

MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA DETERMINAR INFLUENCIAS DEL ÍNDICE NIÑO 1+2 Y LA MJO SOBRE LAS PRECIPITACIONES EN GUAYAQUIL DURANTE ENERO-FEBRERO-MARZO Y ABRIL

LINEAR REGRESSION MODEL TO DETERMINE THE INFLUENCE OF EL NIÑO 1+2 AND MJO INDEXES OVER THE PRECIPITATION IN GUAYAQUIL DURING JANUARY-FEBRUARY-MARCH AND APRIL

Por: Juan José Nieto ⁽¹⁾

Resumen

Se elaboraron modelos de regresión lineal múltiple para determinar el grado de ajuste de los índices niño 1+2 y la Oscilación Madden-Julian con la precipitación acumulada en Guayaquil para enero, febrero, marzo y abril. Los modelos de correlación se construyeron por separado para cada mes con series de tiempo de 1978 al 2004 de precipitación mensual acumulada, anomalía del índice niño1+2 y el índice de la oscilación Madden-Julian en la longitud 120W, con un nivel de confianza de 95%. Se encontró que para enero, marzo y abril, el modelo de correlación lineal se ajusta con coeficientes R2 de 0.80, 0.77 y 0.82 respectivamente. Los niveles de precipitación de febrero no muestran correlación significativa con los índices, parecen responder a otros factores, presumiblemente atmosféricos. Se hicieron comparaciones entre los valores calculados a partir de los modelos y los niveles de precipitación observados de los años 2005, 2006 y 2007 a fin de determinar el grado de acierto de los pronósticos. Para efectuar las comparaciones se estimaron tres rangos en base a los terciles de la serie de tiempo de datos observados. Los rangos se definieron como bajo lo normal (tercil inferior), normal (tercil intermedio) y sobre lo normal (tercil superior). Se encontró que para abril el grado de acierto es de 100%, sin embargo en enero y marzo, los valores pronosticados mostraron un bajo nivel de acierto. El modelo de regresión encontrado puede servir como herramienta para encontrar las relaciones entre las precipitaciones, temperatura del mar y la convección atmosférica.

Palabras Clave: *Precipitación atmosférica, Circulación Atmosférica, Temperatura superficial del mar, variaciones estacionales, Análisis matemático*

Abstract

There were elaborated multiple linear regression models to determine the adjustment degree of niño1+2 and Madden-Julian Oscillation indexes with the accumulated precipitation in Guayaquil, for January, February, March, and April. The correlation models were built separately for each month with time series from 1978 to 2004 of monthly accumulated precipitation, niño 1+2 anomaly, and Madden-Julian Oscillation index for 120W, with a significance level of 95%. It was found that for January, March, and April the linear correlation adjusts with coefficients R2 of 0.80, 0.77, and 0.82 respectively. The precipitation levels for February does not showed significant correlation with the indexes, it seems to respond to other factors, presumably atmospheric.

There were made comparisons between the values calculated with the models and the observed precipitation levels of the years 2005, 2006, and 2007 in order to determine the hit score of the forecast. To perform comparisons there were established three ranges based on the terciles of the observed data time series. The ranges were defined as below normal (lower tercil), normal (middle tercil), and above normal (upper tercil). It was found that for March the skill score is 100%, however, forecasted values for January and March, showed a low skill score. The regression model found may serve as a tool to find relations between precipitations, sea temperature and atmospheric convergence.

Key Words: *Atmospheric precipitations, Atmospheric circulation, Sea Surface Temperature, seasonal variations, Mathematical analysis*

⁽¹⁾ Instituto Oceanográfico de la Armada, INOCAR, Av. 25 de Julio Base Naval Sur, P.O. Box 5940 Guayaquil-Ecuador

Introducción

El clima en Guayaquil y en general en la costa del Ecuador está influenciado a más de los sistemas atmosféricos por las condiciones térmicas del océano pacífico circundante. La estación de lluvias en la costa se inicia en términos generales entre diciembre y enero y finaliza entre abril y mayo. Los niveles de precipitación varían de año a año, oscilando en un amplio rango, entre estaciones muy secas a estaciones muy húmedas debido especialmente a las manifestaciones de los eventos ENSO y a la influencia estacional de los sistemas atmosféricos.

