Pinnacle Environmental & Earth Science

http:/www.pjpub.org

Research Article

Using Time Series Analysis to support the Water Resources Management in the Upper Basin of the Suquía River

Reyna, Teresa[1], Guillén Nicolás[2], Lábaque María[3], Álvarez Javier[4], Eder Matías[5] & Funes Federico[6]

[1,2,3,4,5,6] Facultad de Ciencias Exactas,
Físicas y Naturales. U.N.C,
Av. Vélez Sarsfield 1600,
Ciudad Universitaria.
Argentina.

Accepted 4 March, 2015

ABSTRACT

A time series analysis is a valuable tool to get information about analyzed data structures and their components, which provides a good basis for successful future predictions. This information is useful to make decisions for the best management of this resource. The region where the province of Cordoba is located is subject to strong variations in their hydrological cycles. The strong population growth experienced in the recent years in the city of Cordoba and its area of influence (Area of Greater Córdoba) with more than 1.4000.000 inhabitants has resulted in conflicts in the uses of water. During the last decades this area has experienced prolonged droughts and floods. The Suquía River is the most important source of water supply to the City of Córdoba in Argentina. This study deals with the modeling and forecasting of rainfall time series in the area of upper basin of Suquía River. The series analyzed were San Antonio, El Condor, El Cajon and Barrio El Canal Station. The time series models that were proposed for the analysis of each record were ARMA models. The results show that the time series analysis is a tool that offers valuable information on the characteristics pluviographs.

Keywords: basin, rainfall, Suquía, time series, series of rainfall.

Introduction

Access to water in the world is becoming one of the most significant problems for wildlife conservation and cities development. A proper planning in the use and management of water is considered essential in the developed world, which has become conscious about the environment and the conservation of nonrenewable resources.

The region where the province of Cordoba is located is subject to strong variations in their hydrological cycles. Within this context, the strong population growth experienced in the recent years in the city of Cordoba and its area of influence (Area of Greater Córdoba) has resulted in conflicts in the uses of this resource. In the last decades situations of prolonged droughts and floods have been seen to be accompanied by changes in the land use, and new urban developments in watersheds (particularly Suquía). All these have been competing for water. Increasing pollution in the watersheds, siltation and eutrophication in the reservoirs have only added to the complex situation, where the demands exceed the available water in the system.

Fig. (1): Location of Córdoba within Argentina.

Corresponding Author: Teresa Reyna[1]
Facultad de Ciencias Exactas, Físicas y Naturales. U.N.C, Av. Vélez Sarsfield 1600, Ciudad Universitaria. Argentina.
Email address: teresamaria.reyna@gmail.com

Rainfall is a natural phenomenon resulting from atmospheric and oceanic circulation (local convection, frontal or orographic patterns). In the Suquía River Basin in Argentina, the prediction of rainfall is important for proper mitigation and management of floods, droughts, environmental flows, water demand by different sectors, maintaining reservoir levels, and planning and preparing for disasters. For this purpose, in our study, a time series analysis was applied as a valuable tool to get information about analyzed data structures and their components, and provides a good basis for their successful future predictions.

This article deals with the modeling and forecasting of rainfall time series for three rain-gauge stations in the area of the San Antonio River. This river is one of the most important tributaries of Suquía River.

Methodology

This section is dedicated to giving a brief overview of the methodology used in data analyses.

Many attempts have been made in the recent past to model and forecast rainfall using various techniques, with the use of time series techniques proving to be the most common. Time series analysis techniques usually involve some method of filtering out noise in order to make the pattern more silent.

Stochastic models are mainly based on the optimization of information in the time series of| given variable; reproduce the historical statistics of the time series (Box and Jenkins, 1970). The key concept in this type of model is the dependence between dates; your application has some advantages, for example has a lot of flexibility to work in different scales.

A probabilistic model that describes the structure of a sequence of observations is a stochastic process. A time series of N successive observations $Z' = (Z_1, Z_2, ..., Z_N)$ is taken as an example or embodiment taken from an infinite population of such samples, which could have been generated by the process. A major objective of the statistical research is to infer the properties of the population from the properties of the sample. To do this it needs the skills to describe stochastic processes and time series, and also have classes of stochastic models able to describe situations that occur in practice.

The present work exclusively deals with a time series forecasting model, in particular, the autoregressive moving average (ARMA) model. These models were described by Box and Jenkins (1970). An autoregressive model of order p is conventionally classified as AR(p) and a moving average model with q terms is known as MA(q). A combined model that contains p autoregressive terms and q moving average terms is called ARMA(p,q) (Gujarati, 1995).

