

## APLICACIÓN DE LOS PATRONES PRINCIPALES DE ACOPLAMIENTO AL ESTUDIO DE LA PREDECIBILIDAD EN EL ATLÁNTICO NORTE

Emilia SÁNCHEZ GÓMEZ y M<sup>a</sup> José ORTIZ BEVIÁ  
*Departamento de Física, Universidad de Alcalá de Henares*

### RESUMEN

En este trabajo se presenta una aplicación de los patrones de variabilidad acoplada a la predicción. Los patrones se obtienen mediante la Descomposición en Valores Singulares de la matriz de covarianza cruzada entre dos campos climatológicos: el predictor y el predictando. La matriz de covarianza cruzada se calcula con un determinado desfase entre el campo predictor (el que precede) y el campo predictando. Se han realizado diversos experimentos de predicción, con tres predictores: las temperaturas de la superficie del mar y la temperatura del aire a 850 hPa en la cuenta noratlántica, y las temperaturas de la superficie del mar en todo el Atlántico. El predictando en ambos casos es el campo de la temperatura del aire a 850 hPa en el Atlántico Norte. También se han llevado a cabo distintas pruebas atendiendo al filtrado preliminar de los datos. Los resultados del modelo muestran una “habilidad de predicción” útil en dos regiones: la primera en el subtropical y la segunda centrada en la Península Ibérica. En algunas de estas zonas se puede llegar a predecir hasta un 50% de la varianza del predictando original.

**Palabras clave:** Predictando, predictor, desfase, Descomposición en Valores Singulares, temperatura del aire, temperatura de la superficie del mar, habilidad de predicción, filtro temporal, validación cruzada, predicción, variabilidad acoplada, persistencia, periodo de entrenamiento.

### ABSTRACT

*In this work the patterns which represent the coupled features between two fields are used with forecast purposes. These patterns are obtained by the Singular Value Decomposition of the cross-covariance matrix between the predictor and predictand fields at the lag requested for the forecast. Several forecast experiments have been carried out, using three different predictor fields: the anomalies of the sea surface temperature and air temperature anomalies at 850hPa over the North Atlantic basin, and sea surface temperature anomalies over the whole Atlantic ocean. As predictand field the 850hPa air temperature anomalies over the North Atlantic has been considered. Also other experiments have been performed with three smoothing methodologies applied to the predictand and predictor fields. Results show good values of the skill of forecast over two regions: the first one centered over Bermuda (near the tropics) and the latter situated around Iberian Peninsula. In these zones it is possible to forecast up to about a 50% of the variance of the original predictand field.*

**Key words:** Predictand, predictor, lag, Singular Value Decomposition, air temperature, sea surface temperature, skill, temporal filtering, crossvalidation, forecast, coupled variability, persistence, training sample.

## 1. INTRODUCCIÓN

Las técnicas empírico-estadísticas orientadas a la predicción han sido aplicadas con éxito en numerosos trabajos. Hasta ahora, entre las más destacadas están el Análisis de la Correlación Canónica (ACC), Linear Inverse Modeling (LIM) y el Multichannel Singular Spectrum Analysis (MSSA). Éstas han dado buenos resultados para las zonas tropicales, como los trabajos de BARNSTON y ROPELEWSKI (1992), PENLAND y MAGORIAN (1993) entre otros, en los que han conseguido predecir el estado de El Niño-Oscilación del Sur. Lamentablemente los resultados de las predicciones empíricas no son tan buenos en las latitudes medias, como muestran los trabajos de BARNETT *et al.* (1984), BARNETT y PREISENDORFER (1987), SHABBAR y BARNSTON (1996), JOHANSSON *et al.* (1998) y VAUTARD *et al.* (1998). En estas latitudes debido a la variabilidad interna de la atmósfera, que se puede considerar como ruido blanco, y a la dificultad de establecer las interacciones atmósfera-océano, el estudio de la predecibilidad se complica enormemente.

Las técnicas empíricas ofrecen ventajas frente a las predicciones dinámicas, además del ahorro de tiempo de cómputo, el análisis estadístico permite simplificar las relaciones entre las variables climatológicas y elaborar los pronósticos con varios meses de antelación. Básicamente el método consiste en la elección de un predictor óptimo, es decir, una variable relacionada con el predictando (lo que se quiere predecir) y que presente el mismo comportamiento que éste pero precediéndolo en el tiempo.

