Nordic Hydrology, 29 (1), 1998, 21-40

Long-Term Analysis and Short-Term Forecasting of Dry Spells by Palmer Drought Severity Index

V. K. Lohani,[1] G. V. Loganathan,[2] and S. Mostaghimi[3]
Depts. of Eng. Fundamentals,[1] Civ. Eng.,[2] and
Biological Systems Eng.,[3]
Virginia Tech., Blacksburg, VA 24061, USA

This paper presents a non-homogeneous Markov chain approach for analyzing drought characteristics using the Palmer Drought Severity Index (PDSI). The probability mass functions of occurrence of different drought severity classes, their durations, and times of return to a particular drought class are obtained and are in turn utilized to generate the needed statistics and forecasts. Two methods of forecasting drought severity classes for one, two, and three months lead times are put forward. The methodology is applied to ninety six years of PDSI data corresponding to two climatic divisions in Virginia, USA. Comparison between the analytical results and the empirical estimates supports the utility of the method. The method can be used in a planning mode for developing buffer storage in drought prone regions and in an operational mode for optimal rationing of water among competing needs as drought progresses.

Introduction

Analysis of patterns of dry spells is necessary for planning long-term policies of water resources development in a region. It identifies drought prone regions and determines the appropriate buffer source development as a mitigation measure. In the operational mode during an ongoing drought, appropriate regulations for a meaningful allocation of water among uses can be devised. In this paper long-term records of Palmer Drought Severity Index (PDSI) are analyzed to assess dry spell patterns in Virginia, USA (see Fig. 1). Van Bavel and Lillard (1957) analyzed the historical pattern of droughts in Virginia and observed that from June through September (part of the growing season) the probability of moisture deficiency had been at least 3 years in 10. Vellidis *et al.* (1985) further ascertained the observations of Van Bavel and Lillard (1957). In a recent study conducted by the State Water Control Board (SWCB 1990), 9 drought years have been identified in Virginia during the period of 1957-87 which is again about 30% probability. Short-term forecasting is considered to mean prediction over a three month period during a crop growing season. The Palmer index takes into consideration meteorologic and hydrologic variables in a comprehensive manner to assign a quantitative measure to dryness and wetness (Palmer 1965). A well detailed analysis of the formulation of the PDSI is given in Karl (1983; 1986) and Alley (1984). Detailed reviews are also given in Guttman (1991) and Johnson and Kohne (1993). In the following a brief description of Palmer index computational scheme is given.

To begin with a long-term monthly water balance for the chosen region is carried out. In his original formulation Palmer used the Thornthwaite equation (Thornthwaite 1948) for estimating the potential evapotranspiration. The soil moisture is depleted at the potential rate minus the precipitation as long as moisture storage permits it; otherwise, entire available moisture is used up. The soil moisture storage is divided into two layers namely, the surface layer and the underlying layer. The evapotranspiration requirement is first met by the surface layer. When its storage is used up, transfer from the underlying layer is initiated. For soil moisture recovery, the surface layer storage must be full, before transfer to the underlying layer could take place. Surface runoff is computed by subtracting evapotranspiration and soil moisture storage deficiency amounts. The aforementioned balancing scheme is performed for a long period to obtain the average evapotranspiration, soil moisture recharge (refilling of moisture during large rain events), runoff, and soil moisture loss. When rainfall is set to zero, the evapotranspiration losses are the maximum called the potential loss; potential recharge is the maximum recharge possible for any given period and is taken as the total soil moisture storage minus the available soil moisture storage; potential runoff is taken as total soil moisture storage minus the potential recharge; and the potential evapotranspiration is obtained from the Thornthwaite equation.

The ratios of long-term averages of evapotranspiration, recharge, runoff, and losses to the long-term averages of their respective potential values are computed. These ratios are used as multipliers for the potential values and the resulting evapotranspiration, recharge, and runoff are added and soil moisture losses are subtracted to obtain that month's climatically appropriate for existing conditions (CAFEC) precipitation. The difference between the CAFEC precipitation and the actual precipitation is the precipitation deficit for that month. This deficit is rescaled by multiplying by a constant specific for a region and for a month to obtain an index that is spatially and temporally comparable among different locations. This attribute has been discussed in the literature (Alley 1984; Guttman *et al.* 1992). The resulting moisture anomaly index, called the *Z*-index, becomes part of the Palmer index. The *Z*-index indicates the wetness or dryness of an individual month. The current cumulative deficit index is taken as the weighted sum of the previous cumulative deficit index and the current *Z*-index. The current cumulative deficit index is then subjected to a backtracking scheme to assess moisture depletion or recovery phase with an associated change in the PDSI value itself. The National Climatic Data Center (NCDC) in Ashville, North Carolina computes the PDSI for the contiguous United States. For this purpose the country has been divided into 344 climatic divisions. The boundaries of these divisions were fixed in the late fifties to conform with climate influencing physical features (Johnson and Kohne 1993). Karl (1986) has grouped the PDSI values into seven classes with class 1 (PDSI greater than or equal to 4.00) being the wettest and class 7 (PDSI less than or equal to -4.00) being the driest class. The normal class (class 4) is considered when PDSI is in the range of -1.49 to 1.49. In his original formulation Palmer(1965) used climatic and other data from Kansas and Iowa for development of the PDSI. Since then the index has been used in several other states in the U.S. and various other countries for different purposes (Wheaton 1990; Johnson and Kohne 1993; Jones *et al.* 1996). Heddinghaus and Sabol (1991) reported that the principal uses of the PDSI are in monitoring hydrologic trends, crop forecasts, and assessing potential fire severity.