Durante los eventos de la fase positiva del ENSO donde la anomalía de la temperatura del mar puede llegar hasta +4°C (Barnston, 1997) los niveles de precipitaciones alcanzan excesos de hasta 300%. Esta particularidad no se repite necesariamente durante eventos moderados o débiles, donde las precipitaciones parecen obedecer también de forma significativa a otros factores.

La temperatura del mar es monitoreada a través de varios sistemas de boyas y satélites colocadas en el océano pacífico ecuatorial, para fines de estudio, se lo ha dividido en 4 regiones claramente definidas, entre la parte oriental y occidental. De la zona se tiene registros desde la década de los 50, de valores promedio mensuales observados y de anomalías. Para el presente estudio se han utilizado las anomalías medidas en la región niño1+2, delimitada entre 0-10S-90W-80W, (Fig. 1).

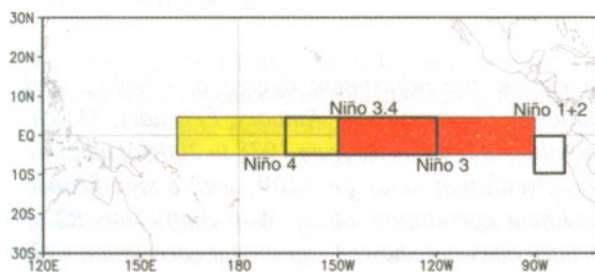


Figura 1. Regiones Niño delimitadas a lo largo del pacífico ecuatorial. Nótese al este el pequeño recuadro blanco de la región niño1+2.

Existen importantes niveles de correlación entre las precipitaciones de la costa del Ecuador y la región niño 1+2 (Hoerling, 2001), la cual explica más del 50% de la variabilidad total del régimen de precipitaciones. El resto de la varianza puede ser explicado por el aporte de los sistemas atmosféricos globales y regionales (Kousky, 1984). Se ha buscado encontrar el aporte modulador de la Oscilación Madden-Julian MJO, la cual es una fluctuación intraestacional que ocurre en los trópicos globales. La MJO es responsable por la mayoría de la variabilidad

del tiempo en los trópicos. La misma provoca variaciones en varios parámetros oceánicos y atmosféricos (Gottschalck, 2006) como velocidad y dirección del viento de altura y superficie, nubosidad, lluvia, temperatura superficial del mar y evaporación en la superficie del océano. Es la componente natural del sistema acoplado océano-atmósfera con un ciclo aproximado de 30-60 días. Existe evidencia que la MJO puede modular el ciclo ENSO en la velocidad de desarrollo y probablemente en la intensidad de los episodios (Kessler and Kleeman, 2000; Zhang and Gottschalck, 2002).

Metodología

Se analizaron series de tiempo de precipitación mensual en una estación meteorológica ubicada al sur de Guayaquil, con registros desde 1964 hasta 2007, en los que se determinó los rangos de percentiles 33% y 66% para establecer niveles de precipitación normal (33% < precipitación < 66%), bajo lo normal (precipitación < 33%) y sobre lo normal (precipitación > 66%) en esa escala porcentual. Con la serie mensual de anomalía de los índices niño 1+2 desde 1950 hasta 2007, se graficó la dispersión de los datos para observar la respuesta y un probable ajuste lineal. Se calculó un coeficiente de correlación múltiple y una correlación cruzada para determinar posibles desfases.

El coeficiente de correlación múltiple entre las series, se calculó mediante la ecuación:

$$\rho_{x,y} = Cov(X, Y) * (\rho_x \rho_y)^{-1}$$

Una vez comprobada la linealidad, con la asunción de normalidad e independencia lineal (Ripley, 1981), se utilizó la serie de tiempo de los índices MJO correspondientes a la longitud 120W (la más cercana a la ubicación de la estación meteorológica 80W) y se calculó una regresión lineal múltiple para cada mes con un nivel de confianza del 95%, tratando de construir modelos múltiples mensuales, a través del método de mínimos cuadrados:

$$b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}, \quad b_0 = \bar{y} - b_1 \bar{x}$$

Para finalmente construir un modelo de la forma:

$$Y = b_0 + b_1 * X_1 + b_2 * X_2 + \dots + b_p * X$$

Donde Y es la precipitación acumulada para cada mes, b_0 es el coeficiente de determinación, X_1 la anomalía de temperatura en el Niño₁₊₂ y X_2 es el índice MJO.

