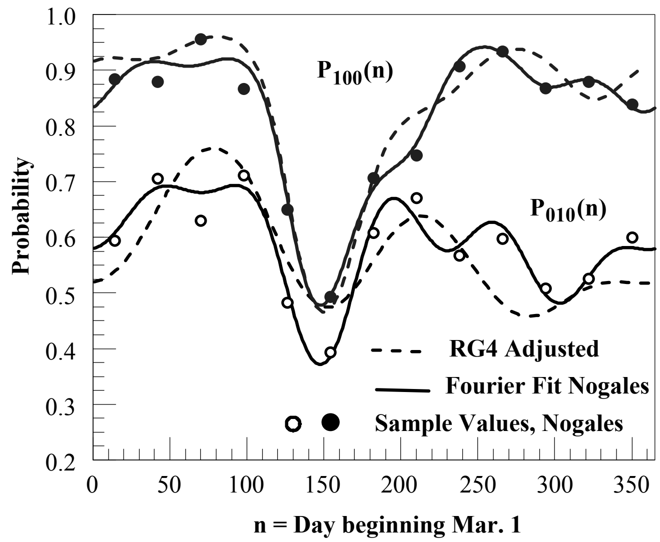

Figure 4b. Markov chain parameters P_{100} and P_{010} for Nogales, AZ and as Estimated from Walnut Gulch RG4.

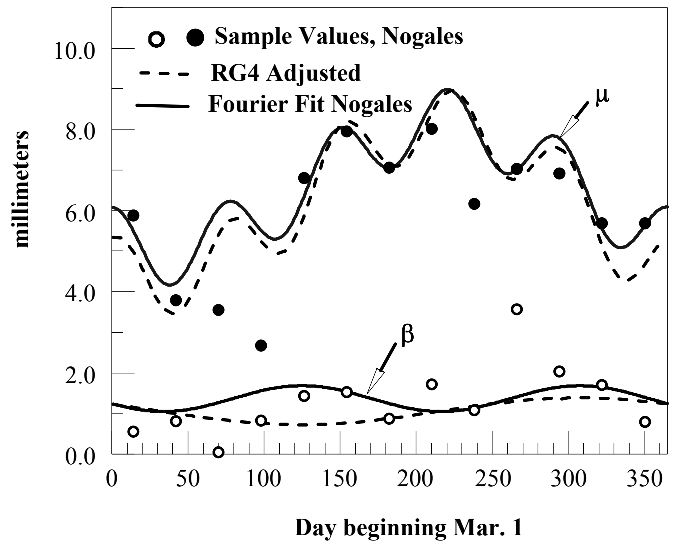

Figure 4c. Mixed Exponential parameter μ for Nogales, AZ and as Estimated from Walnut Gulch RG4.

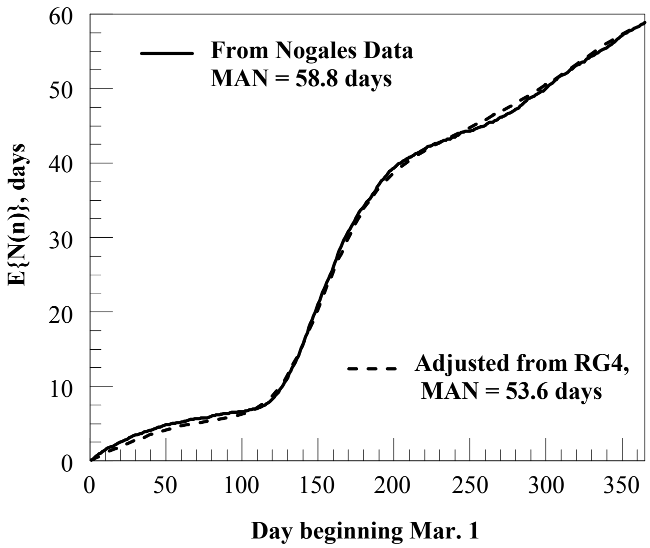

Figure 5a. Sample Accumulated Number of Wet Days Function for Nogales, AZ and as Estimated from Walnut Gulch RG4.