In this paper the focus is on a typical crop growing season. The feasibility of irrigation schemes depends on availability of water from primary sources like wells, impoundments, and streams. Because PDSI incorporates the deviations from the climatically appropriate runoff and the moisture deficit within the soil for a region, it is used as a measure of irrigation water deficit. Since Palmer's original formulation has a Markovian form, there have been several attempts to apply the Markov model to predict its behavior. For drought related analyses the Markov model has been applied in two forms. The first form is an autoregressive order one time series model in which the focus is more on expected values than the probabilities themselves. The second form is a Markov chain in which the variable takes values from discrete classes and the probabilities for various events can be readily computed. Rao and Padmanabhan (1984) adopted time series models to forecast the Palmer drought index. Eltahir (1992) applied an autoregressive order 1 model to generate rainfall values and performed a run analysis to select the drought events. The Palmer's original PDSI equation itself is an autoregressive model of order 1. However, because of the backtracking scheme used in the PDSI computation to place it in the dry or wet spell in a *posteriori* manner, neither PDSI nor a transformed form of it follows a normal distribution, in general. Therefore, the non-parametric form of the Markov chain provides a good alternative to analyze the PDSI data. There have been numerous applications of Markov chain to model wet/dry state transitions in the context of precipitation occurrence process. Good reviews are given in Bowles and O'Connell (1991). However, the authors are not aware of any non-homogeneous Markov chain approach for predicting drought behavior.

While the focus here is on Markov chain applications, Lohani(1995) presents an

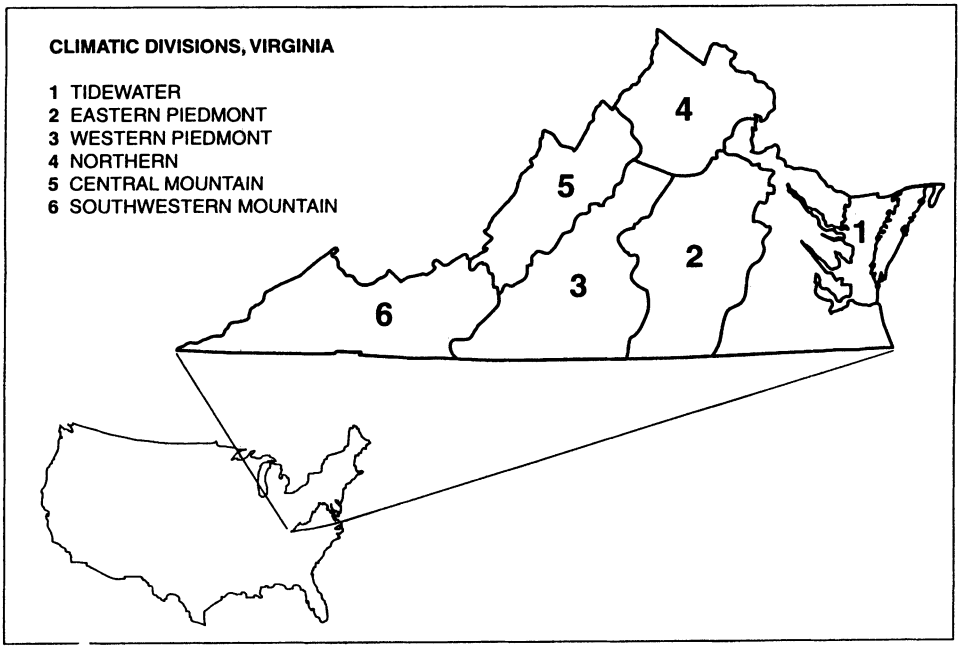

Fig. 1. Virginia Climatic Divisions.

exhaustive review of the drought literature. In this paper a novel non-homogeneous Markov chain approach is described for characterizing droughts in two climatic divisions, CD1 – the Tidewater Region and CD6 – the Southwest Mountains Region of Virginia. The state of Virginia has overall six climatic divisions (Fig. 1). This approach accounts for the non-stationary behavior of climatic variables characterizing droughts. The methodology is applied to assess: i) the probabilities of occurrence of the various drought classes for an area, ii) the duration of stay of a particular drought class, iii) the expected passage time to migrate from one drought class to another, iv) the expected recurrence time of a particular drought class, and v) the expected drought class for future periods.

The theoretical results are supported by the empirical analysis of ninety six years of monthly data for the period 1895-1990 for the two climatic divisions. As a performance measure the Kolmogorov-Smirnov test is applied to check the goodness of fit between the empirical and analytical probability distributions. The paper is organized as follows. In the ensuing section on the formulation detailed theoretical results are presented. The results are also given for the so called mean monthly homogeneous Markov Chain that does not differentiate between months. Within each section, the theoretical conclusions are illustrated with the real data application and the resulting numerical values are tabulated. The section entitled the Monthly Drought Forecast presents two schemes for forecasting the progression of a drought. Because, the method uses the conditional probability, the forecast accuracy can be improved based upon the current state of drought. However, for irrigation related forecast a three-month lead time may be required for planting the crops at the current month. The section on conclusions describes the pivotal contributions.

Non-Homogeneous Markov Chain Formulation

The class delineation of PDSI provides a natural means for considering it as a Markov Chain. Let X_n be the random variable for month, n, representing the drought (wet) class; class 7 being the driest and class 1 being the wettest. For example, $X_1 = X_{Jan} \equiv 1$ represents the occurrence of class 1 in January. A Markov chain completely describes the underlying stochastic process if the transition probabilities denoted by $p_{i,j}^{n,n+1}$ for moving from class i in month n to class j in month n+1 and the initial probability vector, $f^{(0)}$, describing the probabilities of the seven classes for the beginning month are prescribed. In the present study twelve monthly (non-homogeneous) transition probability matrices describing the class transfers from months January to February; February to March; ..., and December to January are formulated. It is hypothesized that these matrices are cyclic in the sense that there is no yearly variation. The transition probabilities depend only on the month and not on the year. The transition probability

$$p_{i,j}^{(n,n+1)} = P\left[X_{n+1} = j \mid X_n = i\right], \text{for } i,j = 1,2,\ldots,7 \text{ and } n = 1,2,\ldots,12 \quad (1)$$

is computed as

$$p_{i,j}^{(n,n+1)} = N_{i,j}^{(n,n+1)} / N_i^{(n)} \quad (2)$$

where $N_{i,j}^{(n,n+1)} \equiv$ number of transitions from class i in month n to class j in month n+1; $N_i^{(n)}$ = number of occurrences of class i in month n. If $N_i^{(n)}$ is zero for some, i, we define $p_{i,j}^{(n,n+1)} = 1/7$ for all j = 1,2,...,7. The monthly PDSI values for CD1 and CD6 of Virginia from 1895-1990 are analyzed.

