Proceedings of 8 ICSHMO, Foz do Iguaçu, Brazil, April 24-28, 2006, INPE, p. 557-561.

SEASONAL RAINFALL FORECAST METHODOLOGY IN REGIONS OF ARGENTINA

Eduardo A.. FLAMENCO, Cesar M. REBELLA, Stella CARBALLO y Rafel RODRIGUEZ [1]
Instituto Nacional de Tecnología Agropecuaria, Buenos Aires, Argentina

1. INTRODUCTION

Medium and long term climatic predictions attain more importance every time, on account of the impact that some extreme events have in the social economic activities of the region.

Its use in Argentina has a transcendental significance, because the occurrence of flood- periods and dry-periods affect considerably the agricultural activity.

The sea surface temperature and pressure variation, in the Pacific, Atlantic, and Indian oceans and other global climatic indicators as the Southern Oscillation, and Northern Atlantic and Pacific Oscillations, influence the rainfall pattern of the analyzed area.

This investigation has been developed over the base of statistic techniques to link those indexes with the seasonal variability of the rainfall and the medium range forecast are shown. It is important to emphasize that the ability of the models improve with the occurrence of the episodes El Niño/La Niña; not meaning that this signal is the most relevant predictor.

What it is meant, is that this natural phenomenon could be causing alterations of the components of the ocean and atmosphere in distant places of the Equatorial Pacific, being these last, the ones that finally would contribute to improve the seasonal rainfall predictions.

2. MATERIALS AND METHODOLOGY

The goal of this investigation is to describe the methodology that is applied for the development of statistic models of seasonal rainfall forecast (bimonthly, July – August trough March - April) in different regions of Argentina, over the base of indicators that represent the global climatic variability.

The area of study is of approximately 650000 km^2 (Figure 1), and includes the Provinces of Buenos Aires, Santa Fe, Corrientes, Entre Ríos, La Pampa northeastern, Córdoba southeastern and areas of Misiones, Formosa, Chaco and Santiago del Estero.

These models use a methodological approach of the statistic type and the techniques that are applied are Multivariate Adaptive Regression Splines and the Principal Components Analysis.

The data that was used in this study was monthly rainfall in 61 rain gauges, period 1959 - 2001, of the Instituto Nacional de Tecnologia Agropecuaria (INTA) and the Servicio Meteorológico Nacional (SMN).

The global climatic indicators, considered as candidate predictors are obtained of the National Oceanic and Atmospheric Administration and are:

a) Sea Surface Temperature (SST), Reynolds and Smith (1994).
b) Sea Level Pressure (SLP), Woodruff (1987).
c) South Oscillation Index (SOI).
d) Niño1+2, Niño3, Niño3.4, Niño4 indexes.
e) North Atlantic Oscillation (NAO).
f) Pacific Decadal Oscillation (PDO), Mantua (1997).
g) Tropical Atlantic Dipole of the SST.

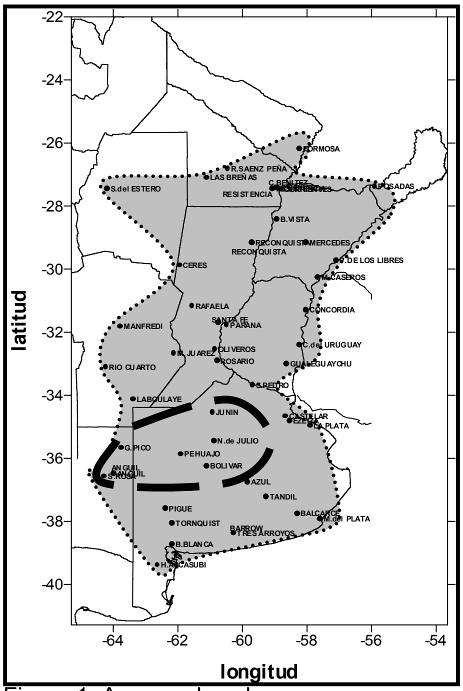

Figure 1. Area analyzed

The Principal Components Analysis (PCA) technique ables to reduce the information of the predictand that are the seasonal rainfall to a less number of variables denominated Principal Components. These are obtained by means of lineal combinations and trying to loose the less quantity of information possible.

The area was divided in regions, which have been delimited in function of the results obtained by applying the ACP technique with the object of capturing the greater temporal - spatial variability of the rainfall according to the analyzed period (variance> 80%).

