

PREDICCIÓN ESTACIONAL DE CONDICIONES ANÓMALAS EN EL ATLÁNTICO NORTE

Emilia SÁNCHEZ GÓMEZ y María José ORTIZ BEVIÁ

Departamento de Física, Universidad de Alcalá

RESUMEN

El objetivo de este trabajo es desarrollar un esquema de predicción empírico que pueda superar los niveles de la habilidad de predicción obtenidos con el modelo de la persistencia. Por este motivo, se ha desarrollado un modelo estadístico de predicción, que tiene como núcleo la Descomposición en Valores Singulares de una matriz. Esta técnica produce pares de patrones que explican la máxima covarianza entre dos campos. Como campo predictor se han escogido las anomalías de la superficie del mar y como campo predictando, las anomalías de la temperatura del aire al nivel de 850 hPa, ambos en la cuenca Noratlántica. El esquema de predicción opera en campos estacionales, es decir el predictor y el predictando se dividen en cuatro conjuntos, atendiendo al ciclo estacional. En el caso del campo predictor, cada conjunto se denomina TOR. Cada TOR predice por separado una estación en el predictando, al variar el intervalo de antelación requerido en la predicción. Los resultados de los pronósticos muestran que la habilidad de predicción es buena en dos regiones del dominio, una región cercana a Bermuda y otra en la Península Ibérica. La habilidad de predicción también presenta una gran dependencia con la estación predictanda, siendo en el verano cuando se alcanzan los mejores pronósticos.

Palabras clave: Campos estacionales, predicción a largo plazo, modelo empírico, atmósfera, océano.

ABSTRACT

The main goal of this work is the developing of an empirical forecast scheme capable of beating the skill levels obtained by the persistence model. We have built a statistical forecast model which has as its core the Singular Value Decomposition of a matrix. This technique yields pairs of patterns that explain the maximum covariance between two fields. We have chosen as predictor field the sea surface temperature anomalies and as predictand field the air temperature anomalies at 850 hPa, both in the North Atlantic basin. The forecasts scheme works on seasonal fields. In this case both, the predictor and predictand fields are divided into four sets, according to the seasonal cycle. For the predictor field, each of the sets is named TOR. Each TOR forecasts separately one predictand season, as the time lead increases. The results show the highest skill levels at two regions: one around Bermuda and the other over the Iberian Peninsula. The skill of forecasts has a strong seasonal dependence. Summer forecasts turn out to be the most successful.

Key words: Seasonal fields, long-term forecast, empirical model, atmosphere, ocean.

1. INTRODUCCIÓN

Las predicciones a largo plazo, especialmente aquéllas capaces de predecir la evolución de la atmósfera con una antelación de dos o tres estaciones, constityen una de las principales motivaciones para la Física del Clima. Las fluctuaciones climáticas relacionadas con los fenómenos de baja frecuencia son muy difíciles de modelizar. Sin embargo, en los últimos años, ha habido una mejora significativa de los Modelos de Circulación General (GCMs), aunque a pesar de todos los esfuerzos, los GCMs no pueden predecir con un éxito razonable la evolución del clima con una antelación de dos estaciones. Una alternativa que se presenta frente a los CGMs, son los modelos empíricos. Estos modelos, no son tan caros desde el punto de vista del coste computacional y además, permiten ahondar en las relaciones físicas de las variables implicadas (SHABBAR y BARNSTON, 1996). Los modelos empíricos de predicción han dado buenos resultados en las latitudes tropicales (PENLAND y MAGORIAN, 1993; PENLAND y MATROSOVA, 1998; RUIZ DE ELVIRA *et al.* (2000)). Sin embargo, en las latitudes medias, los resultados no son tan notables (VAUTARD *et al.*, 1998; SÁNCHEZ GÓMEZ *et al.*, 1999, 2001).