Monthly Drought Class Probabilities

It is of interest to know the probabilities of occurrence of different drought classes for any given month. The drought proneness of a region is reflected in a large probability for the drought classes. As far as monitoring the drought progression is concerned, unless spring snowmelt is involved, the winter months of December through March may not be of interest as opposed to April through November. In Virginia there is no significant snowmelt. The probability mass functions for the various drought classes for these months help to assess the progression of a dry spell during a crop growing season. The probability mass functions related to the duration and times of return to a particular drought class can also be developed. Let $f^{(k)}$ be the class probability vector after k transitions. Using the initial class probability row vector $f^{(0)}$ and the monthly transition matrices we can write

$$f^{(k)} = [f^{(0)}]\,[P_1]\,[P_2] \cdots [P_k] \tag{3}$$

in which $f^{(0)}$ is initial class probability row vector and P_1 = (7 x 7) monthly transition matrix associated with the starting month, say January to February, *i.e.*, $P_1 = P^{(1,2)} = P^{(\text{Jan., Feb.})}$. Of course, the starting month can be any one of the 12 months. Also, due to the cyclic nature of these matrices, the transition matrix for months 14 to 15 denoted by $P^{(14,15)}$ is the same as $P^{(2,3)} = P^{(\text{Feb., Mar.})}$, the February-March transition matrix.

For the long-term, that is as $k \to \infty$ we would like to know whether $f^{(\infty)}$ has steady class probabilities independent of $f^{(0)}$. This will be true if the product of the transition matrices $[P_m]$ through $[P_k]$ denoted by $\phi^{(m,k)}$ called the composite matrix

$$\phi^{(m,k)} = [P_m]\,[P_{m+1}] \cdots [P_k] \tag{4}$$

is a constant stochastic matrix with identical rows (Isaacson and Madsen 1976) for large k for some starting m. For such a constant stochastic matrix it follows from Eq. (3) that $f_m^{(k)}$ will be independent of $f^{(0)}$; furthermore, each class has a steady class probability value corresponding to that class' (column) constant probability of $\phi^{(m,k)}$. However, because the beginning month, m, influences the value of $\phi^{(m,k)}$ the steady class probabilities of $f_m^{(k)}$ will depend on m. It says that the current month's drought behavior does affect the next month's drought pattern. To interpret $f_m^{(k)}$ as $k \to \infty$, consider Eq. (4) as follows. The constant (identical rows) stochastic matrix for January is defined as the product of the sets of the consecutive 12 monthly matrices with the beginning matrix being that of January which is

$$\begin{aligned}\phi^{(1,\infty)} = [\text{Jan}] = \{[P_1]\,[P_2] \cdots [P_{11}]\,[P_{12}]\} \\ \{[P_1]\,[P_2] \cdots [P_{11}]\,[P_{12}]\} \cdots\end{aligned} \tag{5}$$

Because [Jan] is a constant stochastic matrix it follows

$$\text{row}[\text{Jan}] = f_1^{(\infty)} \tag{6}$$

Now consider

$$\begin{aligned}\phi^{(2,\infty)} = [\text{Feb}] = \{[P_2]\,[P_3] \cdots [P_{11}]\,[P_{12}]\,[P_1]\} \\ \{[P_2]\,[P_3] \cdots [P_{11}]\,[P_{12}]\,[P_1]\} \cdots\end{aligned} \tag{7}$$

and we obtain

$$[\text{Feb}] \equiv [P_2][P_3]\dots[P_{11}][P_{12}][\text{Jan}][P_1] \tag{8}$$

from which upon manipulation it follows

$$[\text{Feb}] = [\text{Jan}][P_1] \tag{9}$$

and therefore

$$\text{row}[\text{Feb}] = f_2^{(\infty)} = \text{row}[\text{Jan}][P_1]$$

similarly we can show

$$\begin{aligned}
\text{row}[\text{Mar}] &= f_3^{(\infty)} = \text{row}[\text{Feb}][P_2] \\
\text{row}[\text{Apr}] &= f_4^{(\infty)} = \text{row}[\text{Mar}][P_3] \\
&\dots \\
\text{row}[\text{Dec}] &\equiv f_{12}^{(\infty)} \equiv \text{row}[\text{Nov}][P_{11}] \\
\text{row}[\text{Jan}] &= f_1^{(\infty)} = \text{row}[\text{Dec}][P_{12}]
\end{aligned} \tag{10}$$

Eq. (10) provides a means to evaluate the monthly drought/wet steady class probabilities. It is a system of linear equations in terms of the monthly steady class probabilities that are used to compute these probabilities for each month and class both for CD1 and CD6. Table 1 lists the sample results of steady class probabilities for January, July, and October months obtained using PDSI data for both of the divisions. The table also provides the empirical estimates of these probabilities using 1152 months of data of PDSI. For each month, two values of steady class probability are given for each class. The upper row represents the analytical result and the lower one is the empirical result. For example, the steady class probability of occurrence of class 3 in January month in CD1 is 0.2200 which compares well with the empirical estimate of 0.2187. Likewise, the steady class probability of occurrence of class 3 in January month in CD6 is 0.2037 which also agrees reasonably well with the empirical estimate of 0.1979 (see Table 1). These results imply that if weather conditions are categorized into seven different states using PDSI values, then there is about a 20% probability over long-term that the weather state will be class 3 in January in CD1 and CD6 climatic divisions of Virginia. It is also observed that during July through October steady class probabilities of drought classes (*i.e.* classes 5, 6, and 7) are higher in CD1 division as compared to CD6 division implying that CD1 region is relatively prone to drought conditions as compared to the CD6 division. The steady class probabilities obtained using a mean monthly transition matrix are also given in Table 1. This matrix is computed by considering weather class transitions over the entire data period (1895-1990) without considering the months of such transitions. Because of its inability to differentiate between months, this procedure is not applied for making forecasts. The analysis is elaborated in the section on mean monthly homogeneous Markov chain analysis. Table 1 also contains Karl's

Table 1 – Monthly Steady Class Probabilities, Analytical and Empirical.