En este trabajo la variable predictando son las anomalías de la temperatura del aire a 850 hPa (TA) y el campo predictor son las anomalías de la temperatura de la superficie del mar (TSM). El dominio espacial es el Atlántico Norte para ambas variables. Las anomalías de la TSM han sido consideradas en varios trabajos como buenos predictores, debido a la memoria o inercia del océano, ya que, frente a la atmósfera, su comportamiento puede ser considerado como un ruido rojo (asociado a mayores escalas temporales y espaciales). Siguiendo este razonamiento se puede pensar que hay escalas temporales atmosféricas (alta frecuencia) que no están presentes en la variabilidad oceánica, lo que sugiere que quizás sería necesario un predictor atmosférico, capaz de simular más eficientemente este comportamiento. Esta cuestión se tratará más adelante.

La técnica de predicción que aquí se emplea está basada en la Descomposición en Valores Singulares (DVS) de la matriz de covarianza cruzada entre el campo predictor y el predictando. Esta matriz se calcula desfasando un campo respecto a otro. La DVS, que no es más que una generalización del Análisis en Componentes Principales, está siendo muy empleada recientemente en el campo de la climatología, por su capacidad para encontrar modos acoplados entre dos campos, lo que la convierte en una útil herramienta para los estudios empírico-estadísticos de la predecibilidad.

Atendiendo al campo predictor, se han realizado tres experimentos: 1) con las anomalías de las TSM como predictor, 2) con el propio predictando como predictor de sí mismo (autopredicción) y 3) con las anomalías de las TSM en todo el Atlántico. La metodología también se ha variado, en un caso, se ha realizado una predicción pura, es decir, sin usar ninguna información del "futuro", en otro caso se ha llevado a cabo un experimento de validación cruzada, en el que se extrae de la muestra sólo el año que se quiere predecir, utilizando todos los demás en los cálculos. Por último,

se tratará el tema del filtrado temporal de los datos, previo a los cálculos de DVS, probando con tres tipos de filtros diferentes.

En el apartado 2, se describen los datos y el tratamiento de éstos; en la sección 3, se detalla el método de la DVS; la sección 4 describe la metodología empleada; los distintos experimentos realizados y los resultados obtenidos se presentan en el apartado 5. Por último, la sección 6 es un resumen del trabajo con sus conclusiones.

## 2. DATOS

El campo predictando está representado por las anomalías mensuales de la temperatura del aire (TA) a 850 hPa en el Atlántico Norte. Los datos provienen de los reanálisis del NCEP (KALNAY *et al.*, 1996), y han sido obtenidos de la página web del NOAA\_CIRES Climate Diagnostic Center ([www.cdc.noaa.gov](http://www.cdc.noaa.gov)). El fichero original abarca el periodo 1958-1998 (41 años) y presenta los datos diarios para distintos niveles atmosféricos. Se han calculado las anomalías mensuales, calculando primero las medias mensuales y después restando la climatología de todo el periodo. La elección del nivel de 850 hPa se debe a que la TA a este nivel es muy similar a la de superficie, y se hace menos sensible a los posibles efectos debidos a la topografía. El dominio espacial exacto seleccionado en este trabajo es 20°N-90°N, 90°W-10°E, abarca desde el este de América del Norte al oeste de Europa y desde el subtropical hasta las regiones polares.

Los datos de las temperaturas de la superficie del mar (TSM) (el campo predictor) provienen de otras fuentes distintas. A partir de los datos de COADS (WOODRUFF *et al.*, 1987) y los de IGOSS (REYNOLDS y SMITH, 1994) se ha construido un nuevo campo, mediante la fusión de los dos anteriores. Los COADS son datos mensuales de 1950 a 1993 dispuestos en una rejilla de 2°x2°. Los IGOSS son datos mensuales que provienen de satélite, de 1983 a 1999, su rejilla es 1°x1°. Debido los puntos de los COADS no coinciden con los de IGOSS, éste último ha sido interpolado a la red del primero. El fichero resultante de la unión de los anteriores abarca el periodo 1950-1999 y está en una rejilla 2°x2°. Las anomalías mensuales también se han calculado, restando la climatología de todo el periodo. El dominio espacial que se ha considerado para el campo predictor coincide con el del predictando. La elección del dominio espacial del predictor en el problema de las predicciones empíricas se basa en encontrar la región en la que el predictor tenga su mayor influencia sobre el predictando. En ocasiones es mejor escoger una región de escala global, asociada también a escalas temporales de menor frecuencia, y por tanto de mayor predecibilidad, pero existe el peligro de que haya puntos del dominio poco relacionados con el predictando y que puedan empeorar la predicción con la introducción de algún tipo de ruido.

