Turkish J. Eng. Env. Sci.
28 (2004) , 181 – 190.
© TÜBİTAK

Simulation of Drought Periods Using Stochastic Models

Kadri YÜREKLİ, Ahmet KURUNÇ
Gaziosmanpaşa University, Faculty of Agriculture, Tokat-TURKEY
e-mail: kadriyurekli@yahoo.com
Osman ÇEVİK
Gaziosmanpaşa University, Faculty of Economic and Administrative Sciences, Tokat-TURKEY

Received 07.11.2003

Abstract

This paper presents a methodology for the modeling of the agricultural drought duration in the Tokat region. For this purpose, the study area was divided into 4 hydrologically homogeneous regions. A rain gauge with the longest observation period was selected for each region. Days with rainfall equal to or less than a threshold of 6.4 mm and days with no rainfall were assumed to be drought periods in the 4 hydrologically homogeneous regions. The monthly data sequences were constituted by counting days with rainfall equal to or less than the threshold value of 6.4 mm and days with no rainfall. Linear stochastic models known as ARIMA were used to simulate drought periods in the hydrologically homogeneous regions. For 5 years, the predicted data using the best models were compared to the observed data. The results showed that the predicted data represent the actual data very well for each hydrologically homogeneous region.

Key words: Rainfall, Hydrologically homogeneous region, ARIMA model, Drought period.

Introduction

Drought is one of the most serious problems for human societies and ecosystems arising from climate variability. Although its impact does not come through sudden events, such as flood and storms, drought is one of the most damaging types of natural disasters over long periods. Human beings often increase the impact of drought because of a high use of water that cannot be supported when the natural supply decreases. Drought is difficult to define precisely, but operational definitions often help to define the onset, severity and end of droughts. Le Houerou (1996) stated that droughts are experienced in almost all types of agricultural land in the world, but that arid lands are most susceptible.

Drought is classified as agricultural, hydrological or meteorological. Agnew and Warren (1996) described agricultural drought as a spatial phenomenon that causes significant reductions in agricultural productivity, mainly due to an inadequate supply of soil moisture. Hydrological drought refers to deficiencies in surface and subsurface water supplies (Palmer, 1965). Meteorological drought is usually measured by how far from normal the precipitation has been over a certain period of time (Agnew, 1990).

Ranking the severity of droughts in cropping areas is difficult, due to the varying impact of rainfall at different times of the year. Drought intensity and duration must always be related to a calendar of crop sensitivity to rainfall. Assessing drought severity requires a measure of effective rainfall in relation to soil moisture and plant conditions, rather than just totaling rainfall deficiencies (Wilhite and Glantz, 1985). In addition, agricultural drought durations are affected by soil moisture capacity and evapotranspiration (Okman, 1981). As well as being directly affected by drought, crop yield is a regionally sensitive measure that integrates the temporal and spatial distribution of rainfall anomalies over a region. In a mix of cropping and livestock activities, it is the agronomic indicator most sensitive to rainfall fluctuations (Waring, 1976).

Hersfield *et al.* (1973) stated that days with less than 6.4 mm of rainfall or less than 0.003 mm of runoff cannot supply effective moisture accumulation in soil and cause agricultural drought. Oliver (1961) suggested that rainfall of up to 5.0 mm in 1 day should be regarded as useless since it will evaporate before infiltrating into the soil. Howe and Rhoades (1955) stated that rainfall of up to 3.0 mm in 1 day never contributes to crop water requirements. However, Richey *et al.* (1961) proposed that the temperature in the surroundings decreases due to insufficient amounts of rainfall that cannot infiltrate into the soil but evaporate and crop water consumption is thus reduced. Therefore, they suggested that this rainfall is also useful for crops.

There has been considerable research on modeling for various aspects of drought, such as the identification and prediction of its duration and severity. There exist a variety of techniques and methods to analyze the duration and severity of droughts through probability characterization of low flow, time series methods, synthetic data generation, theory of runs, multiple regression, group theory, pattern recognition and neural network methods. The prediction aspects of drought duration are better developed than the drought severity aspects. A major challenge facing drought research is to develop suitable methods and techniques for forecasting the onset and termination points of drought (Panu and Sharma, 2002).

In times of drought, agricultural productivity (particularly food productivity) declines significantly. A period of only a few weeks without precipitation may cause serious problems for the farmer. In arid and semi-arid regions, drought effects on crops may be reduced or eliminated by reserving enough water from available sources. Therefore, it is very important to know the drought duration in the growing season in terms of scheduling irrigation. In this study, it is proposed to simulate drought durations by the autoregressive integrated moving average (ARIMA) model in cropping areas.