[1] Instituto Nacional de Tecnología Agropecuaria
Instituto de Clima y Agua
Las Cabañas y Los Reseros S/N
1712 - Castelar, Buenos Aires, Argentina.
Iclima@inta.gov.ar

The non - lineal technique Multivariate Adaptive Regression Splines (MARS), Friedman (1991), is used to link the predictand or the selected principals components with the independent variables or predictors, which means, climatic indicators of the ocean – atmosphere state.

From the principals components forecast, the bimonthly rainfall forecasts are calculated for each rain gauge, linking these with the principals components through the Multiple Linear Regression (step-wise).

The cross - validation technique is applied to evaluate the models skill, and the statistic parameters used are R-square, mean square error, test t-Student of each predictor, and F Fisher-Snedecor.

Applying the Incomplete Gamma distribution function to the historic series of each rain gauge, the occurrence probability 80%, 60%, 40% and 20% are calculate which corresponds to the categories: "Very Wet, Wet, Normal, Dry and Very Dry ".

This classification ables to characterize the rain fall regime to expect for each season, in function of the quantitative forecast.

2.1 Teleconection

In this stage cells in the oceans are identified, where the temperatures and pressures to sea level can be linked statistically (teleconnection) with the spatial - temporal variability of the principals components.

The statistical model titled CLIMLAB - Climate Laboratory, Tanco (2000), developed in the International Research Institute (IRI) of the Lamont Doherty Earth Observatory of the Columbia University (N.Y.;USA), is applied in the case of the SST.

In a similar way of the CLIMLAB model, at INTA there has been developed a software that ables to detect teleconections between the principals components and the SLP of the global ocean, Woodruff (1987, 1993, 1998).

Besides there has been delimited zones in the South Atlantic. These conform two dipoles of pressures, the first one at the coasts of Brazil and Argentina and the second between a zone in front the coasts of Argentina and Africa, that have influence in the variability of the climate in Argentina.

It is important to bring out that the observations that have conform the candidate's predictor's matrix are the ones indicated above, which doesn't mean that all are predictors only that the MARS technique has permitted select those of mayor weight in the explanation of the seasonal variability of rainfalls.

3. RESULTS

We will show an example of the application of this methodology, at the delimited sub-region in the Figure 1, to build the model of rainfall forecast for the bimonthly January – February.

As the models will be developed over the base of teleconection between rainfalls and indicators of the global variability, exist antecedents that express that from the 70s decade, the observations of the temperatures of the surface of the sea and pressure are more consistent.

This on account of that the data are rebuilt considering a blended - mixture of observations in situ, as ship observations, tide stations, moored and drifting buoys, and satellite information incorporated in this analysis from the 80s decade.

Consequently, the majority of the models they were calibrated in the period 1980 - 1999.

3.1 Principal Components Analysis

The result of the application of this technique, Davis (1973), for the period 1968 - 2000 is shown in the Table 1, where over a total of 6 variables that are the bimonthly rainfalls, has been able to reduce that quantity to 2 variables that are the Principals Components, explaining them the 80% of the total variance.

We have stand out in this table the load or significant correlation of the rainfalls in January – February in each rain gauge with relation to the respective associate Principal Component. The Principal Component 1 "CP1" is linked to the variability of the rainfalls of Junin, 9 of Julio, Pehuajo and Bolivar rain gauges, while the CP2 to Anguil and Santa Rosa.

Rain Gauge	**CP1**	**CP2**
Junin	**0.87**	-0.001
9 de Julio	**0.84**	-0.08
Pehuajo	**0.85**	0.17
Bolivar	**0.84**	0.32
Anguil	0.07	**0.95**
Santa Rosa	0.09	**0.94**
Variance	**0.48**	**0.32**

Table 1. Principal Components Analysis

3.2 Identification of candidate predictors (SST and SLP)

In this step the ClimLab model has been applied. The figures 2 and 3 are correlation examples between the components CP1 and CP2, period 1980 – 2000 and the SST of October, period 1979 – 1999, being the lag of 2 months, between the forecast period January – February and the candidate predictors SST of October.

In this figures it is observed the important degree of correlation procured between CP1 and the SST of the month of October in the South Pacific (156º – 135º W, 18º – 26º S) and Tropical Atlantic Oceans (32º – 15º W, 7º N – 2º S), as also as between CP2 and the SST at the Atlantic Ocean (20º – 5º W, 12º – 24º S).

