

AJUSTE DE UN MODELO VAR COMO PREDICTOR DE LOS CAMPOS DE ANOMALÍAS DE PRECIPITACIÓN EN CENTROAMÉRICA

ERIC J. ALFARO* – F. JAVIER SOLEY**

Recibido: 3 Setiembre 1999

Resumen

Con el fin de identificar en la región las estaciones pluviométricas con curvas de anomalías similares entre sí, 15 puntos de precipitación de la región fueron sometidos a un proceso de agrupación identificándose cinco conglomerados. Posteriormente se ajustó un modelo Vectorial Autorregresivo VAR, con el objetivo de cuantificar la interacción océano-atmósfera entre distintos índices oceanográficos en el Pacífico y Atlántico Tropical y los regímenes de precipitación en Centroamérica, representados por las primeras funciones ortogonales empíricas de los distintos conglomerados. Este modelo mostró que la principal influencia sobre la región la ejerce el Atlántico Tropical Norte con correlaciones positivas. Por su parte los índices del Atlántico Tropical Sur y del Niño 3 no mostraron una influencia significativa sobre el istmo. Se concluye que al ajustar un modelo estacionario, las anomalías de la temperatura superficial del mar en el Atlántico Tropical Norte influyen más fuertemente sobre la precipitación que aquellas del Niño 3 o del Atlántico Tropical Sur influyendo sobre el grado de formación de la Vaguada Tropical Troposférica Alta.

Palabras clave: ENOS, Variabilidad en el Atlántico, Modelos VAR, Precipitación en Centroamérica.

Abstract

Cluster analysis was used to identify common patterns of 15 precipitation points of the region, using their anomaly time series as grouping variables. Five clusters were identified through this process. A Vector Auto Regressive model was fitted to

*DFAOP-Escuela de Física, Centro de Investigaciones Geofísicas y Centro de Investigaciones en Ciencias del Mar y Limnología, Universidad de Costa Rica, San José, Costa Rica. E-Mail: ejalfaro@cariari.ucr.ac.cr

**DFAOP-Escuela de Física y Centro de Investigaciones Geofísicas, Universidad de Costa Rica, San José, Costa Rica. E-Mail: fjsoley@hermes.efis.ucr.ac.cr

the data to quantify the ocean-atmosphere interaction between the oceanic indices of the Tropical North and South Atlantic, the Tropical Eastern Pacific and the first empirical orthogonal functions of the regional rainfall clusters. This model shows that the Tropical North Atlantic has the largest influence over the region when compared with the influence of the other indices, having positive correlation with all the rainfall. The Tropical South Atlantic and the Niño 3 indices, instead, were found to have no correlation with the rainfall of the region when an stationary model is fitted. This work shows that the variability of the Tropical North Atlantic sea surface temperature anomaly presents stronger associations with the Central America rainfall than the Tropical Eastern Pacific sea surface temperature anomaly. The association is mainly related to the degree of development of the Tropical Upper Tropospheric Trough.

Keywords: ENSO, Atlantic Variability, VAR Models, Central American Precipitation.

Mathematics Subject Classification: 62M10, 62H30.

1. Introducción

A pesar de que las aplicaciones de modelos Vectoriales Autorregresivos-Medias Móviles (VARMA) en el continente americano han sido limitadas (Hastenrath, 1995) se pueden citar dos ejemplos. El primero es el de Alfaro y Cid (1999a) en donde ellos agruparon las curvas de anomalías de 72 estaciones sobre la región centroamericana con el fin de identificar las curvas con comportamientos similares entre sí. Ellos identificaron cinco conglomerados y ajustaron un modelo VARMA entre las primeras funciones ortogonales empíricas (EOF's por sus siglas en inglés) de los distintos conglomerados y los índices de Temperatura Superficial del Mar (TSM) para las regiones del Atlántico Tropical Sur (ATS) y Norte (ATN) y el Pacífico Tropical Este Niño 3. Este modelo mostró que la principal influencia sobre la región la ejerce el ATN con correlaciones positivas sobre la toda la región. Por su parte el índice Niño 3 mostró una influencia más débil, con correlaciones negativas sobre aquellas regiones situadas principalmente en la Vertiente Pacífica de Centroamérica.