Classes → Month ↓	1	2	3	4	5	6	7
CD1							
Jan. (A)	0.0104	0.0521	0.2200	0.5198	0.1146	0.0417	0.0417
(E)	0.0104	0.0521	0.2187	0.5208	0.1146	0.0417	0.0417
July (A)	0.0	0.0626	0.2085	0.4584	0.2083	0.0416	0.0208
(E)	0.0	0.0625	0.2083	0.4583	0.2083	0.0416	0.0208
Oct. (A)	0.0209	0.0625	0.2084	0.3959	0.2604	0.0312	0.0208
(E)	0.0208	0.0625	0.2083	0.3958	0.2604	0.0313	0.0208
CD6							
Jan. (A)	0.0209	0.0214	0.2037	0.4731	0.1978	0.0520	0.0312
(E)	0.0208	0.0208	0.1979	0.4792	0.1979	0.0521	0.0312
July (A)	0.0210	0.0209	0.2191	0.4999	0.1769	0.0520	0.0104
(E)	0.0208	0.0208	0.2188	0.5000	0.1771	0.0521	0.0104
Oct. (A)	0.0209	0.0626	0.1772	0.4897	0.2082	0.0312	0.0104
(E)	0.0208	0.0625	0.1771	0.4896	0.2083	0.0313	0.0104
12 Months' Average							
CD1	0.0113	0.0591	0.1956	0.4886	0.1771	0.0451	0.0234
CD6	0.0174	0.0341	0.1868	0.5130	0.1890	0.0477	0.0121
Mean Monthly Matrix							
CD1	0.0113	0.0590	0.1952	0.4878	0.1771	0.0451	0.0234
CD6	0.0175	0.0342	0.1869	0.5127	0.1889	0.0477	0.0121
Karl (1986)	0.0500	0.0600	0.1700	0.4500	0.1700	0.0600	0.0400
Empirical 12 Months' Average							
CD1	0.0113	0.0590	0.1953	0.4887	0.1771	0.0451	0.0234
CD6	0.0156	0.0338	0.1858	0.5139	0.1892	0.0478	0.0121

(A)= analytical; (E)= empirical

(1986) empirical estimates of the steady class probabilities computed for the entire USA across all months.

Wallis (1993) reported that the probability of PDSI being in class 7 for the month of July for CD1 was between 0.00 and 0.05 and the present analysis yields 0.0208. Further, Guttman *et al.* (1992) observed probabilities of PDSI value being in class 7 in January month as ranging between 0.01-0.05 for CD6. Our analysis gives this value as 0.0312 for CD6. In order to further verify the analytical results the Kolmogorov-Smirnov test was applied with the null hypothesis that the analytical cumulative distribution function and the empirical distribution function are the same. The results indicate that the null hypothesis is true at the 10% significance level. These results validate the use of the non-homogeneous Markov chain technique in evaluating the long-term probabilities of drought classes using the Palmer index. In the following sections the further use of the proposed non-homogeneous Markov chain model in characterizing the durations and the periods of occurrences of the various drought classes is illustrated.

Expected Uninterrupted Residence Time

The expected uninterrupted residence time of the process in a class indicates the duration of that drought class for that region. The process stays in class i without migrating to another class for 'm' time periods when the following event occurs

$$\{X_1 = i = X_2 = \ldots = X_{m-1} \mid X_0 = i\} \tag{11}$$

The probabilities of events specifying uninterrupted stay for different time periods in a particular class, i, can be computed as follows. For example, the probability of one month duration of stay denoted by m=1 starting with the month of January is given by

$$P[X_{\text{Feb}} \neq i \mid X_{\text{Jan}} = i] = P[m=1 \mid X_{\text{Jan}} = i] = 1 - p_{i,i}^{1,2} \tag{12}$$

where: $p_{i,i}^{1,2}$ = probability of moving from class i in January to the same class i, in February. Eq. (12) says that the drought class 'i' is occupied for the thirty one days in January, *i.e.* 1 month and then on the first of February the drought class is no longer 'i' but some other $j \neq i$. This interpretation also says that transition occurs on the last day of the given month. For m=2 starting from the month of January the probability can be computed by

$$P[m=2 \mid X_{\text{Jan}} = i] = P[X_{\text{Mar}} \neq i, X_{\text{Feb}} = i \mid X_{\text{Jan}} = i] \tag{13}$$

$$= P[X_{\text{Feb}} = i \mid X_{\text{Jan}} = i] P[X_{\text{Mar}} = i \mid X_{\text{Feb}} = i] \tag{14}$$

$$= p_{i,i}^{1,2}(1 - p_{i,i}^{2,3}) \tag{15}$$

Likewise, the probabilities of events defining consecutive stay for higher number of time steps can be computed. For example, for staying 12 time steps (*i.e.* 12 months) consecutively in class i, starting in January, the probability will be

$$P[m=12 \mid X_{\text{Jan}} = i] = p_{i,i}^{1,2} p_{i,i}^{2,3} \cdots p_{i,i}^{10,11} \; p_{i,i}^{11,12} \; (1 = p_{i,i}^{12,1}) \tag{16}$$

Table 2 – Expected Uninterrupted Residence Times (months), Analytical and Empirical

Class → Starting Month ↓	1	2	3	4	5	6	7
CD1							
Jan. (A)	2	1.2	2.6	5.3	3.6	2.6	4.0
(E)	2	1.2	2.8	6.3	2.9	2.9	4.0
July (A)	1.1	2.2	2.4	4.3	3.5	2.0	1.0
(E)	-	2.2	2.5	5.6	3.4	1.5	1.0
Oct. (A)	1.5	2.2	3.1	5.0	2.8	2.2	5.3
(E)	1.5	2.0	2.7	6.1	2.8	2.3	8.5
CD6							
Jan. (A)	1.5	1,8	2.4	5.2	3.1	1.7	1.7
(E)	1.5	1.5	3.1	7.0	3.2	1.8	1.7
July (A)	1.0	1.0	2.5	4.9	4.2	2.0	6.8
(E)	1.0	1.0	3.0	5.9	4.5	2.0	9.0
Oct. (A)	2.8	1.7	2.8	4.8	3.5	2.6	3.8
(E)	3.0	1.7	3.9	6.1	4.3	2.3	6.0