The study was performed using the software statgraphics (http://www.statgraphics.com) STATGRAPHICS Plus Version 5.1

Autoregressive Models

A stochastic model that can be extremely useful in the representation of time series is called autoregressive model. In this model the current process value is expressed as a finite linear sum of previous values of the process, plus a shock at random. Notice to process values sampled at equally spaced time intervals *t, t-1*, ..., as Z_t, Z_{t-1}, Z_{t-2}, ... Also note deviations from the mean (μ) (Salas *et al*., 1985) with

$$\tilde{Z}_t = Z_t - \mu$$

$$\tilde{Z}_t = \phi_1.\tilde{Z}_{t-1} + \phi_2.\tilde{Z}_{t-2} + \ldots + \phi_p.\tilde{Z}_{t-p} + a_t$$

is an autoregressive model of order p, AR (p). The reason for this name is that a linear model of the form.

$$\tilde{Z} = \phi_1.\tilde{X}_1 + \phi_2.\tilde{X}_2 + \ldots + \phi_p.\tilde{X}_p + a$$

linking a 'dependent' variable Z with a set of "independent" *X1, X2, X3, ..., Xp* plus an error term, is called regression model variables, and it says that Z is "returned" on *X1, X2, ..., Xp*. In the first equation the variable z is returned on previous values of itself; Therefore the model is self - regressive. This model contains unknown parameters $p + 2$: μ, $Ø_1$, $Ø_2$, ..., $Øp$, σ_a^2, which, in practice, must be estimated from the data. The additional parameter, σ_a^2, is the variance of white noise at.

Models Moving Averages

The autoregressive model expresses the deviation $\tilde{z}_t$ of the process as the finite weighted sum of the deviations of the process, plus a random shock *a*. Equivalently expressed $\tilde{z}_t$ as an infinite sum of the weighted a. Other models of great importance in the representation of time series is called finite process mean or moving averages. Here we make a linearly dependent on a finite number of previous shocks. Then the process:

$$\tilde{Z}_t = a_t - \theta_1.a_{t-1} - \theta_2.a_{t-2} - \ldots - \theta_q.a_{t-q}$$

is called a moving average process of order *q, MA (q)*. This model contains unknown parameters q + 2, which, in practice they will be estimated from the data.

Study Area and Data Collection

The Saws of Cordoba Province are located between 30° and 34° south latitude and 63° and 65 degrees west longitude, consisting of three strands north-south orientation: Oriental, Central and West.

Suquía River or Primero River is one of the most important rivers in the Province of Córdoba. It has its source where several rivers and creeks join in the area of the Pampean Hills to the West of the Province of Córdoba, especially in Punilla Valley. San Francisco River or Río Grande de Punilla, which flows from North to South, joins, by the city of Cosquín, the Yuspe River, which then joins Las Mojarras creek and Los Chorrillos Creek, both having their sources in Los Gigantes.

Cite this Article: Reyna Teresa, Guillén Nicolás, Lábaque María, Álvarez Javier, Eder Matías & Funes Federico "Using Time Series Analysis to support the Water Resources Management in the Upper Basin of the Suquía River" Pinnacle Environmental & Earth Science, Vol. 2 (1) 2015, Article ID pees_179, 531-539, 2015.

Finally, in the South area of Punilla Valley, they join San Antonio River, whose origin is in Altas Cumbres or High Hills. All of these rivers flow into San Roque Dam, forming a reservoir with the same name, as the dam wall placed at the west and high side of Quebrada (Ravine) de Bamba, thus originating Suquía River. When leaving the reservoir, it continues with west-east direction and goes through an area where most of the City of Córdoba lies, and almost in the center of the city, it receives on the right or southern bank of the river a creek called La Cañada.

In this paper the rainfall in the area of the upper basin of the Suquía River was studied (San Antonio Basin). The series were provided by the SSRH of the Nation (http://www.hidricosargentina.gov.ar/sistema_sistema).

Three series were considered: Station San Antonio - El Condor, Station El Cajon and Station Barrio Canal. The localization of them is presented in the Figure 2.

Station: San Antonio - El Condor: The station San Antonio - El Condor, located at coordinates 31° 34'50'' South Latitude and 64° West Longitude 47'20''

This series is daily rainfall time series which starts on 02/13/1992 and ends on 30/12/2013. This series has a total of 7719 measured data and has a shortfall of 273 values (3.42% of the values of the series). Longer intervals without consecutively measured values correspond to the following periods 04/01/1999 - 04/30/1999 and 04/01/2006 - 04/30/2006 both with a total of 30 values.

Station El Cajon: The station El Cajon, located at coordinates 31° 34'50'' South Latitude and 64° West Longitude 47'20''

The time series starts on 12/01/ 1992 and ends on 08/30/2010. This series has a total of 5988 measured data and has a shortfall of 314 values (4.98% of the values of the series). The longer interval without consecutively measured values corresponds to the period 04/01/1996 - 05/31/1996 with a total of 61 values.