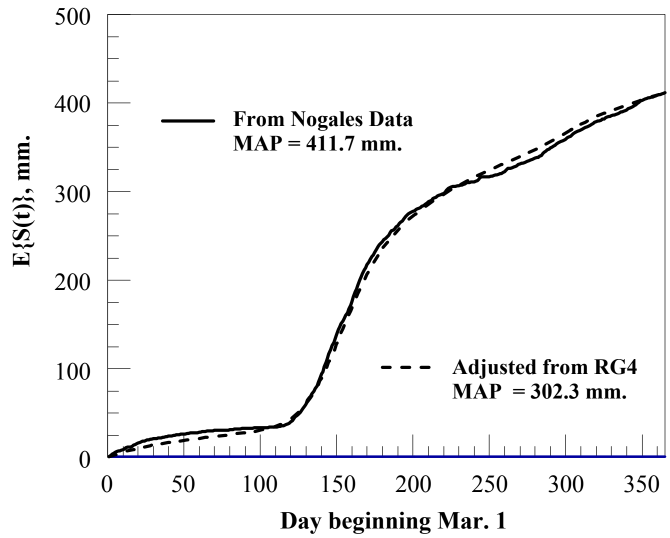

Figure 5b. Sample Accumulated Precipitation Function for Nogales, AZ and as Estimated from Walnut Gulch RG4.

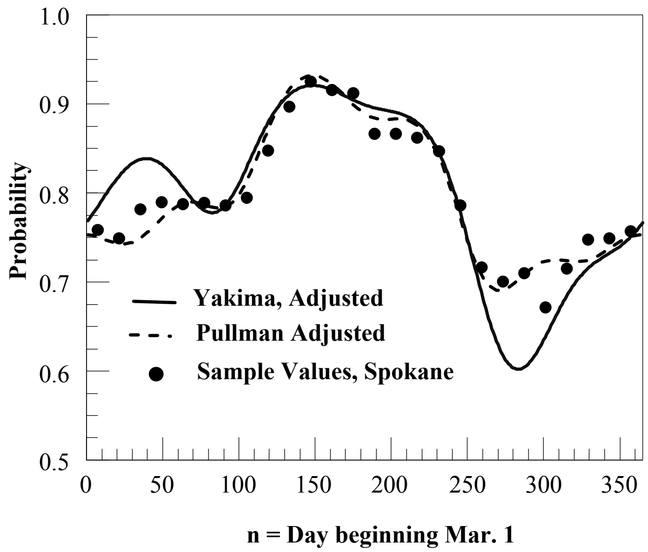

Figure 6a. Sample Markov chain parameter P_{000} for Spokane, WA and as Estimated from Pullman and Yakima.

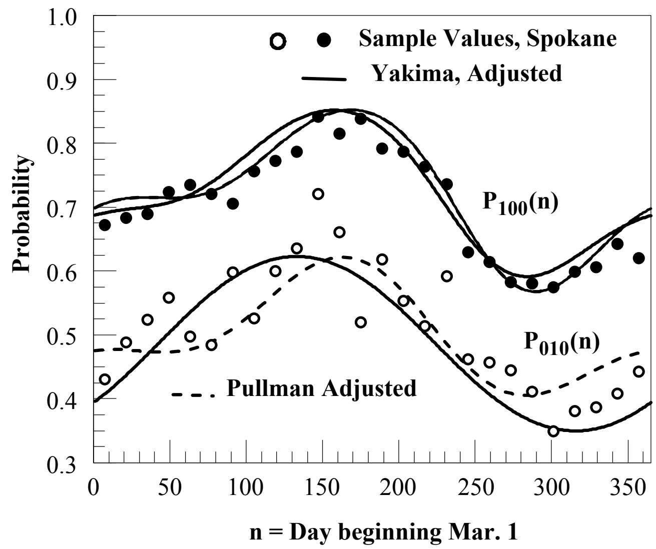

Figure 6b. Sample Markov chain parameters P_{100} and P_{010} for Spokane, WA and as Estimated from Pullman and Yakima.