Cada modelo fue evaluado en cuanto a su relevancia

estadística mediante el estadístico F para determinar la probabilidad de que las varianzas de las series no presentes diferencias significativas. Los resultados de la regresión lineal fueron tabulados para decidir con qué meses se construirían las ecuaciones de regresión, basados en los coeficientes de correlación y el estadístico F.

Se realizaron graficas de las series de tiempo de los datos de precipitación observados y calculados, utilizando los modelos mensuales en aquellos meses en que la correlación múltiple sobrepasa el 70%, como respuesta de la explicación de su varianza en base a la temperatura de mar en la región Niño₁₊₂ y el índice de la MJO.

Para comprobar la validez de los modelos mensuales se comparó el resultado de pronóstico para un rango fuera del período de construcción, esto es para los años 2005, 2006 y 2007 en la estación Guayaquil. Los

niveles pronosticados de precipitación por los modelos mensuales fueron contrastados con los valores observados, otorgándole una calificación de acierto, si el valor pronosticado se ubica dentro del rango del tercil calculado para la serie mensual de la estación y de no acierto, si el pronóstico se ubica fuera del rango. Los límites de los rangos de los percentiles fueron redondeados al entero más cercano para fines prácticos.

Resultados

La estimación inicial de distribución lineal en los diagramas de dispersión muestra linealidad con coeficientes de correlación superiores a 0.60 en enero, marzo y abril (Fig. 2a, 2c, 2d), no así en febrero ($r^2 < 0.50$), donde la dispersión de los datos de precipitación hacen presumir que ésta no responde de forma clara al forzamiento de la temperatura del mar (Fig. 2b)