Las predicciones empíricas que se van a presentar en este trabajo formaron parte de los objetivos del proyecto europeo conocido como PROVOST y están encaminadas a conseguir un modelo empírico que a la par de sencillo sea capaz de superar las predicciones obtenidas con el modelo de la persistencia. En las primeras etapas del trabajo (SÁNCHEZ GÓMEZ *et al.*, 2001, 2002) se trabajó con campos filtrados. Ambos campos, predictor y predictando se filtraban en el tiempo para eliminar variabilidad de alta frecuencia asociada a una baja predecibilidad. Los resultados mostraron unos niveles bastantes aceptables de la habilidad de predicción, sin embargo no se consiguió una mejora significativa frente a la persistencia. En este trabajo se presenta una novedad que se ha introducido en el esquema de predicción. Antes de elaborar los pronósticos, ambos campos, el predictor y el predictando, se dividen en cuatro conjuntos atendiendo al ciclo estacional. En el caso del predictor, cada grupo se denomina TOR. Esta metodología es muy similar a la que presentaron JOHANSSON *et al.* (1998) en sus pronósticos empíricos. Veremos que con este nuevo esquema, se consigue una mejora muy importante frente a la persistencia.

El modelo empírico del cual derivan nuestras predicciones tiene como núcleo la propiedad matemática conocida como Descomposición en Valores Singulares (SVD) de una matriz. Esta técnica puede aislar pares de patrones de los campos predictor y predictando de tal modo que la covarianza entre ellos sea máxima. Aunque esta técnica ya se había empleado previamente por los climatólogos como instrumento de análisis, se implementó por primera vez como técnica predictiva en SÁNCHEZ GÓMEZ *et al.* (1999).

La elección del campo predictando se basa en dos factores: por un lado las predicciones tienen que ser relevantes para el clima en el Atlántico Norte. Se escogió la temperatura del aire a 850 hPa (T850) porque es una variable que está muy bien representada en los reanálisis de NCEP/NCAR y además está conectada con la temperatura en superficie. Por otro lado, esta fue la variable que se usó en todas las integraciones del proyecto europeo PROVOST. La elección del campo predictor se basa en la noción de que la predecibilidad de la atmósfera reside principalmente en aquellos forzamientos de los componentes del sistema climático que varían en escalas temporales largas, como es el caso del océano (BRANKOVIĆ *et al.*, 1994). Por esta razón se han escogido las temperaturas de la superficie del mar (SST) como campo predictor.

El trabajo se distribuye de la siguiente manera: en la sección 2 se presentan los datos, en la sección 3 se describe brevemente la metodología y el esquema estacional, en la sección 4 se presentan los resultados y por último, las conclusiones y el resumen se encuentran en la sección 5.

2. DATOS

Los valores del campo predictando, las anomalías de T850, se han obtenido de los reanálisis de NCEP/NCAR. A partir de los datos diarios, se han calculado las medias mensuales. El periodo que se ha seleccionado es 1948-2000 y la climatología de este periodo se ha restado a los datos mensuales para obtener las anomalías. Los valores de la variable se distribuyen en una rejilla $2.5^{\circ} \times 2.5^{\circ}$.

La capacidad del océano para predecir la T850 se ha evaluado mediante el uso de las anomalías de SST como campo predictor. Los datos de SST se han obtenido de la fusión de los datos de COADS (1950-1982) y los datos de IGOSS (1982-2000). Las anomalías mensuales se han obtenido de la misma manera que las del campo predictando. Estos datos se encuentran distribuidos en una rejilla $2^{\circ} \times 2^{\circ}$.

El dominio espacial de ambos campos es el Atlántico Norte: de $90^{\circ}W$ a $10^{\circ}E$, y de $20^{\circ}N$ a $90^{\circ}N$.

3. MÉTODO EMPÍRICO DE PREDICCIÓN

Si se quiere predecir a l meses de antelación, se calcula la matriz de covarianza cruzada C entre el campo predictando $z(t)$ y el campo predictor $y(t)$:

$$C = \langle z(t)y^T(t-l) \rangle$$

La SVD de la matriz C se puede escribir:

$$C = UWV^T$$

donde \mathbf{u}_k y \mathbf{v}_k son los vectores singulares asociados al campo predictando y predictor respectivamente, y w_k son los valores singulares de la matriz C . Estos tres elementos forman el k -ésimo modo singular.