Materials and Methods

Tokat province, selected as the study area, is bounded by latitudes 39° 45′ N and 40° 45′ N, and longitudes 35° 30′ E and 37° 45′ E, covering 10,160.7 km^2. About 30% of the area is occupied by cropland. Wheat is the major food crop (the average sowing area is 68.5% of the total cropped area) not only in the district, but in all of Turkey. The major sources of irrigation are rainfall, canals and groundwater.

Rainfall amounts vary spatially within the area covered by a given storm. Therefore, this area should be divided into hydrologically homogeneous regions in which rainfall amounts recorded at the rain gauges are assumed to be identical to obtain reliable results in hydrologic studies related to rainfall (Wisler and Brater, 1959; Okman, 1994). For this reason, the studied area was divided into 4 hydrologically homogeneous regions, west (W), central north (CN), central south (CS) and east (E), considering the mean, standard deviation and standard error of monthly rainfall recorded from the rain gauges and altitudes of the rain gauges in Tokat province. These 4 regions are separated from each other by Thiessen polygons. Average annual rainfall levels are 415.8, 479.6, 413.3, and 557.2 for W, CN, CS, and E, respectively (Figure 1) (Yürekli, 1999). Considering the similarity principle of rainfall amounts from rain gauges in a hydrologically homogeneous region, a rain gauge with the longest observation period was selected for each hydrologically homogeneous region. Days with rainfall equal to or less than a threshold of 6.4 mm and days with no rainfall were assumed to be drought periods in the 4 hydrologically homogeneous regions. To constitute the monthly time series, drought periods were obtained by counting days (for each month of the year) with rainfall equal to or less than the threshold value and days with no rainfall.

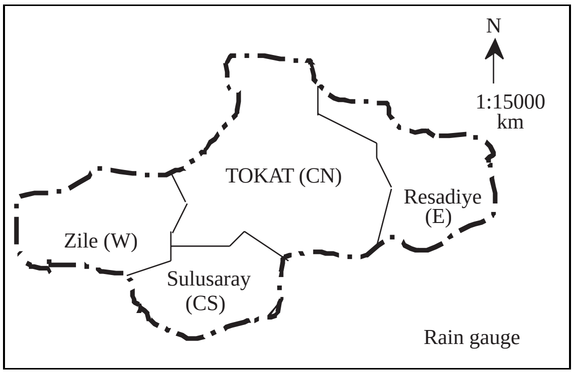

Figure 1. Hydrologically homogeneous regions of the study area.

Time Series Analysis for Drought Periods

In order to analyze time series for drought periods in the W, CN, CS and E hydrologically homogeneous regions, linear stochastic models known as either Box-Jenkins or ARIMA were used. A seasonal

ARIMA model is written as (Janacek and Swift, 1993; Sun and Koch, 2001)

$$\text{Ø}(B)\Phi F(B^s)(w_i - \mu) = Constant + \theta(B)\Theta(B^s)a_i \quad (1)$$

$$w_i = (1-B)^d(1-B^s)^D x_i \quad (2)$$

In Eq. (1), w_i should be taken as z_i if the series is stationary.

Box and Jenkins (1976) recommended that model development consist of 3 stages (identification, estimation and diagnostic check) when an ARIMA model is applied to a particular problem.

The identification stage is intended to determine the differencing required to produce stationarity and also the order of both the seasonal and nonseasonal autoregressive (AR) and moving average (MA) operators for a given series. By plotting original series (monthly series), seasonality and nonstationarity (stochastic trend in the mean and variance) can be revealed.

Many hydrologic time series processes may be stationary or nonstationary. Nonstationary time series can occur in many different ways. The occurrence of a stochastic trend in the mean and variance in a hydrologic time series can result from gradual natural or man-induced changes in the hydrologic environment producing the time series. In stochastic modeling studies in particular nonstationarity is a fundamental problem. Therefore a time series that has nonstationarity should be converted into a stationary time series. A nonstationary time series may be transformed into a stationary time series by using a linear difference equation. Therefore, nonstationarity is the first fundamental statistical property tested for in time series analysis (Huff and Changnon, 1973; Wei, 1990). The following nonparametric test (Spearman's Rho test) can be applied to decide whether a trend exists in the monthly data. The Spearman's Rho test recommended by Gibbons (1997) is given as

$$R_{sp} = 1 - \frac{6 * \sum D_i^2}{n * (n^2 - 1)} \quad (3)$$

$$D_i = K_{xi} - K_{yi} \quad (4)$$

$$t_{cal} = R_{sp} * \left[\frac{(n-2)}{(1-R_{sp}^2}\right]^{1/2} \quad (5)$$

To determine whether there is a trend, the calculated t_{cal} value in Eq. (5) should be compared to the t-critical value from the tables. If the t_{cal}value lies within a 5% significance interval, the data has no trend.