Note : — denotes unavailability of data to compute empirical value; (A)= analytical; (E)= empirical

It is readily observed that the computation of probabilities for various events defining an uninterrupted stay in class i involves the multiplication of i^{th} row and i^{th} column entries (diagonal elements) of the consecutive transition matrices. If any one of these entries is zero, the computation stops at that point because all the remaining probabilities for higher duration of stay go to zero. Once the probabilities for uninterrupted stay for various time periods are computed, the expected uninterrupted residence time for class j, $E[R_{uj}]$, is given by

$$E\left[R_{uj}\,|\,\text{starting month}\right] = \sum_{k} k\,P\left[m=k\,|\,\text{starting month}\right] \qquad (17)$$

where $R_{uj} \equiv$ random variable describing uninterrupted stay in class j.

Eq. (17) is used to compute expected uninterrupted residence times for each starting class and month in both CD1 and CD6. It is obvious that computations of average uninterrupted residence times in respect of dry classes 5, 6, and 7 will indicate duration of drought in the region of interest. Table 2 gives sample results for starting months of January, July, and October in respect of both CD1 and CD6. There are two values reported for each month and class. The upper value gives the analytical estimate and the lower one gives the empirical value. It is seen that the analytical results agree with the empirical observations. It is noted that the empirical estimates are generated from only a fewer occurrences for certain drought classes. It can be observed in Table 2 that if weather state 7 (extreme drought condition) occurs in October month in CD1 it is expected to last on an average for 5.3 months while in CD6 it is expected to last on an average of 3.8 months. It is also observed that the rows in Table 2 are not significantly different for various months which may be interpreted as that the average uninterrupted residence time does not depend on the months but depends only on the class of the weather. Further, the longest residence time is for class 4 for both CD1 and CD6 which implies that the weather tends to stay normal for relatively longer period for both the Tidewater and Southwest Mountain regions. The average continuous residence times for drier classes (5, 6, and 7) are greater than that of the wet classes (1, 2, and 3) which indicates that a drought spell, once occurred, would stay for a relatively longer period than a wet spell. In order to examine the closeness between the analytical and empirical distributions the Kolmogorov – Smirnov test was applied. The results indicate that the analytical distribution is the same as the empirical distribution at 10% significance level. The results of the residence time analysis are useful in interpreting the persistence of various weather classes in different climatic regions.

Expected First Passage Times

The expected first passage time is defined as the average time period taken for the process to go to a class, *j*, for the first time starting from some class, *i*, and is denoted by $m_{i,j}$. For a non-homogenous chain the starting month, *n*, is crucial in deciding the expected first passage time and therefore we let $m_{i,j}^{(n)}$ as the first passage time for a process to reach class, *j*, starting from class, *i*, in month, *n*. Mathematically it attains the form

$$m_{i,j}^{(n)} = (1)\,p_{i,j}^{(n,n+1)} + \sum_{k \neq j} p_{i,k}^{(n,n+1)} \left(m_{k,j}^{(n+1)} + 1\right) \qquad (18)$$

in which the first term says that class *j* can be reached in one step in month $(n+1)$ or the process can go to some class $k \neq j$ in one step as is indicated by 1 in summation term and it takes $m_{k,j}^{(n+1)}$ steps to reach *j*. This equation is simplified by combining the first term probability $p_{i,j}^{(n,n+1)}$ with the remaining sum of probabilities in the summation term to yield

$$m_{i,j}^{(n)} \equiv 1 + \sum_{k \neq j} p_{i,k}^{(n,n+1)} m_{k,j}^{(n+1)} \qquad (19)$$

For example, for $n=1$ we obtain

$$m_{i,j}^{(1)} \equiv m_{i,j}^{\text{Jan}} \equiv 1 + \sum_{k \neq j} p_{i,j}^{(1,2)} m_{k,j}^{\text{Feb}} \tag{20}$$

in which $p_{i,k}^{(1,2)} \equiv$ January-February transition probability for reaching class, k, from class, i. The solution of system of linear equations in Eq. (19) yields the expected first passage times. The average time to return to the same class, called the mean recurrence time, m_{ii}, can also be computed from Eq. (18) as

$$m_{i,i}^{(n)} = (1) p_{i,i}^{(n,n+1)} + \sum_{k \neq i} p_{i,k}^{(n,n+1)} (m_{k,i}^{(n+1)} + 1) \tag{21}$$

which simplifies to

$$m_{i,i}^{(n)} \equiv 1 + \sum_{k \neq i} p_{i,k}^{(n,n+1)} m_{k,i}^{(n+1)} \tag{22}$$

Eqs. (19) and (22) are applied to compute the average first passage and the recurrence times both for CD1 and CD6. The first passage and the recurrence times are

Table 3 – Expected First Passage Time and Recurrence Time to Class 4 (months), Analytical and Empirical

Class → Starting Month ↓	1	2	3	4	5	6	7
CD1							
Jan. (A)	8.1	4.7	4.6	1.8	5.2	6.8	11.2
(E)	4.0	3.8	5.2	1.4	4.3	8.0	15.5
July (A)	6.5	7.4	4.4	2.6	6.6	6.2	5.1
(E)	-	7.2	5.2	2.1	7.2	12.5	3
Oct. (A)	5.7	6.0	4.5	2.3	4.9	9.8	10.1
(E)	2.5	6.2	4.6	2.5	5.6	11.3	17
CD6							
Jan. (A)	6.5	3.0	3.3	1.8	5.4	6.3	3.5
(E)	28.0	4.5	6.0	1.6	5.9	4.6	3.7
July (A)	7.0	4.5	3.7	1.9	7.2	6.0	10.4
(E)	34.0	10.5	4.1	2.0	7.3	5.4	9.0
Oct. (A)	5.3	4.1	3.9	2.0	6.9	7.7	7.4
(E)	2.5	6.3	4.9	1.7	7.2	10.0	6.0