La SVD encuentra pares de patrones de tal manera que la covarianza entre ellos es máxima. Si se proyectan los campos originales en los vectores singulares asociados a a cada campo se obtienen unos coeficientes temporales $a_k(t)$ y $b_k(t)$ del campo predictando y predictor respectivamente. Se cumple entonces la siguiente propiedad:

$$\langle a_1(t)b_1(t-l) \rangle = u_1 C v_1^T = \max = w_1$$

La SVD garantiza que la elección de \mathbf{u}_1 y \mathbf{v}_1 se hace de tal forma que la covarianza entre los coeficientes temporales es máxima y además viene dada por el valor singular asociado.

Los coeficientes temporales se pueden relacionar a través de otros coeficientes que se estiman mediante los mínimos cuadrados. Si el predictando $z(t)$ se reconstruye utilizando los coeficientes

$a_k(t)$ ponderados por sus vectores singulares:

$$z(t) = \sum_{k=1}^q a_k u_k$$

La predicción de $z(t)$ en el instante t_p vendrá dada por:

$$\hat{z}(t_p) = \sum_{k=1}^q c_k b_k(t_p - l) u_k$$

donde los coeficientes c_k se han estimado mediante los mínimos cuadrados.

El desarrollo se ha truncado en el q -ésimo modo singular. De esta manera sólo se retienen los modos singulares más relevantes, el resto se eliminan porque son considerados como ruido. El criterio para escoger el número de términos a retener se basa en la fracción de covarianza explicada por los primeros modos singulares.

En las predicciones empíricas, los campos predictando y predictor se dividen en periodo de entrenamiento y periodo de validación. Los parámetros del modelo son estimados a partir del periodo de entrenamiento. El periodo de validación es el intervalo temporal que se predice y que sirve para evaluar los resultados. El procedimiento que aquí seguimos se ha diseñado para producir auténticas predicciones, es decir, sólo se emplea información del “pasado” para predecir, no se aplica el método de la validación cruzada (SÁNCHEZ GÓMEZ y ORTIZ BEVIÁ, 2001).

3.1. Esquema estacional

Como se ha descrito en la introducción, ambos campos, el predictando y el predictor se dividen en cuatro conjuntos atendiendo al ciclo estacional. Cada predictor estacional se denomina TOR, de esta manera tenemos TOR1 (diciembre, enero, febrero), TOR2 (marzo, abril y mayo), TOR3 (junio, julio y agosto) y TOR4 (septiembre, noviembre y diciembre). Es importante hacer notar que no se realizan promedios estacionales, son datos mensuales. Cada uno de los TORs predice las cuatro estaciones predictandas, de acuerdo con la variación del intervalo de anticipación con el que se predice, como se ilustra en la figura 1.

Los experimentos de predicción se repiten sistemáticamente para los cuatro TORs. El uso de campos estacionales reduce drásticamente la longitud del periodo de entrenamiento, por esta razón se ha optado por usar toda la información previa al instante que se quiere predecir.

a) Significación estadística

El número de patrones singulares que se usan en la reconstrucción del predictando y en la elaboración de los pronósticos oscila entre 15 y 20. Con estos patrones, se puede reproducir en torno a un 90 % de la varianza total de los campos. Cuando el número de términos retenido no es pequeño, pueden aparecer problemas derivados de lo que se conoce como habilidad de predicción artificial. Esto se debe a que, a medida que se avanza en el orden del patrón singular, las restricciones aumentan, de tal modo que pueden aparecer patrones singulares que sean fruto de la matemática y no de procesos físicos reales. Para abordar este problema, lo que se suele hacer es generar un gran