Antes de elaborar las predicciones es necesario un suavizado de los datos. La justificación del filtrado se basa en un análisis espectral. Los campos predictor y predictando han sido promediados sobre 3 bandas latitudinales: la primera de 20°N a 35°N, la segunda de 35°N a 65°N y la tercera de 65°N a 90°N. A continuación se ha realizado un análisis espectral para cada banda. Las altas frecuencias están asociadas a la parte norte de la cuenca noratlántica, mientras que hacia los trópicos los fenómenos son de escala temporal mayor. En la figura 1 se muestran los espectros de las anomalías de la TA (1.a) y TSM (1.b) para la segunda banda. Se encuentran picos significativos

en los periodos cuasibienales (2.5-3.5 años), e intra-anales (8, 4.5 y 3 meses) en la atmósfera. También hay picos en las escalas decadales e inter-anales, pero no son significativos, o bien porque el intervalo de muestreo es demasiado corto (41 años), o bien debido los efectos del promedio por bandas. El espectro de las anomalías de la TSM se puede considerar prácticamente como un ruido rojo frente a la atmósfera. La necesidad de un suavizado previo a los datos está en los picos de 8, 4.5 y 3 meses, presentes en el espectro del aire y también en el del océano, pero en menor medida. Esta escala temporal no es tan importante en la banda 1, pero es necesario eliminarla para mejorar los pronósticos, aunque con ello se consiga predecir menos varianza del campo original.

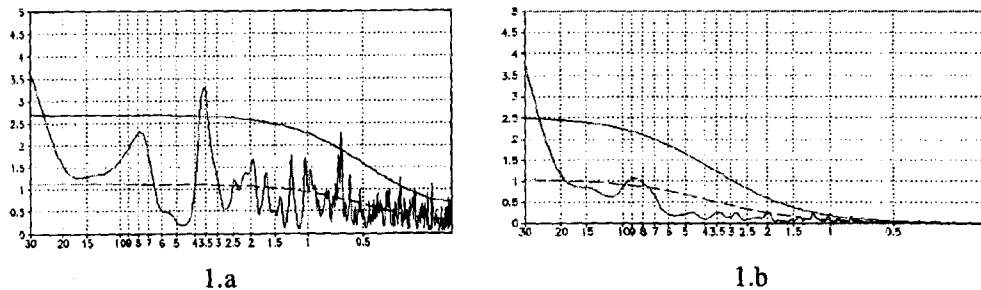


Fig. 1: Espectros de las anomalías de las TA (izquierda) y las TSM (derecha) para los campos promediados en la segunda banda latitudinal. Las curvas teóricas de ruido rojo y la del nivel de confianza al 95% están representadas.

Con el objeto de suavizar los datos, se han realizado tres experimentos con tres formas de filtrar diferentes. La primera de ellas consiste en un filtrado de media corrida de 8 meses ponderada con un coseno, en promedio la varianza del campo queda reducida a un 65%. La segunda es algo más complicada: en primer lugar se filtra el campo original con un filtro paso alto, reteniendo periodos menores a 9 meses. A continuación se calculan las Funciones Ortogonales Empíricas (FOEs) del campo filtrado y se proyecta el campo original en esos vectores, obteniendo así unos coeficientes temporales. Con los vectores y los coeficientes se reconstruye un nuevo campo, que es sustraído al campo original. Con este procedimiento se pretende eliminar del campo toda la información asociada a la alta frecuencia. El tercer experimento es más bien un estudio teórico, se trata de ajustar el campo original a un modelo autorregresivo de primer orden (AR(1)), para cada serie temporal, y quedarnos con la parte modelada de cada una.

En la figura 2 se ha representado la razón entre la varianza del campo predictando filtrado y sin filtrar con los tres suavizados. En el caso de la atmósfera se muestra claramente en las figuras 2.a y 2.b que las altas frecuencias están asociadas al norte del dominio, mientras que a medida que nos acercamos a las zonas subtropicales la razón se aproxima más a 1. Las zonas de alta frecuencia quedan principalmente localizadas en América del Norte y Groelandia. Las frecuencias mayores de 8 meses y mayores no son tan importantes en el océano. La figura 2.c muestra en qué lugares del dominio los campos se aproximan más a un proceso AR(1) puro. Así la atmósfera en las latitudes medias en general queda mal representada con este tipo de modelo, excepto en la zona sub-

tropical.