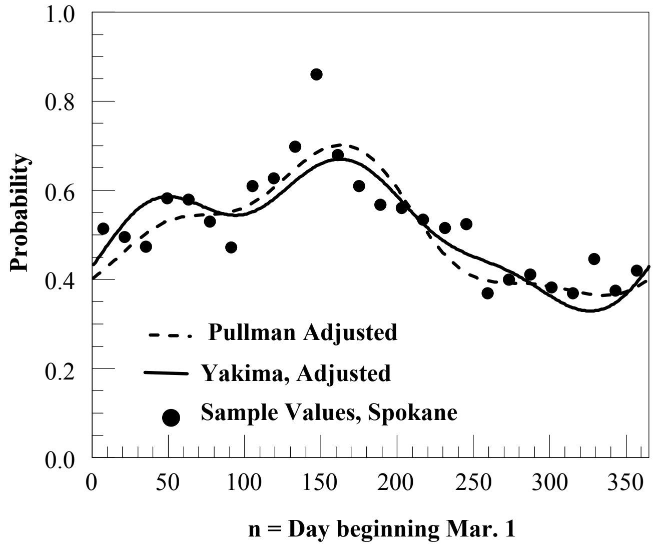

Figure 6c. Sample Markov chain parameter $P_{11\,0}$ for Spokane, WA and as Estimated from Pullman and Yakima.

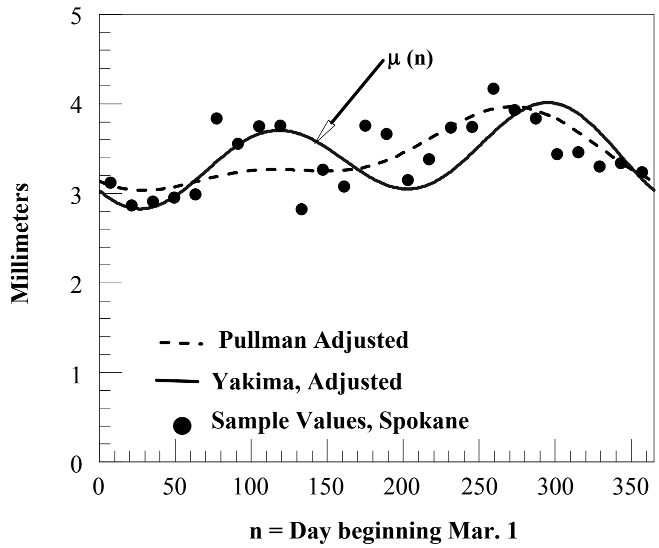

Figure 6d. Sample Mixed Exponential parameter, μ, for Spokane, WA and as Estimated from Pullman and Yakima.

Table 1. Yucca Mountain, Nevada Region Precipitation Data

Station	Elevation, m	Years of Record	MAN, Days	MAP, mm
Rainier Mesa	2283	41	53.3	314.2
Buster Jangle	1240	39	38.2	163.5
Cane Springs	1219	35	38.0	203.8
Desert Rock	991	36	33.3	148.7
Jackass Flats	1043	40	32.1	140.8
Tippipah Springs	1518	39	43.9	224.3
Well5B	939	38	32.7	123.3
Yucca Dry Lake	1196	40	36.0	173.3
Fortymile Canyon, North	1469	40	43.1	203.8
Mercury	1149	37	28.4	121.0
Las Vegas WSO Airport	1099	57	26.6	105.9

Table 2. Regression Coefficients for The Second Order Markov Chain, Yucca Mountain, NV Vicinity.

$$\Gamma_{i\,j\,0} = a + b\ \mathrm{MAP}$$

Parameter	a	b	R^2
$\Gamma_{0\,0\,0}$	3.220	–0.00248	0.877
$\Gamma_{1\,0\,0}$	2.243	–0.00197	0.567
$\Gamma_{0\,1\,0}$	1.058	–0.00417	0.745
$\Gamma_{1\,1\,0}$	1.173	–0.00329	0.708

Table 3. Yucca Mountain Region Regression Statistics for Mixed Exponential Distribution
$Y = a + b*MAP$, except for $C2_{\mu}$ where $Y = a*MAP^{b}$

Parameter, Y	a	b	R^2
α	0.350	—	NS*
$\overline{\mu}$	2.782	0.00899	0.870
$C1_{\mu}$	0.2513	0.00287	0.621
$\phi1_{\mu}$	−1.495	−0.00656	0.414
$C2_{\mu}$	0.0028	1.0923	0.860
$\phi2_{\mu}$	2.357	—	NS
$C3_{\mu}$	—	—	NS
$\phi3_{\mu}$	—	—	NS
$C4_{\mu}$	0.495	—	NS
$\phi4_{\mu}$	−3.955	—	NS
$\overline{\beta}$	0.297	0.00150	0.519
$C1_{\beta}$	0.1965	—	NS
$\phi1_{\beta}$	−3.516	—	NS

*NS is not significant