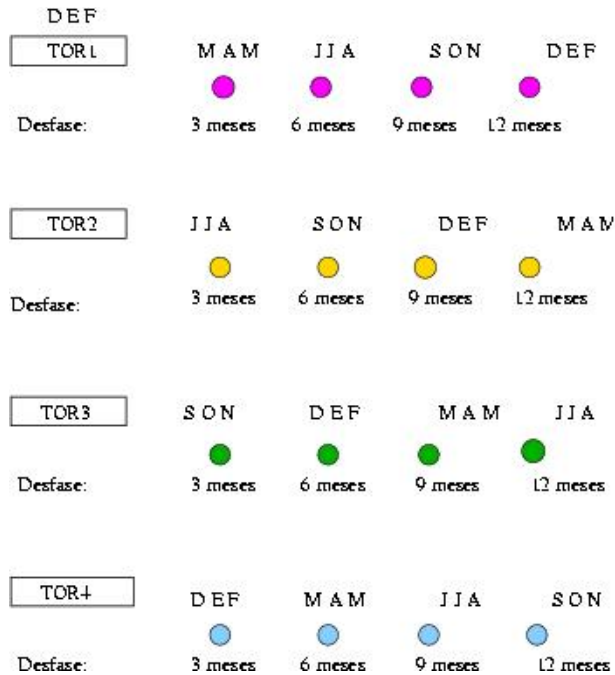


Figura 1: La figura ilustra el esquema estacional de predicción. El campo predictor se divide en cuatro predictores (TORs) según la estación, cada TOR predice todas las estaciones del predictando.

número campos sintéticos con una gran componente aleatoria. El esquema de predicción se aplica sobre estos campos y se compara con los resultados del campo original. La generación de un gran número de campos predictor y predictando sintéticos permite obtener una función de distribución empírica de la habilidad de predicción, y por tanto se puede realizar una prueba estadística. La técnica que hemos empleado para la creación de campos sintéticos se basa en el *bootstrap* (EFRON y TIBSHIRANI, 1993). Los campos predictor y predictando se dividen en bloques de la misma longitud, entonces los campos sintéticos se crean mediante la reorganización aleatoria de los bloques.

b) Persistencia como medida de control

En todos los estudios de predicciones empíricas, la persistencia es usada como una referencia básica. Los pronósticos usados con tu modelo se tienen que comparar con las obtenidas si se asume una persistencia de las anomalías. La persistencia no es más que la autocorrelación del campo predictando al desfase requerido para la predicción.

Tabla 1: SCF1 obtenida en los cálculos de SVD realizados para cada uno de los cuatro TORs predictores.

	3 meses	6 meses	9 meses	12 meses
TOR1	22	26	23	21
TOR2	26	26	23	20
TOR3	25	22	22	24
TOR4	24	23	26	24

4. RESULTADOS

Se han realizado pronósticos estacionales de las anomalías de T850 con diferentes intervalos de anticipación (de 3 a 12 meses). Antes de realizar las predicciones, se han llevado a cabo varios estudios para determinar cuando es más fuerte el acoplamiento lineal entre el campo atmosférico y el campo oceánico.

a) Dependencia estacional de la covarianza entre los campos de anomalías de T850 y SST

El acoplamiento lineal entre los campos predictor y predictando viene dado por los valores singulares de la matriz de covarianza cruzada. En la Tabla 1 se presenta la fracción de covarianza asociada (SCF1) al primer modo singular. Este es el modo que más peso va a tener en la elaboración de los pronósticos. Se han realizado diferentes cálculos de SVD para cada uno de los cuatro TORs y las cuatro estaciones del predictando.

Los valores de la tabla 1 indican que independientemente del intervalo de antelación, el valor de la SCF1 es máximo cuando la estación a predecir es el verano y el otoño. Este hecho apunta a que existe una predecibilidad potencial mayor en verano y otoño.

b) Habilidad de predicción de las anomalías de T850

La habilidad de predicción (H), calculada como la correlación entre el campo observado y el predicho, para el TOR1 de la SST se ha representado en la Figura 2. El periodo de validación es 1970-2000. Las regiones sombreadas indican donde los valores de la habilidad de predicción se encuentran por debajo de los estimados para la habilidad de predicción artificial, con un 95 %.