Conviene hacer notar que, en el primero de los métodos de filtrado, se ha cuidado de no incluir ninguna información del “futuro”. Para cada predicción el conjunto de datos se divide en dos periodos, el de entrenamiento y el de validación. El primero de ellos es el que se utiliza para los cálculos, el segundo es el que se predice. El filtrado de la serie se aplica sólo al primer periodo, pero esta cuestión se tratará con mayor detenimiento en la sección 4.

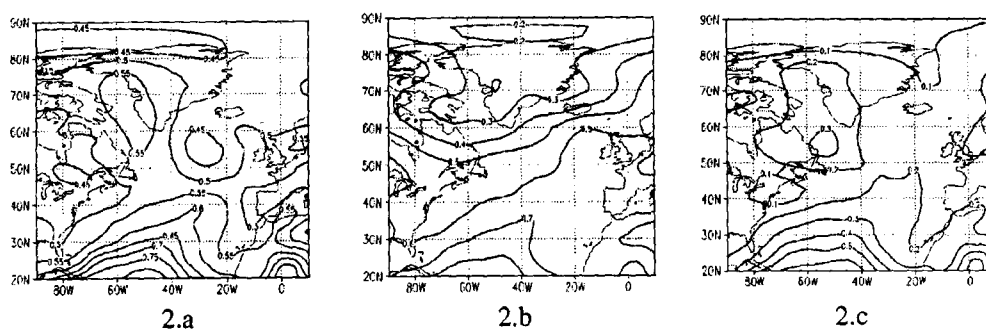


Fig.2: Razón entre las varianzas (campo filtrado/campo no filtrado) de las anomalías de la temperatura del aire para los tres tipos de suavizado.

### 3. ESQUEMA ESTADÍSTICO: DVS APLICADA A LA PREDICCIÓN

La DVS es una técnica estadística lineal que encuentra la variabilidad acoplada entre dos campos diferentes. La DVS junto con el ACC se usa ampliamente en el campo de la climatología. Detalles acerca de esta técnica y sus aplicaciones se pueden encontrar en BRETHERTON *et al.* (1992) y NAVARRA (1993).

Sean  $X_j$  e  $Y_j$  los valores del predictando y predictor en el punto  $j$ , para el intervalo temporal  $t$  ( $t=1, \dots, N_t$ ), entonces la DVS de la matriz de covarianza cruzada ( $C$ ) entre estos dos campos se puede escribir de la siguiente manera:

$$c_{ij} = \sum_{t=l+1}^{N_t} y_i(t-l)x_j(t) = \sum_{k=1}^p \sigma_k u_{ik} v_{kj}$$

donde  $l$  es el retraso temporal del predictando respecto del predictor,  $u_k$  y  $v_k$  los vectores singulares del predictando y predictor respectivamente, y  $\sigma_k$  los valores singulares asociados a cada modo singular  $k$ . Para un determinado modo  $k$ , la evolución temporal de  $u_k$  y  $v_k$  viene dada por la proyección de los campos sobre los vectores, resultando en unos coeficientes,  $a_k$  y  $b_k$ :

$$a_k(t-l) = \sum_{j=1}^p x_j(t-l) u_{kj}$$

$$b_k(t) = \sum_{j=1}^p y_j(t) v_{kj}$$

El esquema predictivo se basa en que teóricamente los patrones espaciales del predictando  $u_k$  representan la situación de los del predictor  $v_k$ , pero  $l$  meses más tarde. Por ello los coeficientes  $b_k$  preceden en el tiempo a los  $a_k$ . La evolución en el tiempo de estos coeficientes está relacionada linealmente por unos coeficientes obtenidos por la estimación de los mínimos cuadrados:

$$a_k(t) = c_k b_k(t-l)$$

y por lo tanto, si el predictando se puede reconstruir a partir de  $u_k$  y  $a_k$ :

$$x_j(t) = \sum_{k=1}^p a_k(t) u_{kj}$$

el campo predicho  $X_m$  se puede obtener:

$$x_{mj}(t) = \sum_{k=1}^p c_k b_k(t-l) u_{kj}$$

#### 4. METODOLOGÍA

Entre las metodologías más empleadas en las predicciones empíricas se destacan tres. La primera de ellas es la validación cruzada, que consiste en eliminar del conjunto de datos el año que se quiere predecir, efectuar todos los cálculos con el resto, y finalmente elaborar la predicción de ese año. El segundo método consiste en dividir los campos en dos periodos, el de entrenamiento (el que se usa para los cálculos) y el de validación, que es el que se predice. Con este segundo método, las predicciones que se obtienen son peores, pero más "puras" en el sentido de que no se usa información del futuro como en la validación cruzada. El tercer método es el de la retropredicción, en el cual los puntos del futuro pueden ser empleados en la predicción del pasado. En este trabajo se ha empleado la segunda metodología, aunque se ha contrastado con la validación cruzada, para comparar resultados.