Autocorrelation function (ACF) and partial autocorrelation function (PACF) should be used to gather information about the seasonal and nonseasonal AR and MA operators for the monthly series (Hipel *et al.*, 1977). ACF measures the amount of linear dependence between observations in a time series. In general, for an MA (0,d,q) process, the autocorrelation coefficient (r_k) with the order of k cuts off and is not significantly different from zero after lag q. If r_k tails off and does not truncate, this suggests that an AR term is needed to model the time series. When the process is an MA (0,d,q)*(0,D,Q), r_k truncates and is not significantly different from zero after lag q+sQ. If r_k attenuates at lags that are multiples of s, this implies the presence of a seasonal AR component. For an AR (p,d,0) process, the PACF ($ø_{kk}$) with the order of k truncates and is not significantly different from zero after lag p. If $ø_{kk}$ tails off, this implies that an MA term is required. When the process is an AR (p,d,0)*(P,D,0), $ø_{kk}$ cuts off and is not significantly different from zero after lag p+sP. If $ø_{kk}$ damps out at lags that are multiples of s, this suggests the incorporation of a seasonal MA component into the model.

The estimation stage consists of using the data to estimate and to make inferences about values of the parameters conditional on the tentatively identified model. In an ARIMA model, the residuals (a_i) are assumed to be independent, homoscedastic and usually normally distributed. However, if the constant variance and normality assumptions are not true, they are often reasonably well satisfied when the observations are transformed by a Box-Cox transformation. The transformations can be expressed as either of the following equations (Wei, 1990):

$$z_{i=1}^n = \lambda^{-1}\left[(x_{i=1}^n + c)^\lambda - 1\right] \qquad \lambda \neq 0 \quad (6)$$

$$z_{i=1}^n = \ln(x_{i=1}^n + c) \qquad \lambda = 0 \quad (7)$$

Box and Jenkins (1976) stated that the model should be parsimonious. Therefore, they recommended the use of as few model parameters as possible so that the model fulfils all the diagnostic checks. Akaike (1974) suggested a mathematical formulation of the parsimony criterion of model building, the Akaike Information Criterion (AIC) for the purpose of selecting an optimal model fit to given data. The AIC mathematical formulation is defined as

$$AIC = -2L\left(\sum a_i^2\right) + 2K \tag{8}$$

where K = p + q +P+Q+ 1, and $L(\sum \hat{a}_t^2)$ = the log of the likelihood function of the Box-Jenkins ARIMA (p,d,q)*(P,D,Q) model. The log likelihood function, $L(\sum \hat{a}_t^2)$, is a monotonically decreasing function of the sum of squared residuals, $\sum \hat{a}_t^2$. The model that gives the minimum AIC is selected as a parsimonious model.

Shibata (1976) has shown that the AIC criterion tends to overestimate the order of autoregression. However, Akaike (1978, 1979) developed a Bayesian extension of the minimum AIC procedure, called BIC. Similar to Akaike's BIC, Schwarz (1978) suggested the following Bayesian criterion for model selection, which has been called the Schwarz Bayesian Criterion (SBC)

$$SBC = -2L\left(\sum a_i^2\right) + Kln(n) \tag{9}$$

The diagnostic check stage determines whether residualsare independent, homoscedastic and normally distributed.

The residual autocorrelation function (RACF) should be obtained to determine whether residuals are white noise. There are 2 useful applications related to RACF for the independence of residuals. The first is the correlogram drawn by plotting $r_k(a)$ against lag k. If some of the RACFs are significantly different from zero, this may mean that the present model is inadequate. The second is the Q(r) statistic suggested by Ljung and Box (1978). A test of this hypothesis can be done for the model adequacy by choosing a level of significance and then comparing the value of the calculated χ^2 to the actual χ^2 value from the table. If the calculated value is less than the actual χ^2 value, the present model is adequate on the basis of the available data. The Q(r) statistic is calculated by

$$Q(r) = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}r_k(a)^2 \tag{10}$$

$$r_k(a) = \sum_{i=k+1}^{n} a_i a_{i-k} / \sum_{i=1}^{n} a_i^2 \tag{11}$$

The following test described by Breusch and Pagan (1979) is very useful for determining whether a transformation of the data is needed. If there is a change in the variance (heteroscedasticity) of residuals, a transformation is necessary for the data. For the test, the residuals from the model fitted to the data are divided into 2 groups. Then the residual sum of squares (ESS_L, ESS_H) for these groups is obtained. The Breusch-Pagan test statistic (F_{cal}) is obtained from the equation below. This test statistic will be distributed F. If F_{cal} is smaller than the actual F value from the table, the residuals are assumed to be homoscedastic.

$$F_{cal} = \frac{ESS_H/(n_H - k_p)}{ESS_L/(n_L - k_p)} \approx F_{table}[(n_H - k_p), (n_L - k_p)] \tag{12}$$

There are many standard tests available to check whether the residuals are normally distributed. Chow *et al.* (1988) stated that if historical data are normally distributed, the graph of the cumulative distribution for the data should appear as a straight line when plotted on normal probability paper.