Los valores de H obtenidos en las regiones del sur son muy similares a los que se obtenían con el esquema de campos filtrados (SÁNCHEZ GÓMEZ *et al.*, 2001). Aunque las predicciones de campos filtrados eran mejores en las zonas continentales, como en la Península Ibérica. Sin embargo, el esquema estacional logra una significativa mejoría cuando comparamos con la persistencia. La Figura 3 representa la diferencia entre la habilidad de predicción del modelo y la obtenida con la persistencia. Vemos como la diferencia es positiva (mejor el modelo) en todas las regiones.

Claramente la calidad de las predicciones empeora a medida que aumenta el intervalo de anticipación con el que se realizan.

c) Dependencia estacional de la habilidad de predicción

En la Figura 2 es fácilmente apreciable que los niveles de H aumentan considerablemente cuando la estación predictanda es el verano. Le sigue el otoño. Este comportamiento de H se manifiesta para los experimentos realizados con el resto de los TORs. Independientemente del TOR predictor, hay una subida considerable de H en el verano. El otoño presenta un máximo secundario.

d) Orígenes de la habilidad de predicción

La estructura espacial de H claramente depende de la estación del predictando que se quiere predecir. Sin embargo, los valores más altos de H siempre están localizados en el sur del dominio y en la Península Ibérica para todos los experimentos.

El origen de la habilidad de predicción se investiga en los patrones singulares. Nos vamos a centrar en el primer modo singular, por ser el que más fracción de covarianza explica y el que tiene más peso en las predicciones. La figura 4 muestra el primer modo singular obtenido cuando se elabora una predicción del verano con 6 meses de antelación, o en otras palabras, cuando el TOR1 (invierno de SST) predice el verano 6 meses después. Las regiones de mayor predecibilidad son aquéllas en las que el patrón tiene su máxima amplitud. Claramente vemos que, en el caso del predictor y del predictando, la predecibilidad está localizada en las región de la bajada del Giro Subtropical (Corriente de Canarias), en la Península Ibérica y en la región oceánica cercana al subtropical.

5. CONCLUSIONES

Hemos estudiado la predecibilidad de las anomalías de T850 en el Atlántico Norte con un modelo empírico basado en la propiedad matricial conocida como SVD. La SVD puede aislar la variabilidad acoplada entre dos campos, produciendo pares de patrones que expliquen la máxima covarianza entre el campo predictor (SST) y el campo predictando (T850). La SVD realiza una descomposición de la matriz de covarianza cruzada entre estas dos variables. El estudio se basa en la hipótesis de Bjerkness (1964), en la que el ruido rojo característico de la variabilidad oceánica, y que en parte resulta del forzamiento atmosférico en escalas temporales cortas, puede a su vez producir anomalías atmosféricas mediante complicados mecanismos de retroalimentación.

Las bases del modelo empírico que aquí se ha empleado fueron introducidas en SÁNCHEZ GÓMEZ *et al.* (2001, 2002). El modelo permite ahondar un poco más en las relaciones físicas entre el océano y la atmósfera ya que produce unos patrones que resumen la variabilidad conjunta entre estos dos campos. En la versión que se empleó en estos trabajos, el modelo se aplicó a campos filtrados, en los que se eliminaron las frecuencias superiores a los 8 meses. En esta nueva versión, se trabaja con campos estacionales. Los niveles de la habilidad de predicción son muy parecidos con los dos métodos, sin embargo, el segundo método predice peor en la región de la Península Ibérica. Un gran avance que supone trabajar con campos estacionales es que los datos no se filtran y por tanto se emplea toda la información contenida en los campos. Las predicciones realizadas con el nuevo esquema estacional logran una mejora significativa frente a las de la persistencia. Este hecho es muy importante porque puede permitir que un modelo empírico se utilice para validar los resultados de los GCMs.