Note: - denotes unavailability of data to compute empirical result; (A)= analytical; (E)= empirical useful in exploring things like time to relief from dry or wet weather conditions. Sample results of the first passage time to class 4 from various classes starting in January, July, and October months for CD1 and CD6 are given in Table 3. There are two values reported for each month and class. The upper one represents the analytical value while the lower one is the empirical value. For example, it is seen that the process will take on an average 6.8 months to go directly to class 4, once class 6 has occurred in January in CD1 which compares well with the empirical result of 8.0 months. For the CD6 region the first passage time from class 6 to class 4 was computed as 6.3 months which again compares reasonably well with the empirical result of 4.6 months. Further, the recurrence time for class 4 in row 1 is 1.8 months in CD6 in January. Intuitively, let us visualize only three classes namely wet, normal, and dry. From Table 1 these classes have approximate occurrence probabilities of 0.25, 0.47, and 0.28, respectively for CD6 in January. From Table 3 to have a recurrence time of 1.8 months for class 4, the process should migrate to wet or dry class for a duration of 1.8 months. It is observed in Table 2 that for CD6 average uninterrupted stay in dry classes (5, 6, and 7) is 2.2 months and wet classes (1, 2, and 3) is 1.9 months for January month which are close to recurrence time for class 4 in January.

It may be noted that there have been limited empirical events in classes 7 and 1 and therefore, the empirical estimates for such cases tend to differ from the analytical values. The Kolmogorov-Smirnov test was applied to test the equality of the analytical and empirical distributions describing the first passage times from class 5 to 4 with the starting month of January. The results indicate that the analytical and empirical distributions are the same at 10% significance level. It is also seen that in general the expected first passage times to class 4 from drier classes is greater than wet classes in both the divisions. This indicates that in these climatic divisions it takes relatively longer time to return to normal class from drought classes as compared to the wet classes.

Monthly Drought Forecast

In this section two schemes for making forecast of weather class are given.

Scheme A : Suppose k^* denotes the most likely weather class for month $(n+1)$, given that during the month n the observed weather class is, i. Then we can write

$$\text{PMODE}\left[X_{n+1} \,|\, X_n = i\right] = k^* \tag{23}$$

wherein k^* is the weather class that relates to $\text{Max}(p_{i,k}^{(n,\,n+1)})$ for all k for given i. Likewise,

$$\text{PMODE}\left[X_{n+2} \,|\, X_{n+1} = k^*\right] \equiv 1^* \tag{24}$$

wherein l^* relates to $\text{Max}(p_{k^*,l}^{(n+1,\,n+2)})$ for all l for given k*. Similarly,

$$\text{PMODE}\left[\underset{=}{X}_{n+3} \,|\, X_{n+2} = 1^*\right] = m^* \tag{25}$$

wherein m^* relates to $\text{Max}(p_{l^*,m}^{(n+2,\,n+3)})$ for all m for given l^*. It is seen that m^* is the forecast of the weather class in (n+3) month, given that during n^{th} month the weather class is i.

Table 4 – Predicted Weather Classes- One, Two, and Three Months Ahead

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
CD6												
Obs. 1930	4	5	5	5	5	6	7	7	7	7	7	7
Pr. (A)*	4	4	5	5	5	5	6	7	7	7	7	7
Pr. (A)**	3	4	4	5	5	5	5	4	7	7	7	7
Pr. (A)***	4	3	4	4	5	5	5	5	4	7	7	7
Pr. (B)*	4	4	5	5	5	5	6	7	7	7	7	7
Pr. (B)**	4	4	4	5	4	4	4	5	7	7	7	7
Pr. (B)***	4	4	4	4	4	4	4	5	5	7	7	7
Obs. 1990	2	2	3	3	2	3	3	3	4	3	3	3
Pr. (A)*	2	2	2	3	3	2	3	3	3	4	3	3
Pr. (A)**	1	2	2	2	3	3	1	3	3	3	4	3
Pr. (A)***	1	1	2	2	2	3	3	2	3	3	3	4
Pr. (B)*	2	2	2	3	3	2	3	3	3	4	3	3
Pr. (B)**	4	4	4	4	4	4	3	4	4	4	4	4
Pr. (B)***	4	4	4	4	4	4	4	3	4	4	4	4
CD1												
Obs. 1930	4	4	5	5	5	5	6	6	7	7	7	7
Pr. (A)*	4	4	4	5	5	5	5	6	6	7	7	7
Pr. (A)**	3	4	4	4	5	5	5	5	6	6	4	7
Pr. (A)***	3	3	4	4	4	5	5	5	5	6	6	4
Pr. (B)*	4	4	4	5	5	5	5	6	6	7	7	7
Pr. (B)**	4	4	4	4	4	5	5	5	4	7	4	7
Pr. (B)***	4	4	4	4	4	5	5	4	5	4	4	4
Obs. 1990	2	3	3	3	2	4	4	4	4	4	4	4
Pr. (A)*	2	3	3	3	3	4	4	4	4	4	4	4
Pr. (A)**	2	3	3	3	3	3	4	4	4	4	4	4
Pr. (A)***	2	3	3	3	3	3	3	4	4	4	4	4
Pr. (B)*	2	3	3	3	3	4	4	4	4	4	4	4
Pr. (B)**	4	4	4	4	4	4	4	4	4	4	4	4
Pr. (B)***	3	4	4	4	4	4	4	4	4	4	4	4

Obs. = observed; Pr. (A)= scheme A prediction; Pr. (B)= scheme B prediction; *= 1 month ahead; **= 2 months ahead; ***= 3 months ahead

Scheme B : In this scheme the p months ahead forecast is defined as

$$\text{PMODE}[X_{n+p} | X_n = i] = r^* \qquad (26)$$

wherein r^* is the weather class that relates to Max $P[X_{n+p}|X_n=i]$ where

$$P[X_{n+p} | X_{n=i}] = \sum_k \sum_l P[X_p = j | X_{p-1} = k] P[X_{n+2} = k | X_{n+1} = 1] P[X_{n+1} = 1 | X_n = i]$$

$$\text{for} \quad i = j = k = 1 = 1, 2, \ldots, 7 \qquad (27)$$

Using the above schemes, the conditional forecast of weather classes 1, 2, and 3 months ahead of time for years 1930 (relatively dry year) and 1990 (relatively wet year) in both CD1 and CD6 are given in Table 4. It is observed from Eqs. (23) and (26) that both schemes yield similar forecasts for one month ahead of time. It is also seen that Eqs. (23) – (25) treat only the modal values as given values for the prediction whereas Eq. (27) considers all possible intermediate classes given by k and l. From Table 4 it is seen that the forecasts are acceptable. Lohani and Loganathan (1997) also point out that in the context of monitoring an ongoing drought, with the aid of Eq. (27), all paths of drought progression can be enumerated.