Station Barrio El Canal - Suquia: The rainfall station Barrio El Canal - Suquía, located at coordinates 31.45° South Latitude and Longitude West 64.49º. The time series starts on 01/01/1992 and ends 08/25/2010. This series has a total of 6599 measured data and no missing values.

Different predictive models were set: ARMA (1,0), ARMA (2,0), ARMA (0,1), ARMA (0,2) and ARMA (1,1) and the results are shown.

Table 1: Total annual precipitation (PTA). * Incomplete Registration

	Hydrological PTA [mm]		
Year	**El Cajón**	**El Canal**	**SA El Cóndor**
1992-93		760	736
1993-94	769	477	910
1994-95	**498**	705	824
1995-96	665	540	1051
1996-97	631	872	784
1997-98	729	711	910
1998-99	706	**439**	**648**
1999-00	909	921	914
2000-01	747	664	878
2001-02	672	699	1154
2002-03	946	716	851
2003-04	914	642	790
2004-05	868	542	932
2005-06	791	478	843
2006-07	863	450	773
2007-08	934	542	830
2008-09	872	700	854
2009-10	455*	238	817
2010-11			734
2011-12			892
2012-13			819
Average	**762.88**	**616.44**	**854.48**
Standard Deviation	**146.97614**	**168.03824**	**108.96145**

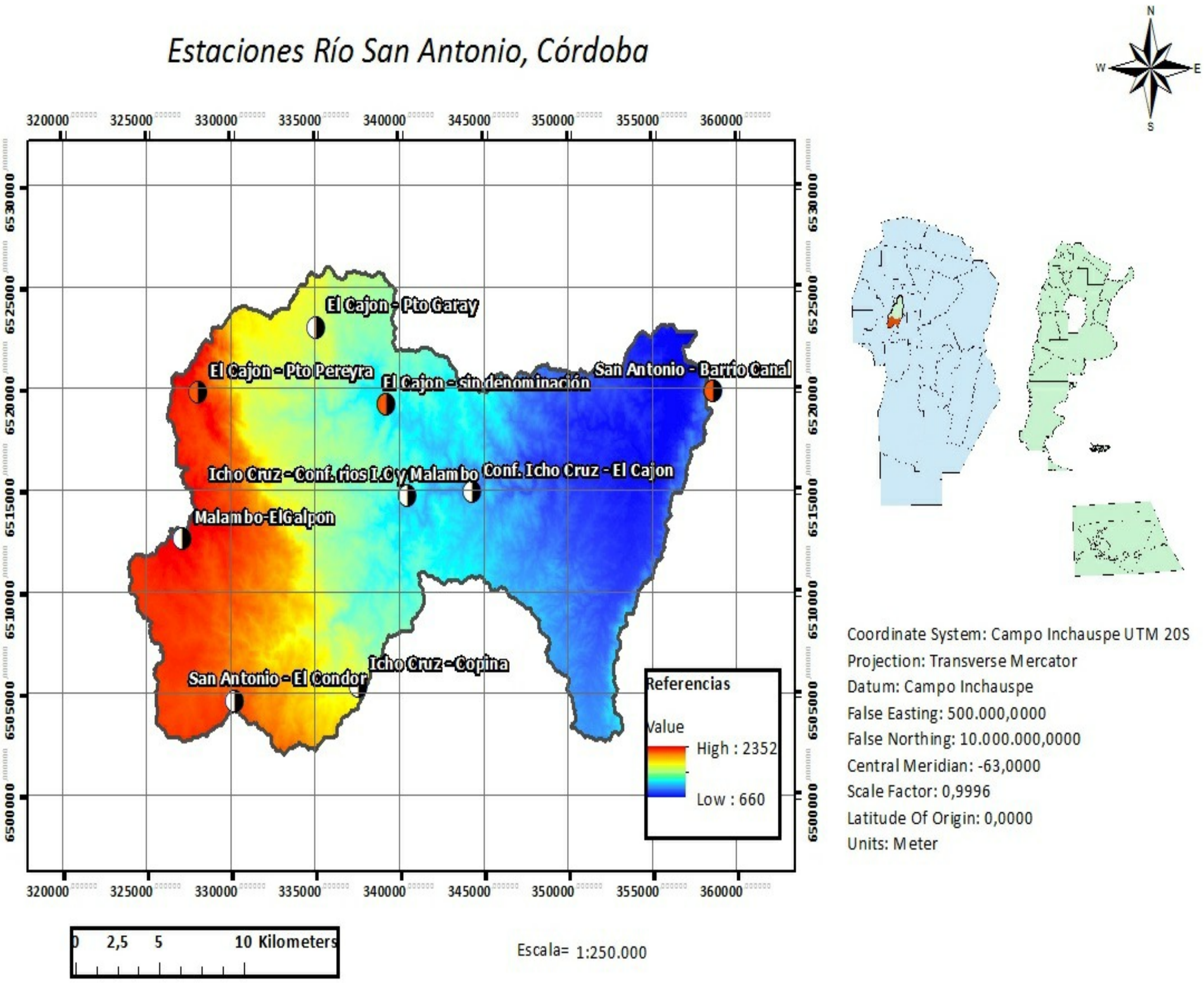

Fig. (2): Location of Station San Antonio River, Córdoba.