El segundo ejemplo lo presenta Liu *et al.* (1998). Como primer caso de estudio ellos usaron un índice de precipitación para las islas del Pacífico abarcando una región comprendida entre los $4^{\circ}N - 12^{\circ}S$ y los $152,5^{\circ}E$ y $82,5^{\circ}W$, y este índice, junto con el del Niño 3, se usaron para ajustar un modelo VARMA mostrando una relación positiva entre la precipitación insular y la TSM del Niño 3. Como segundo caso de estudio ellos utilizaron el mismo tipo de modelos para estudiar las relaciones entre los caudales de los ríos Nare y Grande en Colombia y el índice Niño 3, encontrando una relación negativa entre los caudales y la TSM en esa región oceánica.

De estos dos trabajos citados anteriormente, se desprende que la familia de modelos VARMA (dentro de la cual se incluyen los modelos Vectoriales Autorregresivos o VAR) son una herramienta útil en el pronóstico de las anomalías de variables meteorológicas e hidrográficas. Es por ello que el objetivo principal de este trabajo es optimizar un modelo de esta familia, en su representación de espacio de estados, que permita validar el modelo de Alfaro y Cid (1999a) para cuantificar la interacción océano-atmósfera entre distintos índices en el Pacífico y Atlántico Tropical y los regímenes de precipitación en Centroamérica, usando el conjunto de datos preparado por Xie y Arkin (1997). Esto es

importante pues permite describir cuantitativamente las relaciones estacionarias entre las series, asignándole un valor de probabilidad a las relaciones ya que la variable climática más importante de predecir en Centroamérica es la precipitación. Además de lo anterior, este estudio permitiría evaluar el uso potencial de datos de precipitación preparados en arreglos equiespaciados, lo cual podría ser usado como guía para ajustar salidas de modelos numéricos tal como el *MM5*, actualmente en uso en el Centro de Investigaciones Geofísicas de la Universidad de Costa Rica.

La región que comprende este estudio abarca desde el sur de México en el Norte, hasta Costa Rica y Panamá en el sur o sea el istmo centroamericano enmarcado en la Fig. 1 entre los entre $8,75 - 18,75^{\circ}N$ y $78,75 - 93,75^{\circ}W$. Esta limita al noreste con el Mar Caribe y al suroeste con el Océano Pacífico, ambos con sistemas de corrientes superficiales de deriva hacia el noroeste y temperaturas superficiales promedio de $28^{\circ}C$. Los principales sistemas montañosos de Centroamérica están ubicados en un eje noroeste-sudeste cercano a la costa pacífica con cumbres que sobrepasan los $3000m.s.n.m.$ Este eje cordillerano es bastante irregular y divide el territorio centroamericano en dos vertientes principales: Caribe y Pacífico.

2. Datos y Metodología

Los índices de TSM del Atlántico Tropical y Niño 3 fueron los mismos usados y descritos ampliamente por Alfaro y Cid (1999a) y definidos por Enfield (1996), debido a que ellos están asociados con áreas oceánicas que presentan alguna influencia sobre la precipitación en Centroamérica (e.g. Alfaro *et al.*, 1998; Enfield y Alfaro, 1999). Las áreas del Atlántico Tropical Norte (ATN) y Sur (ATS) usadas fueron $22^{\circ}N - 6^{\circ}N, 80^{\circ}W - 15^{\circ}W$; y $2^{\circ}N - 22^{\circ}S, 35^{\circ}W - 15^{\circ}E$, respectivamente. El área usada para la construcción del índice Niño 3 fue $6^{\circ}N - 6^{\circ}S, 150^{\circ}W - 90^{\circ}W$. Los índices oceanográficos usados en este trabajo, fueron calculados como el promedio de los puntos de la rejilla dentro del área oceánica determinada. Estos fueron obtenidos del Atlantic Oceanographic and Meteorological Laboratory (AOML), National Oceanic and Atmospheric Administration (NOAA) de Estados Unidos de América (EUA). El período utilizado de estos índices y de los registros de precipitación, comprendió de enero de 1979 a diciembre de 1995.