Nuestro esquema de predicción se resume como sigue: las predicciones se van a realizar desde enero de 1970 hasta diciembre de 1998. Para predecir un determinado mes, se usan los 12 años anteriores a él en los cálculos de la matriz de covarianza cruzada. Los 12 años componen el periodo de entrenamiento y el punto que se predice es el periodo de validación. Entre el periodo de entrenamiento y la predicción hay un intervalo de tiempo, que indica la anticipación con la que se predice. Por ejemplo, si se quiere predecir enero de 1970 con 3 meses de antelación, se empleará en los cálculos el periodo que va desde enero de 1950 hasta septiembre de 1969. El mismo periodo pero adelantado 3 meses se considera en el campo predictor. Una vez que se ha realizado el

pronóstico de un mes, se pasa al siguiente, y para ello el periodo de entrenamiento se desplaza también. El número de meses del periodo de entrenamiento se ha escogido después de realizar numerosas pruebas de predicción y concluir en que el periodo óptimo es de 12 años, a pesar de que parezca un intervalo demasiado corto para una estadística fiable. Los 12 años son filtrados en el tiempo sin utilizar meses del futuro, es decir, en la media corrida estos puntos quedan excluidos de la suma.

Existen muchos criterios para determinar el número de modos singulares que se escogen en la reconstrucción de los campos  $X$  y  $X_m$ . En este estudio se ha truncado el desarrollo en serie en los 15 primeros patrones ( $p=15$ ). Con este número se llega a reconstruir un 80% de la varianza total del campo predictando. Este paso equivale a un segundo filtrado de los datos.

Una vez que se han elaborado los pronósticos para las tres décadas, la bondad de los resultados se pueden medir de diferentes maneras: correlación entre el campo predicho y el observado, error cuadrático medio, los coeficientes del error lineal en el espacio de probabilidad (LEPS scores), etc. La medida de la "habilidad" de predicción ( $S$ ) que se emplea en este trabajo es la correlación entre  $X$  y  $X_m$  para el intervalo de predicción. La significación estadística de las correlaciones se ha determinado con la aproximación para muestras con un número considerable de puntos.  $S$  es significativa al 95% de confianza si  $S > 2/T^{1/2}$ , donde  $T$  es el número de puntos con el que se calcula la correlación. Si el periodo de predicción lo componen 348 meses, el límite de la significación se encuentra por encima de 0.1.

## 5. EXPERIMENTOS Y RESULTADOS

Se han llevado a cabo tres experimentos de predicción atendiendo a los métodos de filtrado que se han descrito en la sección 2. La "habilidad" de predicción para todo el periodo (1970-1998) ha

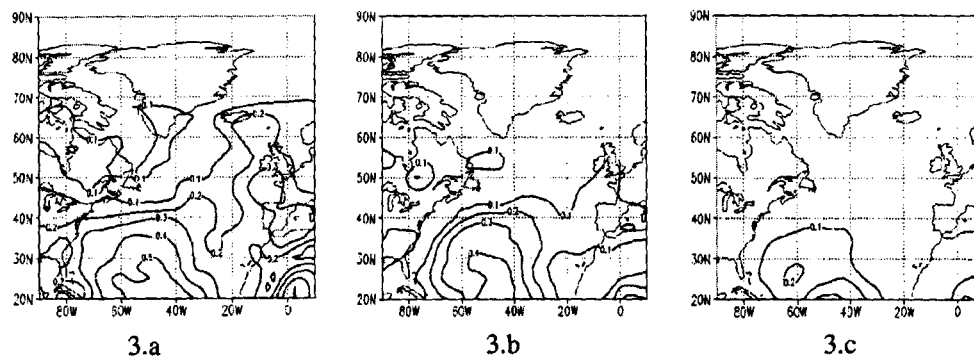


Fig. 3: Valores de  $S$  multiplicados por la razón entre las varianzas de la figura 2 y por 0.8 (por retener 15 patrones en el desarrollo). Los números de la figura 3 multiplicados por 100 representan el porcentaje de varianza del campo original que se ha conseguido predecir. 3.a representa el primer tipo de suavizado, 3.b al segundo y 3.c a la aproximación de la serie a un proceso AR(1).

sido calculada. El desfase entre predictor y predictando al igual que el tiempo de antelación se ha variado de 3 a 18 meses.