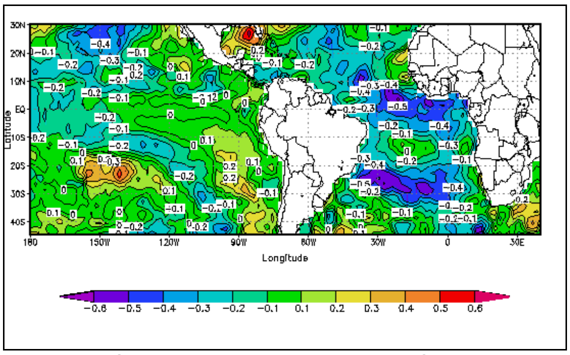

Figure 2. Spatial correlation between CP1 January-February period 1980 - 2000 and SST October period 1979 – 1999.

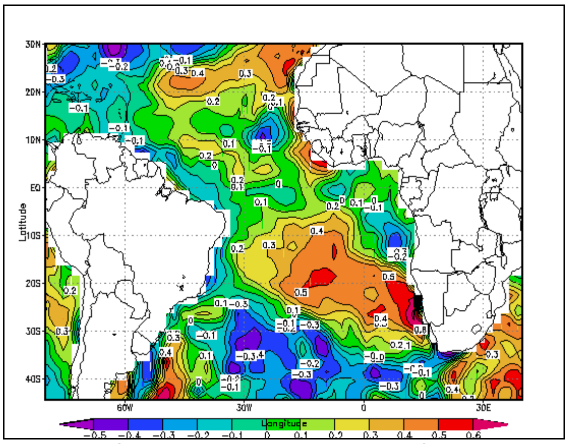

Figure 3. Spatial correlation between CP2 January-February period 1980 - 2000 and SST October period 1979 – 1999.

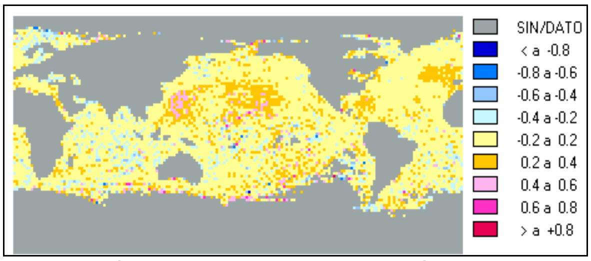

Figure 4. Spatial correlation between CP1 January-February period 1968 - 1998 and SLP October period 1967 – 1997.

Likewise, Figure 4 shows the significant correlation (0.4 < r < 0.6) between the CP1, period 1968 – 1998 and the pressure at sea level of the month of October period 1967 – 1997, at the region located at the North Pacific Ocean (129º - 141º E, 33º – 17º N).

3.3. Statistic models of seasonal rainfall forecast

Applying the MARS technique, the expression (1) is the best obtained model for the region under study and in this case for the First Principal Component is:

$$\mathrm{CP1}_{\mathrm{ef}(t)} = \alpha * \mathrm{FB1} + \beta * \mathrm{FB2} + \varphi * \mathrm{FB3} + \varepsilon \quad (1)$$

Where:

$$\mathrm{BF1} = \text{máx}\,(0, \mathrm{slpOPNW}_{10(t-1)} - 1010.5$$
$$\mathrm{BF2} = \text{máx}\,(0, \mathrm{sstOAT}_{10(t-1)} - 26.75$$
$$\mathrm{BF3} = \text{máx}\,(0, \mathrm{sstOPS}_{10(t-1)} - 0.136$$

R^2 = 0.83
F(3,17) = 27.5
Standard error of the estimations = 0.42.

Being:

CP1 ef(t): standardized index of the First Principal Component bimonth January – February, year [t].

BFi: Basic Functions (splines)

slpOPNW10(t-1): SLP [milibars] month of October year [t – 1] of the region located in the North Pacific Ocean (Figure 5).

sstOAT10 (t-1): SST [ºC] month of October year [t-1] of the region located in the Tropical Atlantic Ocean (Figure 3).

sstOPS10 (t-1): SST [ºC] month of October year [t - 1] of the region located in the South Pacific Ocean (Figure 3).