Mean Monthly Homogeneous Markov Chain Analysis

In order to characterize the mean monthly transition behavior of weather, transition probabilities are computed giving emphasis on transitions among classes, irrespective of the months in which these take place during the year. If N_{ij} is the number of times the process transits from class i to class j, regardless of the month, and N_i is the total number of times the process is in class i, then

$$p_{ij} = \frac{N_{ij}}{N_i} \qquad (28)$$

denotes the mean monthly probability of transition from class i to class j. In this manner the mean monthly transition matrix can be defined as $P = [\, p_{i,j} \,]$ for i,j = 1,2, ..., 7. The mean monthly matrices represent a closed communicating class of all weather classes (1, 2, ..., 7). This is logical since in reality any class of weather is possible in a mean monthly transition. Further, the absence of any transient or absorbing classes indicates that neither there exists a weather class from which the system disappears forever nor there exists a permanent weather class in which the system is trapped.

Drought Class Probabilities

The mean monthly matrix represents an irreducible aperiodic Markov chain which has limiting probabilities which are independent of starting class. These steady class probabilities, denoted by the vector λ, are computed as the non-negative solution of (Ross 1989)

$$\lambda_j = \sum_{i=1}^{7} \lambda_i p_{ij} \quad \text{and} \quad \sum_{j=1}^{7} \lambda_j = 1.0 \tag{29}$$

The steady class probabilities computed using mean monthly matrices are computed for both CD1 and CD6. The results for CD1 and CD6 are given in Table 1. It is seen that these probabilities are quite close to the average values computed using the 12 different monthly matrices and the empirical probabilities. The long-term probability for class 7 (extreme drought class) is found 0.0234 in CD1 which is only 0.0121 in CD6 (see Table 1). Similarly, for normal weather class (class 4) the long-term probability in case of CD1 is 0.4878 and for CD6 it is 0.5127. These results further indicate that CD1 is slightly more prone to drought conditions as compared with the CD6.

Mean Recurrence Time

The mean recurrence time, m_{jj}, for a class j is defined as the expected number of transitions until a Markov chain, starting in class j, returns to that class. Because for once in m_{jj} a visit to j occurs, the smaller the m_{jj} more visits occur for j; and the larger the m_{jj} fewer visits occur for j. It can be shown that (Hoel *et al.* 1972)

$$m_{ij} = \frac{1}{\lambda_j} \tag{30}$$

Table 5 gives the mean recurrence times of class 4 computed using average of 12 monthly matrices computations, mean monthly matrix, and observed data. These results in general agree well except for some discrepancies in the extreme classes. It is seen that class 4 has a recurrence period of about 2 months for both CD1 and CD6.

Mean First Passage Time

The mean monthly transition matrix can be analyzed to the find the mean first passage time which gives an estimate of the number of transitions the process takes, on the average, to go from one particular class to another for the first time. The mean first passage time is computed as

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj} \quad (31)$$

where m_{ij} ≡ mean first passage time to go from class i to j; p_{ik} = one step transition probability to go from class i to k; m_{kj} ≡ mean first passage time to go from class k to j. Table 5 gives the mean first passage times from various classes to class 4 in both CD1 and CD6. It is seen that on an average it takes longer to return to normal class from drought classes in CD1 as compared to CD6. This further indicates that CD1 region has higher persistence of drought, once occurred, as compared to CD6 region. This inference is in agreement with SWCB(1990) which found that during 1957-1987 the CD1 region experienced 8 drought years while CD6 region had only 5 drought years.

Table 5 – Mean Recurrence Times and Mean First Passage Times to Class 4

Location ↓ Class →	1	2	3	4	5	6	7
CD6							
12 Monthly Matrices' Average	6.2	4.7	3.5	1.9	6.3	6.9	5.7
Mean Monthly Matrix	6.1	4.8	3.6	2.0	6.3	7.0	5.5
Empirical	24.8	9.7	5.1	1.9	6.5	5.9	5.9
CD1							
12 Monthly Matrices' Average	6.6	5.8	4.6	2.1	5.6	7.9	8.5
Mean Monthly Matrix	6.0	5.9	4.6	2.1	5.8	7.6	8.2
Empirical	4.3	5.4	5.0	2.1	5.9	10.7	13.2

Table 6 – Average Uninterrupted Residence Times (months) and Empirical Average Residence Times

Location ↓ Class →	1	2	3	4	5	6	7
CD6							
12 Monthly Matrices' Average	2.5	1.6	2.5	4.8	3.4	2.0	2.7
Empirical Average	2.5	1.8	2.5	4.9	3.5	2.1	2.8
Mean Mon. Matrix	2.5	1.8	2.5	4.9	3.5	2.1	2.8
CD1							
12 Monthly Matrices' Average	1.3	1.8	2.7	4.8	3.3	2.3	3.2
Empirical Average	1.4	1.9	2.7	4.9	3.3	2.2	3.9
Mean Mon. Matrix	1.4	1.8	2.8	4.9	3.3	2.3	3.5

Mean Uninterrupted Residence Times

Table 6 gives average uninterrupted residence times computed using the 12 monthly matrices, the mean monthly matrix, and the empirical data. The results using the three methods agree very well. It is observed that on an average the drought classes tend to stay uninterrupted slightly longer in CD1 as compared to CD6 which again indicates higher persistence of drought classes in CD1 as compared to CD6.