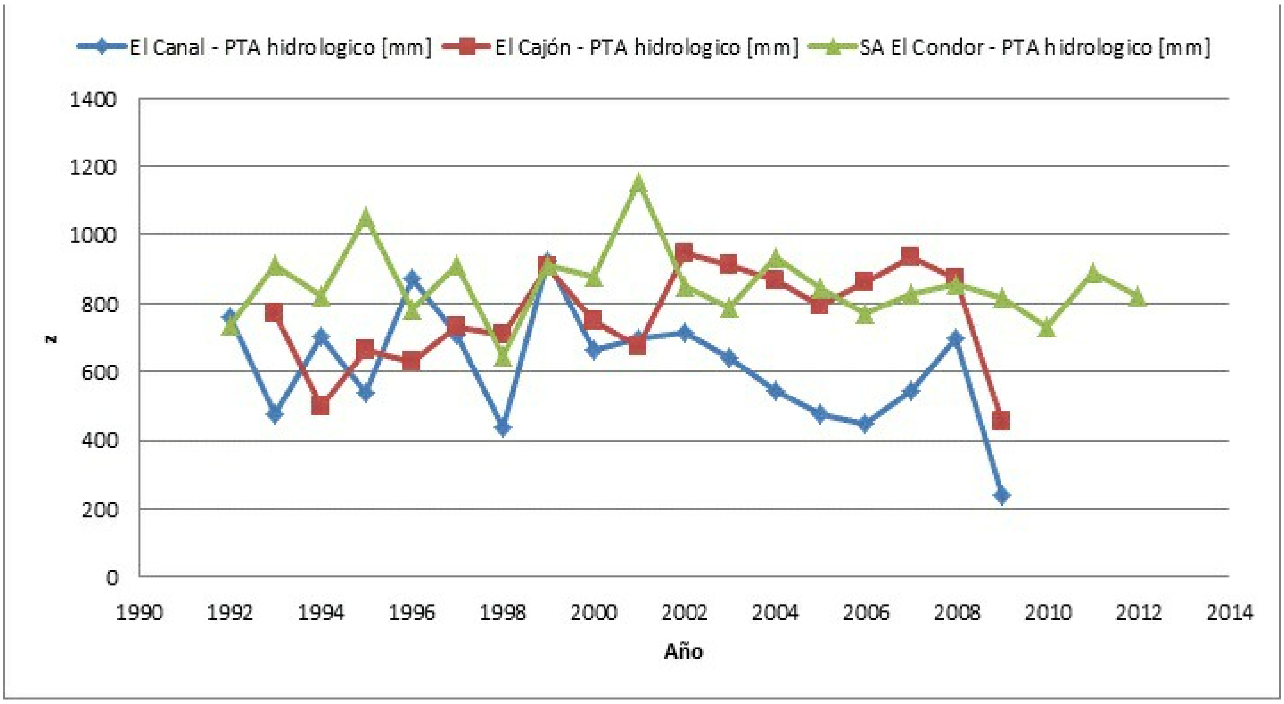

Fig. (3): Total Annual rainfall in mm

Table 2: Monthly rainfall Station Barrio El Canal - Suquía, El Cajón and SA El Condor.

	Hydrological PTA [mm]		
Month	**El Cajón**	**El Canal**	**SA El Cóndor**
July	6.5	6.27	7.41
August	5	5.27	8.86
September	24.06	22.77	28.59
October	58.06	53.61	62.91
November	109.76	79.83	140.86
December	145.56	102.94	151.68
January	136.28	110.83	144.32
February	102.33	82.61	118.32
March	101.17	82.33	120.59
April	37.61	43.61	46.27
May	19.33	15.44	22.27
June	20.94	3.33	5.09