α, β, φ, ε: regression coefficients

The analysis of the statistics R^2, F – Fisher, standard error of the estimations, t- Student of the predictors prove that the model calibrated is statistically very significant.

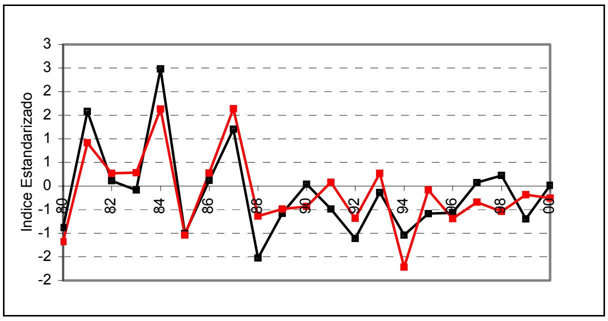

Figure 5. CP1 observed (black line) versus forecasted (red line).

In Figure 5 there has been presented the series observed and forecasted historically of the CP1 applying the technique of crossed validation.

With the purposed of categorically evaluating the results for each region, there has been carried out an analysis of contingency between observations and forecasts of each principal component considering the categories detailed above.

The results, in this case for the CP1, are shown in the Table 2, and we can conclude that of a total of 21 forecast, we have obtained 8 directs coincidences, that represent the 38% of success.

If it is considered a category of difference between observation and forecast; for example the observation corresponds to a Dry period and the forecast to a very Dry one or Normal, the quantity of coincidences is of 17 asserts, which means the 80% of coincidences.

The value calculated of χ^2 (Chi2)=29.98 exceeded the 98% of the value of that distribution, with 4 degrees of freedom, which means, that the probability of having a forecast at random is only of 2%.

Table 2. Contingence analysis

		Forecast				
		VD	D	N	W	VW
Observed	VD	2	2	0	0	0
	D	1	1	1	1	0
	N	1	0	0	2	1
	W	0	0	3	1	0
	VW	0	1	0	0	4

VD: very dry, D: dry, N: normal, W: wet, VW: very wet

It is important to outstand that there has not been a single downfall over 21 forecasted cases, that is to say, the model has forecasted a Very Wet period meanwhile the observed was Very Dry.

Finally, starting from the historic forecasts of each one of the principal components, are forecasted applying the Lineal Multiple Regression technique, the seasonal rainfall January – February of each observation station.

The described methodology has been applied for the development of the regional models of seasonal forecasts of rainfalls for the bimonthlies July - August year [0] trough March - April year [1].

4. CONCLUSIONES

Figures 6, 7, and 8 show examples of the results obtained for the bimonthly January – February during the events Neutral, El Niño and La Niña in the Equatorial Pacific.

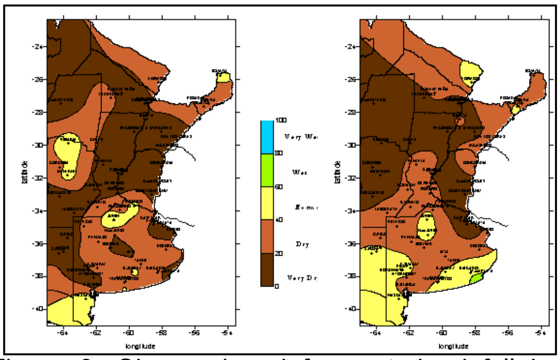

Figure 6. Observed and forecasted rainfall in January – February 1989, expressed in occurrence probability. La Niña event in the Pacific Equatorial Ocean.

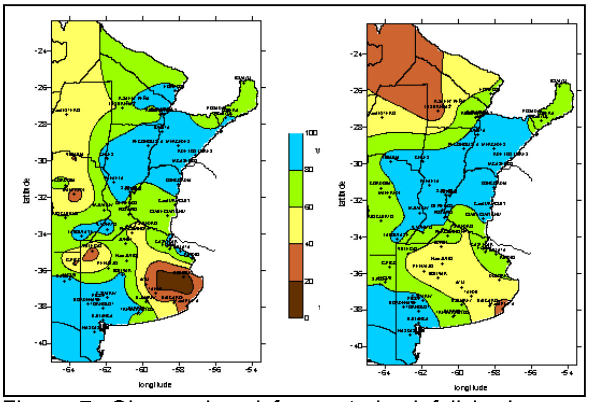

Figure 7. Observed and forecasted rainfall in January – February 1998, expressed in occurrence probability. El Niño event in the Pacific Equatorial Ocean.