La figura 3 muestra los valores de  $S$  para una predicción elaborada con una anticipación de 9 meses. Estos valores han sido multiplicados por la razón entre las varianzas de la figura 2, y por 0.8 (que es la fracción de la varianza que se ha retenido en la reconstrucción de los campos al quedarnos con 15 patrones). Este número es una medida del porcentaje de varianza que se ha conseguido predecir en cada punto. Como medida de control se han considerado los valores de la correlación esperados con la predicción obtenida con la persistencia (no se ha mostrado). A 9 meses, la técnica no mejora mucho frente a la persistencia, pero esto no ocurre si se predice con una antelación de 3 ó 6 meses.

En vista de los valores representados en la figura 3, podemos concluir que las mejores predicciones se obtienen con el primer método de filtrado, mientras que las peores, con el modelado del campo a un proceso AR(1). En algunas zonas situadas en el sur del dominio, se puede llegar a predecir más de un 50% de la varianza del campo original, en otras zonas como en la Península Ibérica solo se predice alrededor de un 30% de la varianza del campo original.

La variación de la calidad de las predicciones en función de la anticipación con la que se predice se presenta en la figura 4, donde se han promediado los valores de  $S$  para cuatro zonas del dominio y se han representado en función de los meses de anticipación. Las cuatro zonas se han seleccionado de acuerdo a los mejores valores de  $S$ , valores intermedios y valores más bajos. Así se ha designado zona I ( $75^{\circ}\text{W}-30^{\circ}\text{W}$ ,  $20^{\circ}\text{N}-40^{\circ}\text{N}$ ), zona II ( $30^{\circ}\text{W}-0^{\circ}\text{W}$ ,  $35^{\circ}\text{N}-50^{\circ}\text{N}$ ), zona III ( $25^{\circ}\text{W}-10^{\circ}\text{W}$ ,  $60^{\circ}\text{N}-70^{\circ}\text{N}$ ) y zona IV ( $85^{\circ}\text{W}-55^{\circ}\text{W}$ ,  $45^{\circ}\text{N}-65^{\circ}\text{N}$ ). La zona I (centrada en el subtrópico y Bermuda) presenta los valores de  $S$  más aceptables, seguida de la zona II (centrada en la Península Ibérica). Los valores de  $S$  en la zona III, situada en las proximidades de Islandia, y en la zona IV, situada en las inmediaciones de la Península del Labrador, son muy bajos para los cuatro casos. Como es de esperar, las predicciones empeoran a medida que aumenta la anticipación (y el desfase entre predictor y predictando). Así los valores de  $S$  disminuyen claramente de 3 a 18 meses.

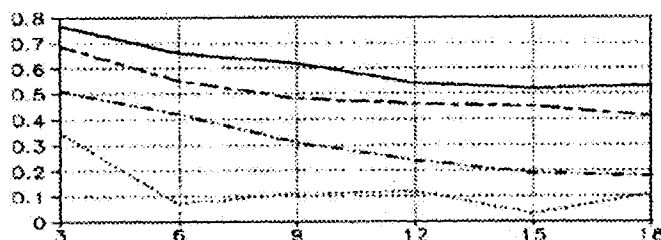


Fig. 4: Variación de la calidad de las predicciones medida por  $S$  en función del tiempo de antelación con el que se predice (en meses). Zona I (línea continua), zona II (punto-rama), zona III (dos puntos-rama) y zona IV (línea punteada).

Se han realizado otros experimentos con los campos filtrados con la media corrida de 8 meses. En primer lugar se ha desarrollado el método de la validación cruzada, comentado en la sección anterior. Los resultados se muestran en la figura 5.a para una predicción con un desfase de 9 meses



entre predictor y predictando. Los valores de  $S$  han sido multiplicados por la razón entre las varianzas de la figura 2.a y por el factor 0.8. Las predicciones mejoran bastante con esta técnica, ya que usan tanto información del pasado como del futuro para predecir.

Con el objeto de mejorar los resultados en el norte del dominio, se ha considerado la posibilidad de incluir un predictor atmosférico por su capacidad de reproducir la variabilidad de alta frecuencia mejor que el océano. Con este propósito se han empleado las anomalías de las TA como predictoras de sí mismas (autopredicción). En la figura 5.b está representado el experimento de autopredicción con 9 meses de antelación. Es de suponer que las cortas escalas temporales asociadas con el norte del dominio puedan ser representadas con más éxito por un campo atmosférico en lugar de uno oceánico. La figura 5.b en comparación con la figura 3.a muestra que esto no ocurre así, ya que las predicciones no han conseguido ser mejoradas.