Haan (1977) stated that the other way to check the normality of residuals is the Kolmogorov-Smirnov method. This is a non-parametric test of the fit of data to a theoretical distribution using the maximum absolute deviation (D_{cal}) between the 2 functions of cumulative distribution. The maximum absolute deviation is (Haan, 1977)

$$D_{cal} = \max|Fn(x) - Fa(x)| \tag{13}$$

The value of the D_{cal} statistic is compared with the critical value D_{Tab} (n, α) obtained from Haan (1977). If D_{cal} is greater than the critical value D_{Tab} (n, α), the null hypothesis related to normality is rejected for the chosen level of significance.

Results and Discussions

All the information from rain gauge stations (Figure 1) taken into consideration in order to simulate drought periods in the W, CN, CS and E hydrologically homogeneous regions is given in Table 1. The non-parametric test (Spearman's Rho test) at a 5% significance level was applied to monthly series comprising days with rainfall equal to or less than the threshold value of 6.4 mm and days with no rainfall. The Spearman's Rho test results are given in Table 2. The t_{cal} values of the CN, W and CS hydrologically homogeneous regions for the 6.4 mm threshold and of the CS hydrologically homogeneous region only for days with no rainfall were not between the actual t values (±1.96) from the table at a 5% significance level. This suggests that there are linear trends in these data sequences. The plots of the ACFs drawn for the data sequences were examined in order to identify the form of the ARIMA model. Visual inspections show that the plot of original series and the ACF graph for the 6.4 mm threshold of the E hydrologically homogeneous region follow an attenuating sine wave pattern that reflects the random periodicity of the data and possibly indicates the need for non seasonal and/or seasonal AR terms in the model. For these series, the cyclic seasonal component and linear trend were removed by taking the non seasonal and seasonal differencing operator as one (1).

The ACFs and PACFs were estimated for the monthly data of each hydrologically homogeneous region. All the ACFs were significantly different from zero at lag 60 ($5s < n = 240/4$). Additional to this, the Ljung-Box Q statistics were estimated for lag 60. The Q(r) statistics of the W, CN, CS and E hydrologically homogeneous regions at lag 60 for the 6.4 mm threshold are 240.0, 515.8, 271.1 and 422.2, respectively. The Q(r) statistics of the hydrologically homogeneous regions at lag 60 for days with no rainfall are 1204.0, 1108.9, 1145.4 and 717.6, respectively. These results are greater than 77.9 (the actual χ^2 value). Therefore, they emphasize that the ACFs obtained from the monthly data sequences are different from zero. In other words, there was a linear dependence between the selected drought periods. However, the ACFs did not cut off but rather damped out. This may suggest the presence of AR

Table 1. Hydrologically homogeneous regions for the study area.

Hydrologic regions	Rain gauges	Average rainfall (mm)	Area (km^2)
West	Zile, Turhal, Boztepe, Zile, Reşadiye	34.7	2010
Central North	Tokat, Erbaa, Almus, Almus Dam, Doğanyurt, Pazar, Dökmetepe, Hacıpazarı, Niksar	47.0	4650
Central South	Artova, Çamlıbel, Sulusaray, Ekinli	36.3	1539
East	Çamiii, Reşadiye, Bereketli	47.7	1961

Table 2. The ARIMA models selected for hydrologically homogeneous regions.

Hydrologic region	ARIMA model	Model Statistic					
		Trend (t_t)	SBC	Probability of Q(r) > 0.05	K-S Test $D_{cal} < D^*_{Tab}$	F_{cal} >0.05	Const.
$W_{6.4}$	(0,1,1)(1,0,1)	2.67	-1289.6	0.714	0.072	0.661	—
W_{NR}	(0,0,1)(1,0,1)	0.04	1346.5	0.963	0.051	0.853	20.95
$CN_{6.4}$	(2,1,0)(1,0,1)	2.69	993.1	0.061	0.076	0.792	—
CN_{NR}	(2,0,1)(1,0,1)	0.74	1358.1	0.525	0.062	0.922	20.98
$CS_{6.4}$	(0,1,2)(1,0,1)	2.56	996.6	0.827	0.081	0.709	—
CS_{NR}	(2,1,0)(1,0,1)	3.79	1415.5	0.089	0.038	0.732	—
$E_{6.4}$	(1,0,1)(1,1,1)	1.76	874.7	0.320	0.071	0.974	—
E_{NR}	(1,0,2)(1,0,1)	0.95	1317.3	0.158	0.077	0.894	—