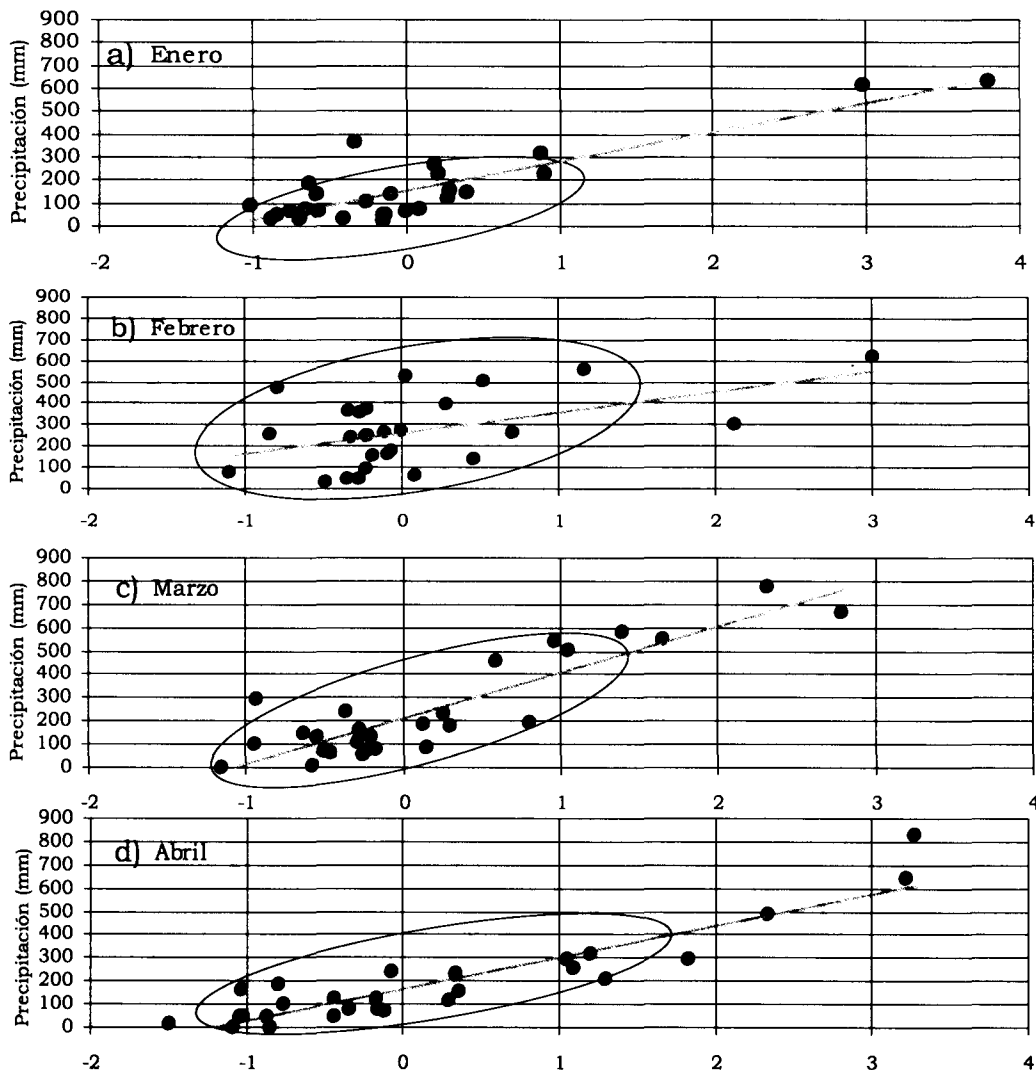


Figura 2. Diagramas de dispersión de Precipitación vs. Anomalía Niño1+2. a) enero, b) febrero, c) marzo, c) abril. Nótese el incremento de dispersión para febrero. La línea continua representa el ajuste lineal de los datos. Los puntos que se desagregan del grupo de forma marcada son los eventos El Niño 82-83 y 97-98.

Los coeficientes de ajuste de los modelos de regresión lineal múltiple generados, muestran una situación similar en cuanto a la respuesta mensual de la precipitación en función de las anomalías nino1+2, el coeficiente r^2 ajustado para enero, febrero y marzo es > 0.75 , mientras que para febrero es de 0.19 (tabla 1). El cálculo del estadístico F que muestra que el

ajuste de las variables independientes no se produce por el azar, en todos los casos lo confirma, sin embargo en febrero el estadístico y su valor crítico es bastante cercano. Con estos argumentos se decidió no generar un modelo de regresión para febrero puesto que podría llevar a conclusiones erróneas.