El último de los experimentos se basa en la posible influencia del océano Atlántico Tropical en las latitudes superiores. A fin de investigar esta cuestión, se ha construido un nuevo campo predictor con la ampliación de las anomalías de las TSM a todo el océano Atlántico. La figura 5.c muestra los valores de  $S$  multiplicados por los valores de la figura 2.a y por 0.8. Los resultados se han mostrado para la predicción con 9 meses de antelación. Al comparar las figuras 3.a y 5.c, se llega a la conclusión de que los resultados no han mejorado significativamente, incluso en algunas zonas como el norte de las islas Británicas o la Península Ibérica los valores de  $S$  son más bajos. El empeoramiento de la "habilidad" de predicción es un indicativo de que el Atlántico Tropical no ejerce apenas influencia y por tanto no es un buen predictor de las situaciones anómalas presentes en estas regiones.

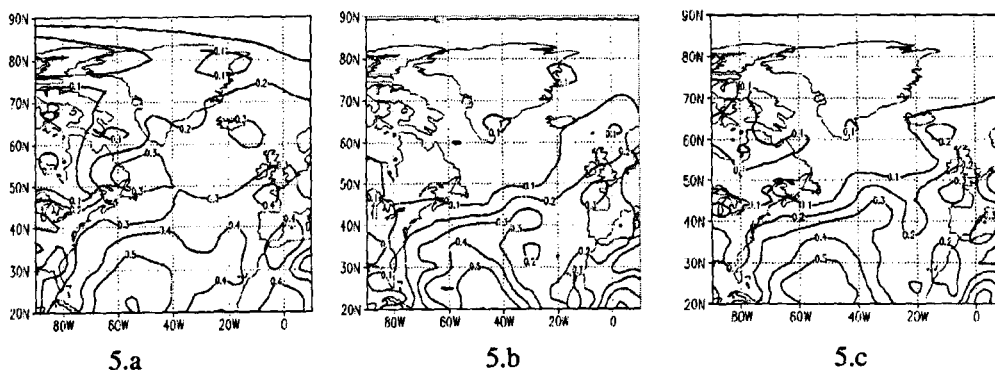


Fig. 5: Valores de  $S$  multiplicados por la razón entre las varianzas de la figura 2.a y por 0.8 (por retener 15 patrones en el desarrollo). Los números de la figura 5 multiplicados por 100 representan el porcentaje de varianza del campo original que se puede predecir. 5.a corresponde al experimento de validación cruzada, 5.b al de autopredicción y 5.c al considerar las TSM de todo el Atlántico como predictor.

## 6. RESUMEN Y CONCLUSIONES

En este trabajo se ha presentado una aplicación de los patrones de variabilidad acoplada entre dos campos a la predicción empírica. La técnica se basa en la Descomposición en Valores Singulares (DVS) de la matriz de covarianza cruzada entre el campo predictor y el campo que se quiere predecir o predictando. Esta matriz se calcula a un determinado desfase entre estos dos campos. La DVS es una técnica que encuentra aquellas combinaciones lineales de un campo que estén linealmente relacionadas con combinaciones lineales de otro campo, con la condición de que además sean capaces de reproducir la mayor porción de la varianza de éstos. El esquema de predicción desarrollado en este estudio es muy sencillo y se basa en la elección de un periodo de entrenamiento de 12 meses para la predicción de un determinado mes. Las predicciones se han elaborado para casi tres décadas de enero de 1970 a diciembre de 1998. Como campo predictor se han escogido las anomalías de las temperaturas de la superficie del mar en el Atlántico Norte y como variable predictando, las anomalías de la temperatura del aire a 850hPa en el mismo dominio espacial. Antes de la aplicación de la DVS es necesario efectuar un suavizado de los datos, justificado por un estudio espectral. Se ha experimentado con tres procedimientos de filtrar diferentes y con el que mejor resultados se obtiene es un sencillo filtro de media corrida de 8 meses. La razón entre las varianzas campo filtrado/campo no filtrado, representada en la figura 2, indica en qué regiones unas escalas temporales son dominantes frente a otras. Así la alta frecuencia, asociada a una predecibilidad baja, se encuentra localizada en el norte dominio.

Una vez elaborados los pronósticos, se calcula la “habilidad” de predicción  $S$  como la correlación entre el campo predicho y el campo observado. Los valores de  $S$  son elevados en la mitad sur del dominio, especialmente en dos zonas, una situada en la zona de calmas cerca del Trópico, y la otra centrada en la Península Ibérica. Los valores sobre Islandia aunque significativos, no son demasiado elevados. Los peores valores de  $S$  se encuentran en la zona del Labrador y Groelandia.