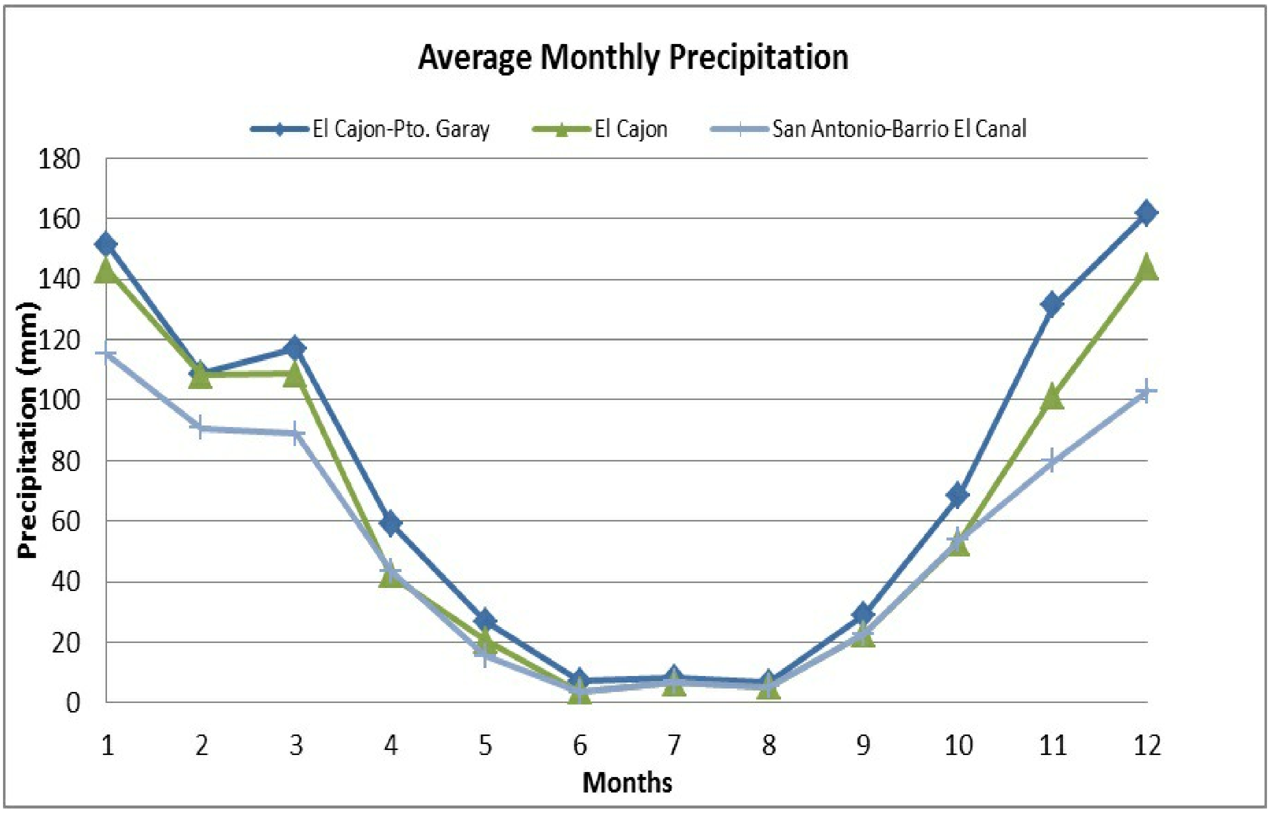

Fig. (4): Rainfall Stations Monthly Measurements El Cajon, El Canal and El Condor.

In this step, the model that seems to represent the behavior of the series is searched through autocorrelation and partial autocorrelation functions (ACF and PACF) for further investigation and parameter estimation (McLeod *et al*, 1977).

The models considered in each series annual were (A) ARIMA(1,0,0), (B) ARIMA(2,0,0), (C) ARIMA(0,0,1), (D) ARIMA(0,0,2), (E) ARIMA(1,0,1).

The study was performed using the software statgraphics (http://www.statgraphics.com).

The results are presented in the table 3.

Table 3: Results of different forecasting models.

EL CANAL STATION										
Modelo	*RMSE*	*MAE*	*MAPE*	*ME*	*MPE*	*RUNS*	*RUNM*	*AUTO*	*MEDIA*	*VAR*
(A)	1.02045	0.76147		-0.00336		OK	OK	OK	*	OK
(B)	0.98387	0.68633		-0.04653		OK	OK	OK	*	OK
(C)	1.02531	0.76452		-0.00358		OK	OK	OK	*	OK
(D)	1.02596	0.70947		-0.01408		OK	OK	OK	*	OK
(E)	1.00684	0.73559		-0.07589		OK	OK	OK	OK	OK
EL CAJÓN STATION										
Modelo	*RMSE*	*MAE*	*MAPE*	*ME*	*MPE*	*RUNS*	*RUNM*	*AUTO*	*MEDIA*	*VAR*
(A)	0.997115	0.654617		-0.00395		OK	OK	OK	OK	*
(B)	1.01568	0.650132		-0.00137		OK	OK	OK	OK	*
(C)	1.00025	0.65949		-0.00452		OK	OK	OK	OK	*
(D)	1.02784	0.660567		-0.00892		OK	OK	OK	OK	*
(E)	1.0266	0.657082		-0.01068		OK	OK	OK	OK	*
EL CONDOR STATION										
Modelo	*RMSE*	*MAE*	*MAPE*	*ME*	*MPE*	*RUNS*	*RUNM*	*AUTO*	*MEDIA*	*VAR*
(A)	0.997115	0.654617		-0.00395		OK	OK	OK	OK	*
(B)	1.01568	0.650132		-0.00137		OK	OK	OK	OK	*
(C)	1.00025	0.65949		-0.00452		OK	OK	OK	OK	*
(D)	1.02784	0.660567		-0.00892		OK	OK	OK	OK	*
(E)	1.0266	0.657082		-0.01068		OK	OK	OK	OK	*