*0.088, D_{Tab}critical value from the table
SBC, Schwarz Bayesian Criterion
Const, constant in ARIMA model

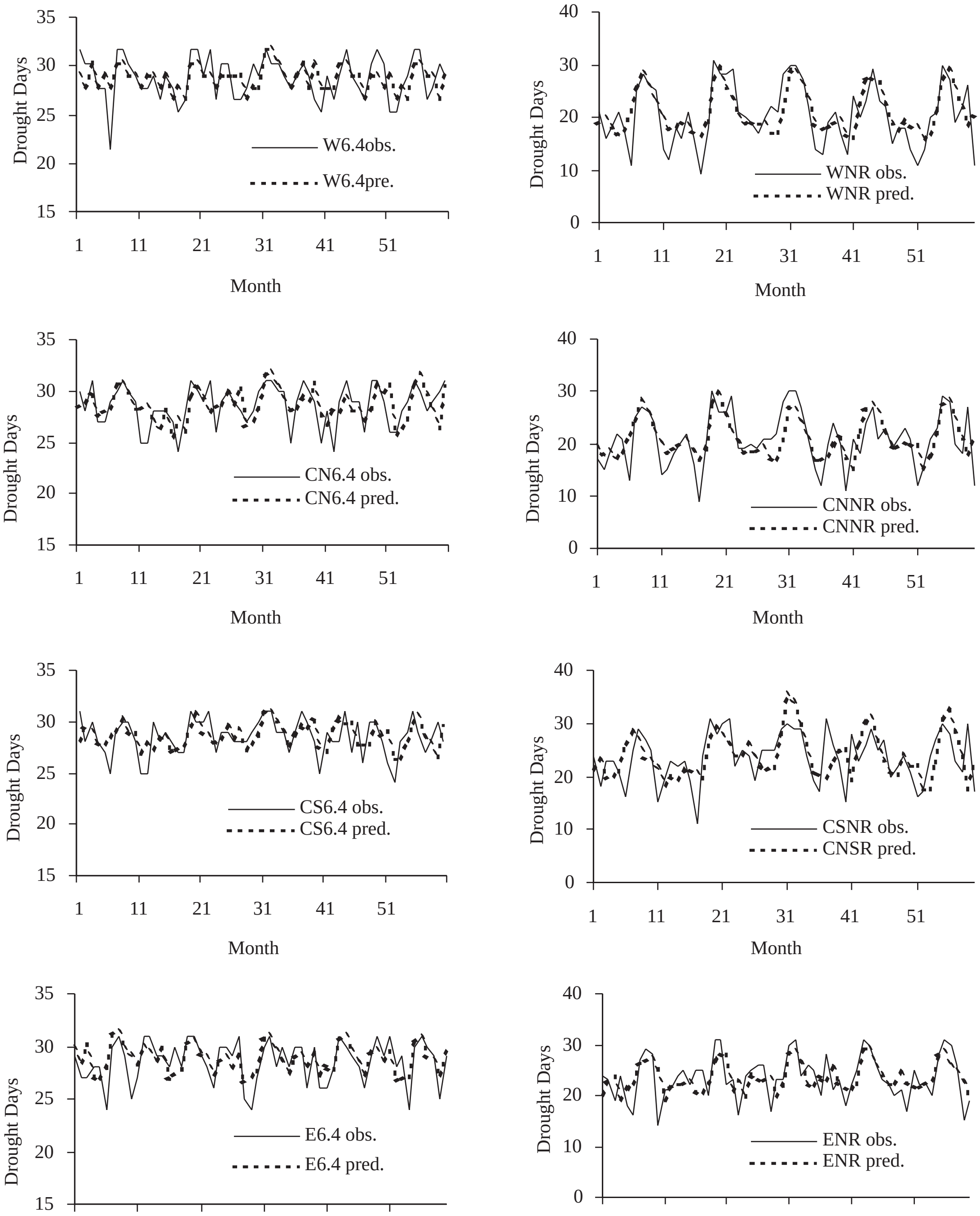

Figure 2. Comparison of the observed data to the predicted data for 5 year.

terms. The PACFs possess significant values at some lags but rather tail off. This may imply the presence of MA terms. The ACFs have significant values at lags that are multiples of 12. This may stress that seasonal AR terms are required but that these values attenuate. There are peaks on the graphs of the PACFs at lags that are multiples of 12 that may suggest seasonal MA terms, but these peaks damp out.

Alternative ARIMA models were estimated by considering the ACF and PACF graphs from the monthly data obtained for the hydrologically homogeneous regions. The SBC was taken into account for obtaining a parsimonious model among these alternatives. The model with the lowest SBC was assumed to be parsimonious. In addition, model parameters were analyzed at a 5% significance level by using the t-test to select the best model fit to the data. If there was any parameter significant at a level greater than 5% , it was eliminated.