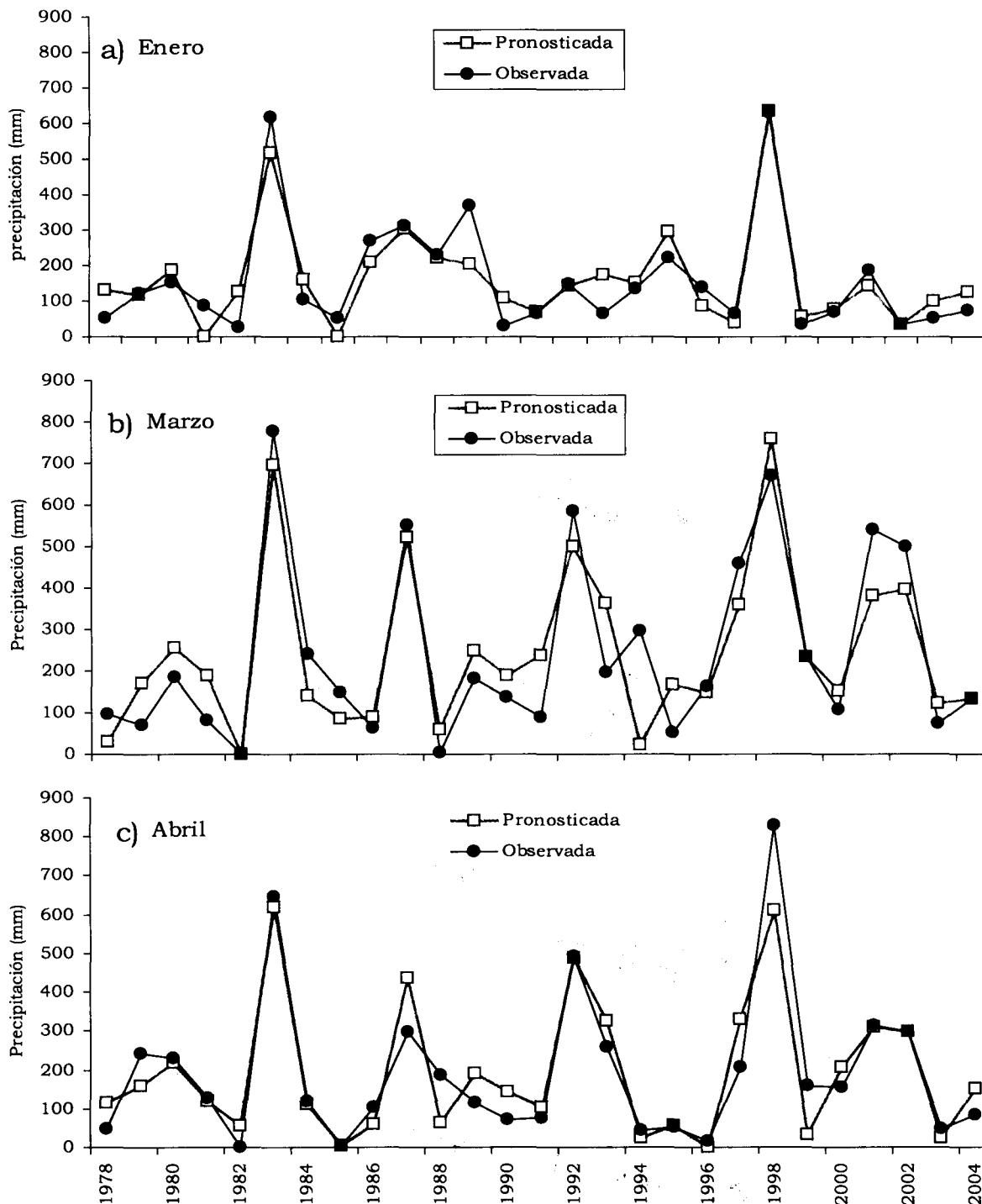


Figura 3. Reconstrucción de serie de tiempo de precipitación en base a los modelos de regresión múltiple. a) enero, b) marzo y c) abril. La línea negra representa el valor real observado de precipitación, la línea en tono de gris representa el valor de precipitación pronosticado. No se ha representado el modelo de febrero por la baja correlación encontrada entre las variables. Los coeficientes de correlación r^2 de las series pronosticada vs. observada son 0.92, 0.89 y 0.92 respectivamente.

	0.80	54.3	$1.2 \cdot 10^{-9}$	145.05	153.18	70.84
	0.19	4.0	0.03	*	*	*
	0.77	44.7	$8.0 \cdot 10^{-9}$	212.01	183.46	-29.37
	0.83	63.2	$2.7 \cdot 10^{-10}$	164.73	130.61	-31.89

Tabla 1. Resumen de resultados del cálculo de regresión lineal utilizados para elaborar las ecuaciones del modelo de regresión. Para febrero no se calcularon coeficientes de regresión en vista de la pobre correlación de las variables.

$$\text{Precipitación}_{\text{marzo}} = 212.01 + 183.46 \cdot \text{anomalía}_{\text{nifo1+2 marzo}} - 29.37 \cdot \text{MJO}_{120\text{W marzo}}$$

Tabla 2. Ecuaciones del modelo de regresión lineal múltiple a partir de las cuales se reconstruyó la serie de precipitación para cada mes.

Utilizando las ecuaciones del modelo de regresión lineal múltiple se reconstruyó la serie con valores pronosticados, las cuales se contrastan con los valores observados (Fig. 3). Se encontró una alta correlación entre ambas series, con coeficientes r superiores a 0.85.

La calificación de aciertos de los pronósticos para los años posteriores al período de análisis, se realizó en base a los percentiles mensuales 33% y 66% (tabla 3). Los rangos se utilizaron para calificar las precipitaciones observadas y pronosticadas como, bajo lo normal, normal y sobre lo normal.