References table

RMSE = Root Mean Squared Error Test runs up and down excessive.

RUNM = Test excessive above and below the median runs.

AUTO = Box-Pierce test for autocorrelation excessive.

MEDIA = Test for difference in means between the 1st half and 2nd half.

VAR = Test for difference in variance between the 1st half and 2nd half.

OK = not significant (p> = 0.05).

* = Marginally significant (0.01 <p <= 0.05).

** = Significant (0.001 <p <= 0.01).

*** = Highly significant (p <= 0.001)

El Canal Station: it can be seen that the model that best describes the series of rainfall in the station is an ARMA (1,1).

El Cajon Station - El Condor Station: the model AR (1) is the one that best represents the data for both stations. Then graphics serial analysis of each of the records presented considered. The results are presented in the figure 5.

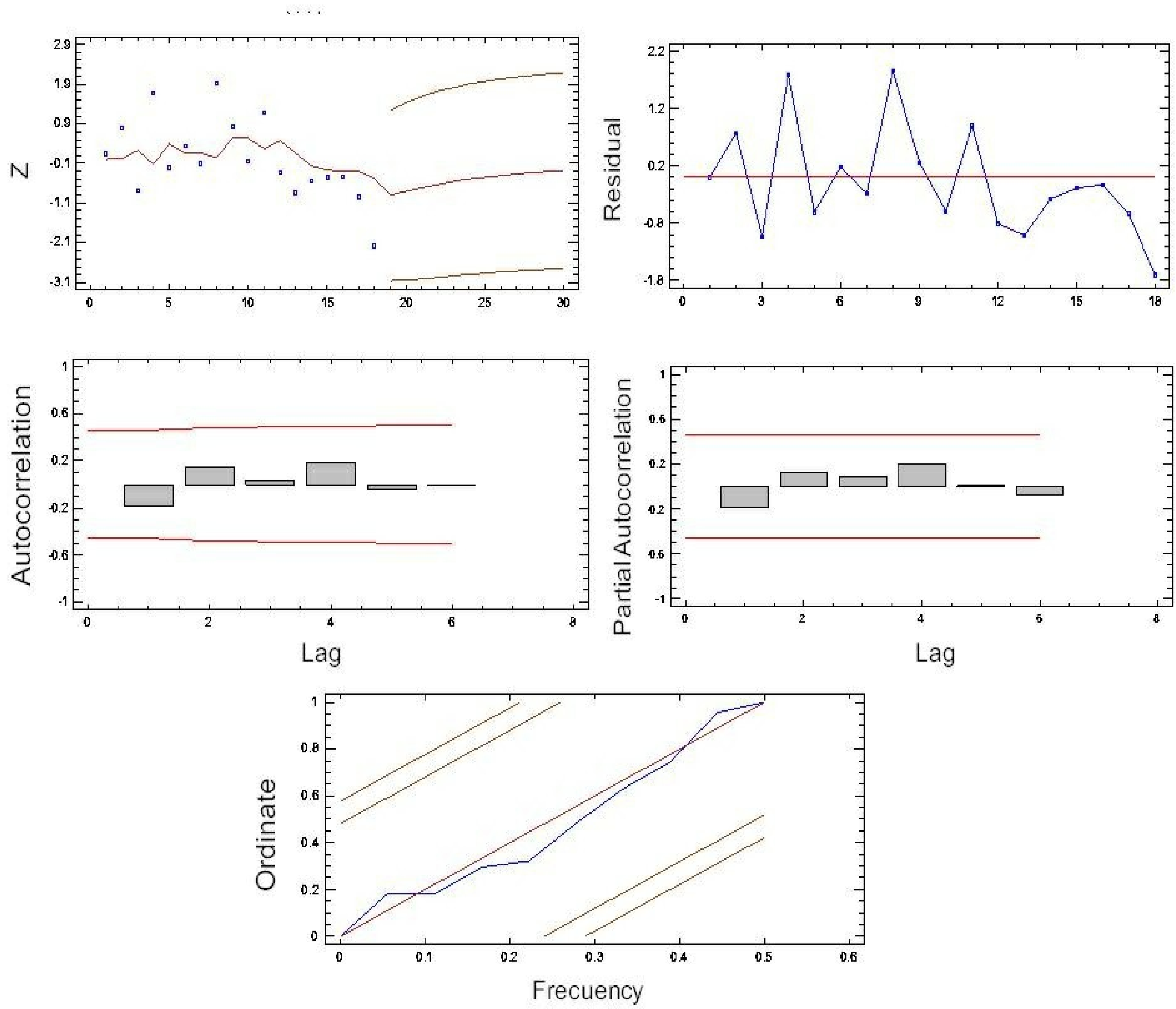

Figure 5: Graph of the original series and the number of residues Barrio Station Canal - Suquía. Selected model ARMA (1,1). Graphics autocorrelation function and partial autocorrelation of the residual series. Periodogram Accumulated of the residual series. The same methodology was followed at El Cóndor and El Cajón series.