Diagnostic checks were applied in order to determine whether the residuals of the selected models from the ACF and PACF graphs were independent, homoscedastic and normally distributed. The models that had the minimum SBC among the selected models fulfilled all the diagnostic checks and these were selected as the best model for days with rainfall equal to or less than the 6.4 mm threshold and days with no rainfall for the hydrologically homogeneous regions. The selected best models are given in Table 2. The critical assumption of independence for the RACFs of the residuals was done by using the χ^2 distributed Ljung-Box Q statistic. In addition, the probabilities of the Q statistics calculated for the best models are given in Table 2. In this table, test results from the Kolmogorov-Smirnov method for the normality and those from Breusch-Pagan for the homoscedascity of the residuals are also given. Table 2 shows that all the diagnostic checks for the residuals are fulfilled.

The value (V) of the parameters, associated standard errors (SEV), t-ratios and probabilities (<5%) for the standard errors are listed in Table 3. The SEVs calculated for the model parameters were rather small compared to the parameter values. Furthermore, even at the 1% significance level, all of the parameters are significant and these parameters should be included in the models (Table 3).

Since the residuals from the models are normally distributed and homoscedastic, a Box-Cox transformation (Eqs. (3) and (4)) to the monthly data sequences was not necessary except for the 6.4 mm threshold for the W hydrologically homogeneous region. In Eq. (4), c is substituted by 100 for this series. Granger and Newbold (1976) pointed out that the transformations can change the type of the model being estimated. Therefore, it is better to estimate a model from non-transformed data.

Comparison of means and standard deviation for observed data and predicted data from the model are given in Table 4. To detect whether there was a significant difference between the mean and standard deviation values from the observed and predicted data, a Z-test for the means and F-test for the standard deviation were applied (Haan, 1977; Devore and Peck, 1993). Since Z_{cal} values related to means were between Z-critical table values ($\pm$1.96 for two-tailed at a 5% significance level), the data support the claim that there is no significant difference between the mean values of observed and predicted data. Similarly, the F_{cal} values of standard deviation were smaller than the F-critical table values at a 5% significance level. Thus, the results show that predicted data preserve the basic statistical properties of the observed series. In addition, Figure 2 shows the relationships between the observed data for 5-years and the predicted data for the same years from the best model selected for days with rainfall equal to or less than the 6.4 mm threshold and days with no rainfall in each hydrologically homogeneous region. For each hydrologically homogeneous region and selected threshold values, the predicted data follow the observed data very closely.

Conclusion

Based on the analysis to simulate drought durations by the ARIMA model for hydrologically homogeneous regions shown above, the following conclusions were drawn:

ARIMA model applications to the W, CN, CS and E hydrologically homogeneous regions showed that predicted data preserved the basic statistical properties of the observed series.

The ARIMA model equation for each hydrologically homogeneous region given below can be used successfully for the simulation of drought durations.

Table 3. Statistical analysis for the model parameters.

Hydrologic regions	Model parameters	Variables in the model: Value of parameters	Standard error	t-ratio	Probability < 0.05
$W_{6.4}$	θ_1	0.832	0.052	16.12	0.000
	Φ_1	0.982	0.023	43.23	0.000
	Θ_1	0.868	0.078	11.14	0.000
W_{NR}	θ_1	-0.193	0.062	-3.13	0.002
	Φ_1	0.998	0.005	186.80	0.000
	Θ_1	0.942	0.094	9.98	0.000
	Constant	20.95	1.318	15.89	0.000
$CN_{6.4}$	$\emptyset_1$	-0.509	0.061	-8.30	0.000
	$\emptyset_2$	-0.277	0.062	-4.49	0.000
	Φ_1	0.985	0.018	55.54	0.000
	Θ_1	0.857	0.072	11.90	0.000
CN_{NR}	$\emptyset_1$	0.971	0.149	6.50	0.000
	$\emptyset_2$	-0.282	0.063	-4.44	0.000
	θ_1	0.771	0.139	5.56	0.000
	Φ_1	0.993	0.013	75.75	0.000
	Θ_1	0.914	0.084	10.88	0.000
	Constant	20.98	0.605	34.67	0.000
$CS_{6.4}$	θ_1	0.592	0.066	9.03	0.000
	θ_2	0.253	0.065	3.89	0.000
	Φ_1	0.986	0.021	47.89	0.000
	Θ_1	0.897	0.074	12.13	0.000
CS_{NR}	$\emptyset_1$	-0.582	0.061	-9.61	0.000
	$\emptyset_2$	-0.306	0.061	-4.97	0.000
	Φ_1	0.993	0.009	111.09	0.000
	Θ_1	0.890	0.066	13.55	0.000
$E_{6.4}$	$\emptyset_1$	0.795	0.135	5.89	0.000
	θ_1	0.887	0.108	8.23	0.000
	Φ_1	-0.166	0.075	-2.21	0.028
	Θ_1	0.866	0.062	13.86	0.000
E_{NR}	$\emptyset_1$	0.998	0.001	868.63	0.000
	θ_1	0.731	0.063	11.54	0.000
	θ_2	0.237	0.062	3.80	0.000
	Φ_1	0.923	0.040	22.98	0.000
	Θ_1	0.681	0.092	7.43	0.000