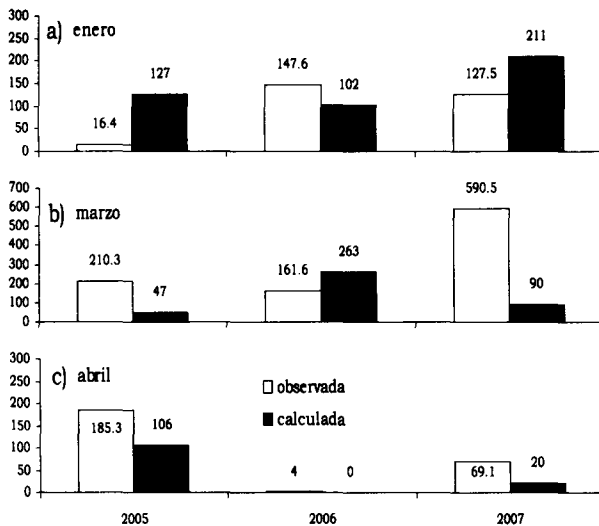


Figura 4. Comparación de los resultados de la precipitación calculada vs. la precipitación observada para los años fuera del rango de construcción de la ecuaciones de regresión. a) enero, b) marzo y c) abril. Las máximas diferencias encontradas entre los observado y pronosticado son 110mm, 500mm y 80mm para enero, marzo y abril, respectivamente. De acuerdo a los rangos de percentiles 33% y 66%, el pronóstico no acertó en enero y marzo de buena forma, son así en abril, en que el porcentaje de acierto fue 100%, basado en el criterio de anomalía en base a percentiles.

CONCLUSIONES

1. Existe una alta correlación entre los datos observados y los pronosticados por el modelo de regresión, es decir que la anomalía de temperatura del mar y en índice de la Oscilación Madden-Julian en la longitud 120W, para los enero, marzo y abril. No así para febrero en que la precipitación presumiblemente responde a condiciones atmosféricas principalmente.
2. El modelo de pronóstico estadístico responde de mejor forma a abril, donde se lograron porcentajes de acierto de 100% bajo el criterio de límites de percentiles para rangos bajo lo normal, normal y sobre lo normal. Bajo este criterio, enero y marzo presentaron un bajo nivel de acierto.
3. El método y las variables usadas no pueden ser utilizados como herramienta de pronóstico, sino como un mecanismo para entender las relaciones entre las precipitaciones, la temperatura del mar y la convergencia atmosférica. Sin embargo con los resultados encontrados es posible establecer posibles tendencias de las precipitaciones a corto plazo.
4. Existen precipitaciones que salen de rango en función de parámetros locales y no responden a forzamientos regionales necesariamente.

Bibliografía

- Barnston, A.G., M. Chelliah and S.B. Goldenberg, 1997. Documentation of a highly ENSO-related SST region in the equatorial Pacific. *Atmosphere-Ocean*, 35, 367-383.
- Hoerling, M. P., A. Kumar, and T. Xu, 2001. Robustness of the nonlinear climate response to ENSO's extreme phases. *J. Climate*, 14, 1277-1293.
- Kousky, V.E., M.T. Kayano, and I.F.A. Cavalcanti, 1984. The Southern Oscillation: Oceanic-atmospheric circulation changes and related rainfall anomalies. *Tellus*, 36A, 490-504.

Aline de H.N. Maia, Meinke Holger and Sarah Lennox, 2004. Assessment of probabilistic forecast 'skill' using p-values.

V. Mossmann, A. Castro, J.L. Sánchez, 2005. Análisis de la tendencia de las precipitaciones en los meses de febrero, marzo y abril en el territorio peninsular de España.

Ripley, B.D., 198. Spatial Statistics. Wiley Series in probability and mathematical statistics, Jhon Wiley & Sons, New York, 252 pp.

Stone, R.C., Hammer, G.L. and Marcussen, T., 1996a. Prediction of global rainfall probabilities using phases of the southern oscillation index. Nature 384: 252-255.

Gottschalck J., Kousky V., Higgins W., and L'Heureux M. The Madden-Julian Oscillation, 2006.

Kessler, W. and R. Kleeman, 2000. Rectification of the Madden-Julian Oscillation into the ENSO cycle. J.Climate, 13, 3560-3575.

Zhang, C. and J. Gottschalck, 2002. SST anomalies of ENSO and the Madden-Julian Oscillation in the equatorial Pacific. J. Climate, 15, 2429-2445.