Monthly Precipitation Analysis

The monthly rainfall of El Cajon station is presented in Figure 6.

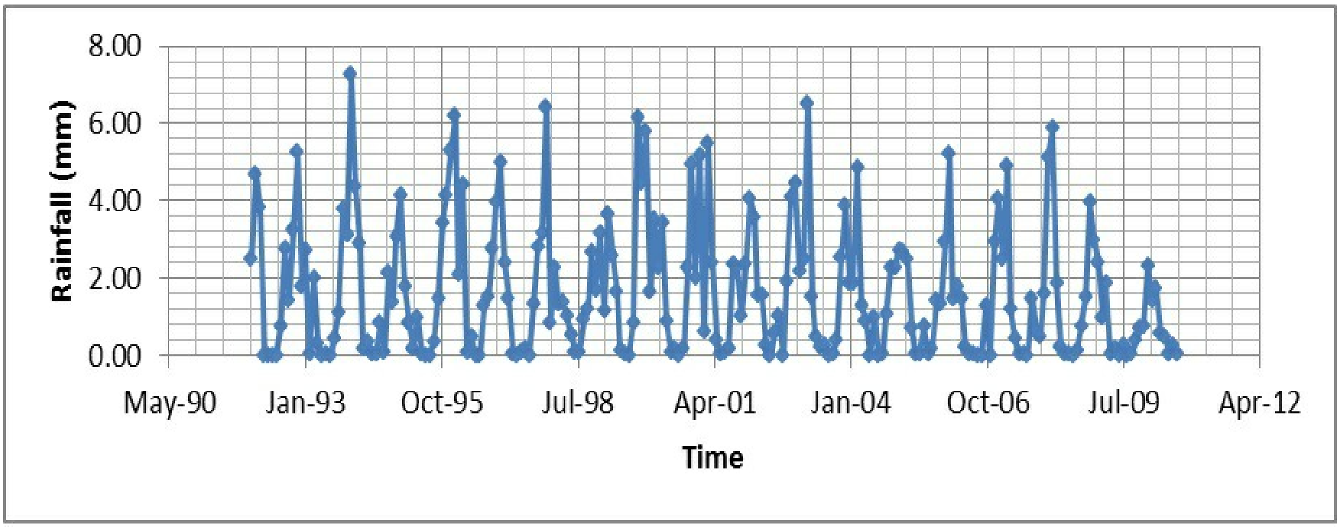

Figure 6. Rainfall Station El Cajon

Cite this Article: Reyna Teresa, Guillén Nicolás, Lábaque María, Álvarez Javier, Eder Matías & Funes Federico "Using Time Series Analysis to support the Water Resources Management in the Upper Basin of the Suquía River" Pinnacle Environmental & Earth Science, Vol. 2 (1) 2015, Article ID pees_179, 531-539, 2015.

The results for the analyzing the monthly series Station El Cajon are presented below. For this case the following models were compared: (A) ARIMA(1,1,1), (B) ARIMA(1,2,1), (C) ARIMA(2,1,1), (D) ARIMA(1,1,2), (E) ARIMA(1,3,1)

Modelo	*RMSE*	*MAE*	*ME*	*RUNS*	*RUNM*	*AUTO*	*MEDIA*	*VAR*
(A)	50.4422	36.5167	0.0551739	*	*	***	OK	OK
(B)	50.4856	36.9216	0.0234208	***	**	***	OK	OK
(C)	50.647	36.8504	0.064663	***	*	***	OK	OK
(D)	50.5315	36.6764	0.0535334	**	OK	***	OK	OK
(E)	66.3836	48.917	4.44181	OK	OK	***	OK	OK

The results shown that best fit are ARIMA (1,1,1). . Figure 7 shows the original series and the synthetic series. It also presents the forecast values and confidence limits of 95 %