$W_{6.4}$, $(1 - 0.982B^{12})\,[(1 - B)z_i] = (1 - 0.832B)\,(1 - 0.868B^{12})a_i$
W_{NR}, $(1 - 0.998B^{12})x_i = 20.95 + (1 + 0.193B)\,(1 - 0.942B^{12})a_i$
$CN_{6.4}$, $(1 + 0.509B + 0.277B^2)\,(1 - 0.985B^{12})\,[(1 - B)x_i] = (1 - 0.857B^{12})a_i$
CN_{NR}, $(1 - 0.971B + 0.282B^2)\,(1 - 0.993B^{12})x_i = 20.98 +$
$(1 - 0.771B)\,(1 - 0.914B^{12})a_i$
$CS_{6.4}$, $(1 -0.986B^{12})\,[(1 - B)x_i] = (1 - 0.592B - 0.253B^2)\,(1 - 0.897B^{12})a_i$
CS_{NR}, $(1 + 0.582B + 0.306B^2)\,(1 - 0.993B^{12})\,[(1 - B)x_i] = (1 - 0.890B^{12})a_i$
$E_{6.4}$, $(1 - 0.795B)\,(1 + 0.166B^{12})\,[(1 - B^{12})x_i] = (1 - 0.887B)\,(1 - 0.866B^{12})a_i$
E_{NR}, $(1 - 0.998B)\,(1 - 0.923B^{12})x_i = (1 - 0.731\,B - 0.237B^2)\,(1 - 0.681B^{12})a_i$

Table 4. Comparison of statistic properties of the observed and predicted data.

Hydrologic regions	M^1_{obs}	M^2_{pre}	Decision $\mid z_{cal} \mid < 1.96$	SD^3_{obs}	SD^4_{pre}	Decision $F_{cal} < 1.24$
$W_{6.4}$	28.7	28.7	0.00	1.94	1.31	0.68
W_{NR}	21.0	21.1	-0.24	5.40	3.75	0.69
$CN_{6.4}$	28.6	28.6	0.00	1.89	1.39	0.74
CN_{NR}	21.0	21.1	-0.24	5.27	3.50	0.66
$CS_{6.4}$	28.5	28.6	-0.68	1.97	1.14	0.58
CS_{NR}	22.4	22.4	0.00	5.21	4.42	0.85
$E_{6.4}$	28.7	28.9	-1.30	1.93	1.39	0.72
E_{NR}	23.4	23.2	1.11	4.40	3.51	0.80

[1]mean of the observations
[2]mean of the predictions from the model
[3]standard deviation of the observations
[4]standard deviation of the predictions from the model

Nomenclature

a_i	white noise time series value at time i
B	backward shift operator
c	constant for Box-Cox transformation
d	order of the nonseasonal differencing operator
D	order of the seasonal differencing operator
ESS_L	the residual sum of square for the low group
ESS_H	the residual sum of square for the high group
Fa (x)	the specified theoretical cumulative distribution function
Fn (x)	cumulative density function based on n measurements
k_p	degree of freedom
K_{xi}	rank of i^{th} observation in the historical data
K_{yi}	rank in the historical data of i^{th} observation in the ascended data
n	the number of observation
n_L	the number of residuals in the low group
n_H	the number of residuals in the high group
Q(r)	Ljung-Box statistic at lag m
r_k(a)	ACF of a_i at lag k
R_{sp}	rank order correlation coefficient
s	seasonal lengh
x_i	discrete time series value at time i
w_i	stationary series formed by differencing the x_i
z_i	transformation of $x_i$series

Greek Symbols

λ	exponent for Box-Cox transformation
μ	mean level of the w_i series (if D + d > 0 often $\mu \approx 0$)
Ø(B)	nonseasonal AR operator of order p
θ (B)	nonseasonal MA operator of order q
Φ (B)	seasonal AR parameter of order P
Θ (B)	seasonal MA parameter of order Q

References

Agnew, C., "Spatial Aspects of Drought in the Sahel", Journal of Arid Environments, 18, 279-293, 1990.

Agnew, C. and Warren, A., "A Framework for Tackling Drought and Land Degradation", Journal of Arid Environments, 33, 309-320, 1996.