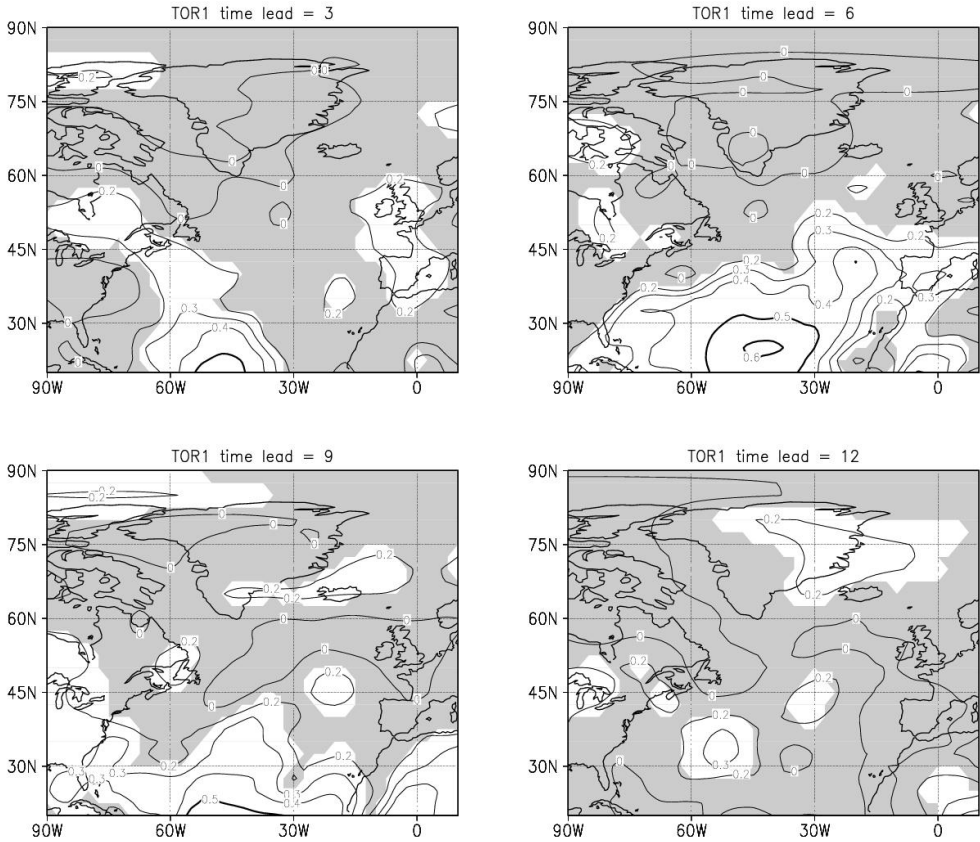


Figura 2: Distribución espacial de la habilidad de predicción para las predicciones de T850 efectuadas por el TOR1 (invierno) de SST. La línea gruesa indica los valores de H superiores a 0.5. Las regiones sombreadas indican dónde H está por debajo de los valores estimados para la habilidad de predicción artificial.