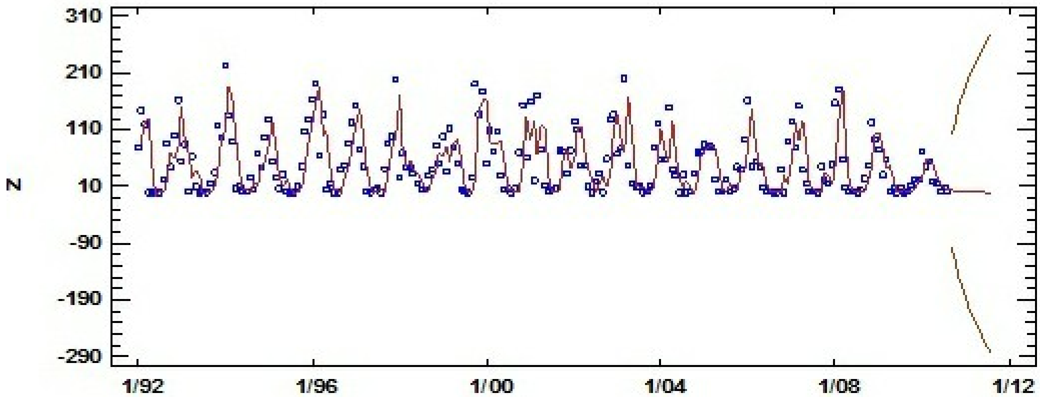

Figure 7: Rainfall Station El Cajon.

Comments and Discussions

Using time series for hydrologic analysis is a powerful tool and allows forecasting especially in the short and medium term. The analyzed times series models indicate that low number of parameters correctly represent the behavior of annual precipitation series.

Among these are the desire to describe variation in variables of interest over time, to gain a better understanding (or explanation) of the data-generating mechanism, to be able to predict future values of a time series, and to allow for the optimal monitoring and control of a system's performance over time (Chatfield, 1989).

Hydrologic time series often exhibit seasonality due to the periodicity of the weather. The models that best represent the series of annual rainfall of Seasons El Cajon and El Condor were ARMA (1,0) models while for Channel station was the ARMA (1,1) model.

The annual average of the stations considered are 763 mm for the El Cajon Station to Station 616 mm and 854 mm Channel for Condor station. With an overall area average of 745 mm. But the patterns of rainfall in the three stations are similar and the low order models represent adequately these records.

The average minimum monthly register for the three series (figure 4) are too similar. The difference exists at the average maximum monthly register.

The minimum register for El Cajon annual series occurred during the period 94/95 (498 mm); for Station El Canal series and El Condor series the minimum ocurred 98/99 (648mm).

This values must be consider for the good water resource management .

Regarding the analysis of monthly series, the ARIMA (1,1,1) model represent adequately the series and will allow planning of water demand for the basin.

Studies also show that it is necessary to increase series studies and monitoring of regional climate variables and policy support management of water resources.

ARIMA models in their general form are used for simulation when one needs to generate many realizations of the process. The hydrologic simulation is particularly useful in the context of water resources management because it yields many different equally likely series.

This results will be useful to make good decisions for the best management of this resource in the province of Cordoba and will permit understand the strong variations in their hydrological cycles. This information will permit ensure too the provision of water and a good quality of life for the city and all its inhabitants.

References

1. Box, G. E. P., Jenkins, G. (1970) Time Series Analysis, Forecasting and Control, San Francisco, Holden-Day.

2. Chatfield, C. (1989). The analysis of time series: An introduction (4th ed.). London; New York: Chapman and Hall.

3. Gujarati, D. N., (1995) Basic Econometrics, 5th Ed. McGraw-Hill Book Co., New York.

4. McLeod, A. E., Hipel, K. W. & Lennox, W. C. (1977). Advances in Box- Jenkins modeling, Applications. Water Resources Research 13(3), 577-586.

5. Reyna, S., Reyna, T., Lábaque, M.; Fulginiti, F. (2012) Diagnóstico de La Provisión de Agua del Gran Córdoba. Alternativas de Solución. XXV Congreso Latinoamericano de Hidráulica. San José, Costa Rica.

6. Salas, J. D., Delleur, J. W., Yevjevich, V., Lane, W. L. (1985) Applied Modeling of Hydrologic Time Series. Water Resources. Littleton, Colorado, USA.

7. http://www.hidricosargentina.gov.ar/sistema_sistema.php

8. http://www.statgraphics.com

Cite this Article: Reyna Teresa, Guillén Nicolás, Lábaque María, Álvarez Javier, Eder Matías & Funes Federico "Using Time Series Analysis to support the Water Resources Management in the Upper Basin of the Suquía River" Pinnacle Environmental & Earth Science, Vol. 2 (1) 2015, Article ID pees_179, 531-539, 2015.