Akaike, H., "A Look at the Statistical Model Identification", IEEE, Transactions on Automatic Control, AC-19, 716-723, 1974.

Akaike, H., "A Bayesian Analysis of the Minimum AIC Procedure", Ann. Inst. Statist. Math., 30A, 9-14, 1978.

Akaike, H., "A Bayesian Extension of the Minimum AIC Procedure of Autoregressive Model Fitting", Biometrika, 66, 237-242, 1979.

Box, G.E.P. and Jenkins, G.M., Time Series Analysis Forecasting and Control, Holden-Day, San Francisco, 1976.

Breusch, T. and Pagan, A., "A Simple Test of Heteroscedasticity and Random Coefficient Variation", Econometrica, 47, 1287-1294, 1979.

Chow, V.T., Maidment, D.R. and Mays, L.W., Applied Hydrology, McGraw-Hill Book Company, New York, 1988.

Devore, J. and Peck, R., Statistics: The Exploration and Analysis of Data, Second Edition, Duxbury Press, California, 1993.

Gibbons, J.D., Nonparametric Methods for Quantitative Analysis, American Sciences Press, Inc., Ohio - USA, 1997.

Granger, C.W.J. and Newbold, P., "Forecasting Transformed Series", Journal of the Royal Statistical Society, Series B, 38, 189-203, 1976.

Haan, C.T., Statistical Methods in Hydrology, Iowa State Press, Iowa, 1977.

Hershfield, D.M., Brakensiek, D.L. and Comer, G.H., Some Measures of Agricultural Drought, Floods and Droughts, Water Resources Publications, Fort Collins Colorado, 1973.

Hipel, K.W., McLeod, A.I. and Lennox, W.C., "Advances in Box-Jenkins Modeling 1. Model Construction", Water Resources Research, 13, 567-575, 1977.

Howe, O.W. and Rhoades, H.F., "Irrigation Practice for Corn Production in Relation to Storage of Plant Development", Soil Science Society of Proceedings, 19, 25, Washington D.C., 1955.

Huff, F.A. and Changnon, S.A., "Precipitation Modification by Major Urban Areas", Bulletin American Meteorological Society, 54, 1220-1232, 1973.

Janacek, G. and Swift, L., Time Series Forecasting, Simulation, Application, Ellis Horwood, New York, 1993.

Le Houerou, H.N., "Climate Change, Drought and Desertification", Journal of Arid Environments, 34, 133-185, 1996.

Ljung, G.M. and Box, G.E.P., "On a Measure of Lack of Fit in Time Series Models", Biometrika, 65, 297-303, 1978.

Okman, C., The Recurrence Probability of Agricultural Drought Spells in Ankara Province, University of Ankara, Faculty of Agriculture, Publication No, 777, Ankara, 1981 (In Turkish).

Okman, C., Hydrology, University of Ankara Faculty of Agriculture, Publication No, 1388, Ankara, 359, 1994 (In Turkish).

Oliver, H., Irrigation and Climate, Edward Arnold Limited, London, 1961.

Palmer, W.C., Meteorological Drought, Research Paper No. 45. United States Department of Commerce, Weather Bureau, 1965.

Panu, U.S. and Sharma, T.C., "Challenges in Drought Research: Some Perspectives and Future Directions", Hydrological Sciences Journal, 47, 19-30, 2002.

Richer, C.B., Jacobson, P. and Hall, C., Agricultural Engineers Handbook, McGraw - Hill Book Company, New York, 1961.

Schwartz, G., "Estimating the Dimension of a Model" Ann. Statist., 6, 461-464, 1978.

Shibata, R., "Selection of the Order of an Autoregressive Model by Akaike's Information Criterion", Biometrika, 63, 117-126, 1976.

Sun, H., and Koch, M., "Case Study: Analysis and Forecasting of Salinity in Apalachicola Bay, Florida, Using Box-Jenkins Arima Models", Journal of Hydraulic Engineering, 127, 718-727, 2001.

Waring, E.J., Local and Regional Effects of Drought in Australia, Australian UNESCO Seminar, Australian Government Publishing Service, Canberra, 1976.

Wei, W.W.S., Time Series Analysis, Addison-Wesley Publishing Company Inc. New York, 1990.

Wilhite, D.A. and Glantz, M.H., "Understanding the Drought Phenomenon: The Role of Definitions", Water International, 10, 111-120, 1985.

Wisler, C.O. and Brater, E.F., Hydrology, John - Wiley and Sons Inc., New York, 1959.

Yürekli, K., Hydrologic Characteristics of Amasya and Tokat Provinces, PhD. Thesis, University of Ankara, Graduate School of Natural and Applied Sciences, Department of Farm Structures and Irrigation, Ankara, 1999 (Unpublished) (In Turkish).