A diferencia del trabajo de Alfaro y Cid (1999a), en este estudio se usaron datos de precipitación mensual acumulada para 15 puntos sobre o muy cercanos al istmo (menos de $100km$ alejados de la costa), de la base de datos de Xie y Arkin (1997), entre $8,75 - 18,75^{\circ}N$ y $78,75 - 93,75^{\circ}W$, en una malla de $2,5^{\circ} \times 2,5^{\circ}$. Estos datos globales fueron construidos como promedios espaciales a partir de estaciones meteorológicas, estimados satelitales y salidas de modelos numéricos. Una de las ventajas del uso de este conjunto de datos es que ellos son fácilmente accesibles por INTERNET, mientras que los datos de las estaciones meteorológicas no lo son y además requieren un proceso exhaustivo de control de calidad. Este trabajo permite por lo tanto el comparar los resultados de estas dos aproximaciones al campo de precipitación. Además de lo anterior, la mayoría de los modelos numéricos y conjuntos de datos globales se encuentran almacenados en rejillas de promedios espaciales por lo que es de interés a la comunidad geofísica establecer la concordancia entre datos de

punto (estaciones) con datos en rejillas.

Como primer paso, se procedió al análisis exploratorio de los datos, para corregir posibles tendencias y puntos atípicos. Luego se obtuvieron las series de anomalías estandarizadas de las estaciones, como la diferencia entre el valor mensual de precipitación y su promedio mensual dividido por su desviación estándar, estas últimas fueron calculadas tomando en cuenta el total del período de la serie. Este procedimiento desestacionaliza la serie al eliminar los picos espectrales correspondientes al ciclo anual y sus armónicos de 6, 4 y 3 meses. Las secuencias estacionales no son estacionarias y deben ser preblanqueadas (“pre-whitened”) con un procedimiento como el seguido para obtener las anomalías y utilizarlas en un modelo estacionario tipo VAR o VARMA (Lütkepohl, 1993).

Con el fin de identificar las estaciones con curvas de anomalías estandarizadas similares entre sí, las series fueron sometidas a un proceso de agrupación, mediante técnicas de análisis de conglomerados (cluster analysis). La agrupación de las estaciones se produjo en dos pasos. Primero se usó el método de conglomerados jerárquicos de varianza mínima de Ward (JMP, 1995). El criterio para la agrupación se fundamentó en el análisis del dendrograma. Se seleccionó el número máximo de grupos con disimilitud del 90 %. El análisis de conglomerados fue validado por medio de análisis discriminante usando como clasificación a priori la determinada por el análisis de conglomerados.

Posteriormente, se procedió al cálculo de las EOF's de los grupos y se conservaron aquellas que retuvieran la información de gran escala. Luego se utilizaron estas funciones, más los índices ATN, ATS y Niño 3, como entrada para el cálculo del modelo de espacio de estados (SAS, 1993), La formulación del modelo, la estimación de los parámetros y el algoritmo usado están descritos ampliamente en Alfaro y Cid (1999a). Este tipo de modelos es definido por la siguiente ecuación de transición de estado

$$\mathbf{z}_{t+1} = \mathbf{F}\mathbf{z}_t + \mathbf{G}\mathbf{e}_{t+1} \quad (1)$$

donde \mathbf{z}_t es el vector de estado de dimensión $s \times 1$, \mathbf{F} es la llamada matriz de transición y sus coeficientes determinan las propiedades dinámicas del sistema, su dimensión es de $s \times s$. \mathbf{G} es llamada la matriz de innovación $s \times r$, ella determina la estructura de la varianza de la ecuación de transición. El vector de entrada \mathbf{e}_t es una sucesión de vectores independientes aleatorios normalmente distribuidos y de dimensión r , con media cero y matriz de covarianza Σ_{ee} , este vector es llamado también vector de innovaciones.

Además de la ecuación de transición, se incluye una ecuación de variables observables \mathbf{x}_t , que nos da sus valores como una función de \mathbf{z}_t y de dimensión $r \times 1$, la ecuación de variables observables en este caso, es la extracción de las primeras r componentes del vector de estado, o sea

$$\mathbf{x}_t = [\mathbf{I}_r \mathbf{0}] \mathbf{z}_t \quad (2)$$

donde \mathbf{I}_r es una matriz identidad de $r \times r$.

El período utilizado para el ajuste y evaluación del modelo fue de 1979 a 1995, o sea el mismo de la serie de tiempo. Para evaluar la calidad del ajuste del modelo, se verificó la aleatoriedad de los residuos por medio de dos criterios propuestos por Vandaele (1983). El primero contempla la obtención del periodograma integrado de los residuos a partir de la distribución normal aproximada (Jenkins y Watts, 1968; Box y Jenkins, 1976) y por