Conclusions

The analysis of the Palmer drought severity index using the non-homogeneous Markov chain provides useful drought characteristics such as long-term drought probabilities, duration of drought termed as uninterrupted residence times, and expected periods of recovery computed as the first passage and recurrence times. The method provides probability mass functions of these parameters which define durations and periods of recurrence of various drought classes. Analysis of historical PDSI data of two climatic divisions (CD1 and CD6) in Virginia indicates that CD1 region has higher persistence of drought, once occurred, as compared to CD6 region. Application of the Kolmogorov-Smirnov test indicates that the analytical distributions of drought characterizing parameters computed using the non-homogeneous Markov chain procedure are the same as the empirical distributions at the 10% significance level. The proposed two schemes for predicting future drought classes with 1, 2, and 3 months of lead time yield good forecasts.

The three month ahead forecasts are quite useful in the context of irrigation. The methodology provides a comprehensive characterization of drought. In a planning mode the methodology identifies drought prone regions. This information can be used in planning and development of proper additional storage facilities. In an operational mode with the aid of the probability mass functions, the future course of a drought can be determined. In fact, in March 1995 the method signaled that 1995 would be a drought year in Virginia which was again confirmed in May 1995 as the dry spell progressed. Eventually, the State Governor declared a drought emergency in Virginia in September 1995. The method confirmed well ahead of time that 1996 would be a wet year in Virginia. In this paper many novel ideas related to drought forecasting have been put forward. The real time testing has so far validated the method.

Acknowledgment

The authors appreciate the constructive comments of the anonymous reviewers.

References

Alley, W. A. (1984) The Palmer drought severity index: limitations and assumptions, *Journal of Climate and Applied Meteorology, Vol. 23*, pp. 1100-1109.

Bowles, D. S., and O'Connell, P. E. (eds.) (1991) *Recent Advances in the Modeling of Hydrologic Systems*, Kluwer Academic Publishers, Boston.

Eltahir, E. A. B. (1992) Drought Frequency Analysis of Annual Rainfall Series in Central and Western Sudan, *Hydrological Sciences, Vol. 37, No. 3*, pp. 185-200.

Guttman, N. B. (1991) A sensitivity analysis of the Palmer hydrologic drought index, *Water Resources Bulletin, Vol. 27*, No. 5, pp. 797-807.

Guttman, N. B., Wallis, J. R., and Hosking, J. R. M. (1992) Spatial comparability of the Palmer drought severity index, *Water Resources Bulletin, Vol. 28, No. 6*, pp. 1111-1119.

Heddinghaus, T. R., and Sabol, P. (1991) A Review of the Palmer Drought Severity Index and Where Do We Go From Here ?, In: Proceedings 7th Conference on Applied Climatology, September 10-13, 1991, American Meteorological Society, Boston, Massachusetts pp. 242-246.

Hoel, P. G., Port, S. C., and Stone, C. J. (1972) *Introduction to Stochastic Processes*, Houghton Mifflin Co., Boston.

Isaacson, D. L., and Madsen, R. (1976) *Markov Chains: Theory and Applications*, John Wiley, New York.

Johnson, W. K., and Kohne, R.W. (1993) Susceptibility of reservoirs to drought using Palmer index, *Journal of Water Resources Planning and Management, Am. Soc. of Civ. Eng., Vol. 119*, No. 3, pp. 367-387.

Jones, P. D., Hulme, M., Briffa, K. R, Jones, C. G., Mitchell, J. F. B., and Murphy, J. B. (1996) Summer Moisture Accumulation over Europe in the Hadley Center General Circulation Model Based on the Palmer Drought Severity Index, *International Journal of Climatology, Vol. 16*, No. 2, pp. 155-172.

Karl, T. R.(1983) Some spatial characteristics of drought duration in the United States, *Journal of Climate and Applied Meteorology, Vol. 22*, pp. 1356-1366.

Karl, T. R. (1986) The sensitivity of the Palmer drought severity index and Palmer's Z index to their calibration coefficients including potential E. T., *Journal of Climate and Applied Meteorology, Vol. 25*, pp. 77-86.

Lohani, V. K. (1995) Characterization of Palmer Drought Index as a Precursor for Drought Mitigation, Ph.D. Dissertation, Virginia Tech, Blacksburg, Virginia.

Lohani, V. K., and Loganathan, G. V. (1997) An Early Warning System for Drought Management Using the Palmer Drought Index, *Journal of the American* Water *Resources Association, Vol. 33, No. 6* (to appear).

Palmer, W. C.(1965) Meteorological drought, U.S. Weather Bureau, Research Paper no. 45, Washington, D. C.

Rao, A. R., and Padmanabhan, G. (1984) Analysis and Modeling of Palmer's Drought Index Series, *Journal of Hydrology, Vol. 68*, pp. 211-229.

Ross, S. M. (1989) *Introduction to probability models*, 4th edition, Academic press, New York.

State Water Control Board (SWCB) (1990) Drought indicator parameters, Information Bulletin No. 582, Richmond, Virginia.

Thornthwaite, C. W. (1948) An Approach Toward a Rational Classification of Climate, *Geological Review, Vol. 38*, pp. 55-94.

Van Bavel, C. H. M., and Lillard, J. H. (1957) Agricultural drought in Virginia, Virginia Agriculture Experiment Station, Technical Bulletin No. 128, Blacksburg, Virginia.

Vellidis, G., Ross, B. B., and Taylor, D. B. (1985) *Agricultural drought in Virginia 1930-83,* Information Series 85-3, College of Agriculture & Life Sciences, Virginia Tech, Blacksburg, VA.

Wallis, J. R. (1993) *The National drought atlas* Proc. 20th Anniversary Conference on Water Management in the '90s, ASCE, Seattle, Washington.

Wheaton, E. E. (1990) Frequency and severity of drought and dust storm, *Canadian Journal of Agricultural Economics, Vol. 38*, pp. 695-700.

Received: 14 June, 1996
Revised: 4 June, 1997
Accepted: 16 July, 1997

Address:
G. V. Loganathan, Civ.Eng.Dept.,
200 Patton Hall,
Virginia Tech.,
Blacksburg, VA-24061-0105,
U.S.A.
Email: gvlogan@vt.edu
V. K. Lohani, Dept. of Eng. Fundamentals,
S. Mostaghimi, Biological Systems Eng.,