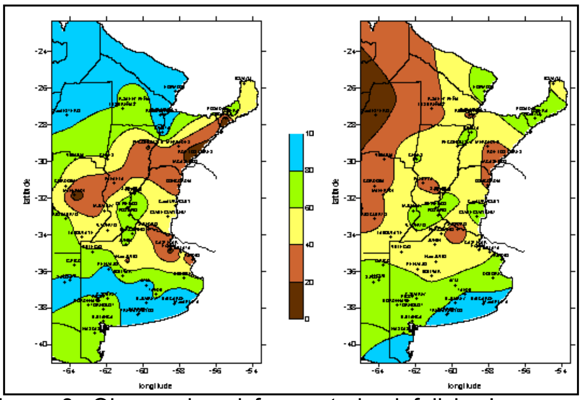

Figure 8. Observed and forecasted rainfall in January – February 2000, expressed in occurrence probability. Neutral event in the Pacific Equatorial Ocean.

This methodology has abled to identify the predictors in different places of the planet, which made possible to obtain good seasonal forecasts in absence of the sign El Niño, La Niña.

Also the strongest El Niño events of the century 1982/83, 1997/98 have been evaluated, where the spatial behavior of the rainfall was different, and these are acceptably forecasted by the developed models, in the same way for the episode La Niña 1988/1989 with dries in all this part of the Argentine territory.

With these criterions, it has been focalized the development of seasonal forecasts of rainfalls in important regions of Argentina, where important agriculture activity exists, that need of this information for the planning and the managing of crops.

We have obtained through the use of statistic techniques to relate rainfalls of different regions of Argentina, with the indexes and the variables that are related to the variability ocean- atmosphere and that the climatology has denominated teleconection, which means that some very important connections exist between the great systems of the atmospheric and oceanic circulation, like the Pacific, Atlantic and Indian Ocean and these regions of the American Continent.

In this study not only it has been considered the data base of the SST, published by the NOAA (NCEP OI Analyses), over the base of the Optimum Interpolation method (Reynolds and Smith, 1994) rather also the atmospheric pressure at sea level (COADS: Comprehensive Ocean and Atmospheric DataSet) , both in a resolution 2º x 2º.

It is also important to have in consideration, that in the case of the SLP, algorithms do not exist as in the case of the SST, that permits having consistent information, being some missing data fundamentally in cells of the South East Pacific and the South Atlantic.

Another important aspect at the moment of compare the observation and the forecasts of rainfalls, is that the net of observation in operation is not the best to represent significant climatic events, if we consider that there has been only used in this study 61 rain gauges irregularly distributed.

Recognition: the authors of this work wish to thank the collaboration of Mrs. Graciela Galvani and Graciela Cazenave.

5. BIBLIOGRAPHY

Davis, J. C., 1973: Statistics and Data Analysis in Geology, pp. 473 – 533.

Friedman, J. H., 1991a: Multivariate Adaptive Regression Splines (with discussion), Annals of Statistics, 10,1-141 (March).

Mantua, N.J. and S.R. Hare, Zhang Y., Wallace J.M., and Francis R.C., 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society, 78, pp. 1069-1079.

Reynolds R. W. and Smith T.M., 1994: Improved global sea surface temperature analyses using optimum interpolation. J. Climate, 7, 929-948.

Tanco R. A. y Berri G. J., 2000: CLIMLAB2000, Version 1.1.0, A Statistical Software Package for Climate Applications, IRI-TR-00/1, International Research Institute (IRI) for Climate Prediction, LDEO, Palisades, N.Y., 10964 USA.

Woodruff, S. D., Slutz R. J., Jenne R. L., and Steurer P. M., 1987: A comprehensive ocean-atmosphere data set. Bull. Amer. Meteor. Soc., 68, 1239-1250.

Woodruff, S. D., Lubker S. J., Wolter K., Worley S. J., and Elms J. D., 1993: Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a: 1980-92. Earth System Monitor, 4, No. 1, 1-8.

Woodruff , S. D., Dias A. F., Elms J. D., and Worley S. J., 1998: COADS Release 2 Data and Metadata Enhancements for Improvements of Marine Surface Flux Fields. Phys. Chem. Earth, 23, 517-527.