Con el fin de simular con más acierto los fenómenos atmosféricos de alta frecuencia y mejorar así los valores de  $S$ , se ha probado con un predictor atmosférico. Para ello las anomalías de las TA han sido empleadas también como predictoras de sí mismas. Los resultados muestran que tampoco la atmósfera es capaz de predecir las escalas de tan alta frecuencia.

Por último, para estudiar la capacidad del Atlántico Tropical como predictor de la temperatura del aire en latitudes superiores, se ha ampliado el dominio del predictor original a todo el Atlántico. Las predicciones no llegan a mejorar significativamente, incluso en algunas regiones  $S$  disminuye claramente.

Según las estructuras espaciales que presentan las figuras 2, 3 y 5 es evidente que la predecibilidad está ligada a los fenómenos de baja frecuencia. La técnica tiene éxito cuando las interacciones entre predictando y predictor tienen una gran componente lineal. La mitad norte del dominio, caracterizada por una gran inestabilidad baroclina, asociada a la variabilidad interna de la atmósfera y de componente fuertemente no lineal, no se predice eficientemente con esta técnica. El análisis espectral efectuado a puntos de estas regiones, muestra que su evolución temporal puede ser considerada prácticamente como un ruido blanco, con un tiempo de decorrelación muy corto. El

problema de la predicción se complica enormemente en estas situaciones. De todas maneras, otras posibles variaciones a la metodología que sean capaces de mejorar los resultados están siendo consideradas.

## 7. AGRADECIMIENTOS

Los autores desean agradecer a César Gutiérrez por su colaboración al principio de este trabajo, y al resto del Grupo de Física de Clima de la Universidad de Alcalá: Montserrat Seisdedos, Antonio Ruíz de Elvira, Francisco Álvarez y William Cabos.

También se quiere agradecer a Christian Olalla García y Manuel Prieto Mateo su ayuda en los aspectos técnicos.

El trabajo ha sido financiado por el Programa Medio Ambiente y Clima de la Unión Europea, contrato ENV4-CT98-0714 (SINTEX).

## 8. REFERENCIAS

- BARNETT, T.P., HEINZ, H.D. y HASSELMANN, K. (1984): "Statistical prediction of seasonal air temperature over Eurasia". *Tellus*, 36A, 132-146.
- BARNETT, T.P. Y PREISENDORFER, R. (1987): "Origins and levels of monthly seasonal skill for United States surface air temperature determined by Canonical Correlation Analysis". *Mon. Wea. Rev.*, 9, 1825-1850.
- BARNSTON, A.G. Y ROPELEWSKI, C.F. (1992): "Prediction of ENSO episodes using Canonical Correlation Analysis". *J. Climate*, 5, 1316-1345.
- BRETHERTON, C.S., SMITH, C. Y WALLACE J.M. (1991): "An intercomparison of methods for finding coupled patterns in climate data". *J. Climate*, 5, 541-560.
- JOHANSSON, A., BARNSTON, A., SAHA, A. Y VAN DER DOOL, H. (1998): "On the level and origin of seasonal variability forecast skill in Northern Europe". *J. Atmos. Sci.*, 55, 103-127.
- KALNAY, E., *et al.* (1996): "The NCEP-NCAR Reanalysis Project". *Bull. Am. Meteor. Soc.*, 77, 437-471.
- NAVARRA, A. (1993): "A new set of Orthonormal Modes for Linearized Meteorological Problems". *J. Atmos. Sci.*, 50, 2569-2583.
- PENLAND, C. Y MAGORIAN, T. (1993): "Prediction of Niño3 sea surface temperature anomalies using Linear Inverse Modeling". *J. Climate*, 6, 1067-1076.
- REYNOLDS, R.W., Y SMITH, T.M. (1994): "Improved sea surface temperature analysis". *J. Climate*, 7, 929-948.
- SHABBAR, A. y BARNSTON, A.G. (1996): "Skill of seasonal climate forecasts in Canada using Canonical Correlation Analysis". *Mon. Wea. Rev.*, 124, 2370-2385.
- VAUTARD, R., PLAUT, G., WAMG, R. Y G. BRUNET (1998): "Seasonal prediction of North America surface air temperature using space-time principal components". *J. Climate*, 10, 389-394.

WOODRUFF, S.D., SLUTZ R.J., JENNE, R.L. Y STEURER, P.M. (1987): "A comprehensive atmosphere-ocean data set". *Bull. Amer. Met. Soc.*, 68, 1239-1250.