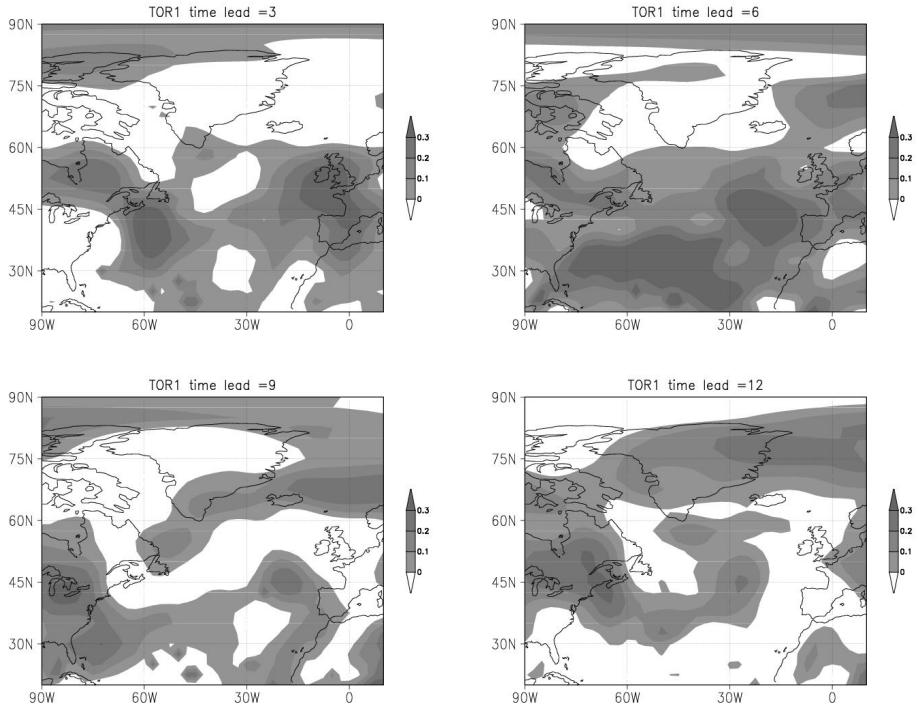


Figura 3: Diferencia entre la habilidad de predicción del modelo empírico y la obtenida asumiendo persistencia.

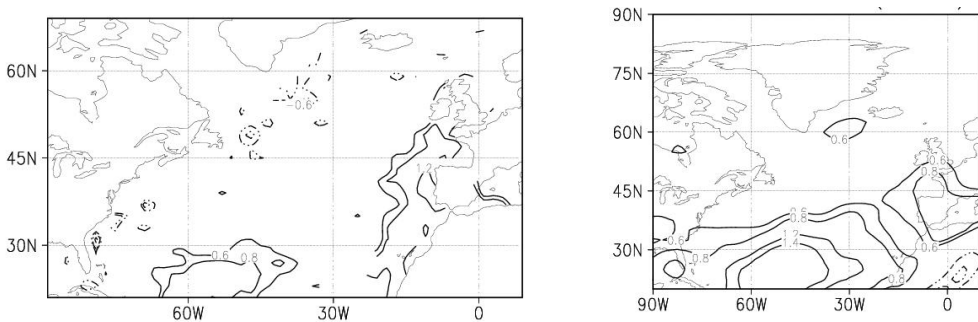


Figura 4: Primer modo singular obtenido de las predicciones del verano de T850 con el TOR1 (invierno) de la SST. El patrón oceánico a la izquierda y el atmosférico a la derecha. Este patrón explica un 26 % de la covarianza entre los dos campos.

Los orígenes de la habilidad de predicción se han investigado en los primeros modos singulares. La estructura espacial de los patrones muestran que la predecibilidad está confinada en las regiones de la bajada del Giro Subtropical y en el sur del dominio. Las regiones del norte casi no tienen predecibilidad.

6. AGRADECIMIENTOS

Los autores quieren agradecer la colaboración del resto del grupo de Física del Clima de la Universidad de Alcalá: Francisco Álvarez García, William Cabos Narváez y Antonio Ruiz de Elvira.

El trabajo ha sido financiado por el Programa Medio Ambiente y Clima de la Unión Europea, contratos ENV4-CT95-0109 (PROVOST) y ENV4-CT98-0714 (SINTEX).

7. REFERENCIAS

- BJERKNES, J. (1964): Atlantic air-sea interaction. *Adv. In Geophys.*, 10, pp. 10-82.
- BRANKOVIĆ, C., PALMER, T.N. y FERRANTI, L. (1994): Predictability of seasonal atmospheric variations. *J. Climate*, 7, pp. 217-237.
- EFRON, B. Y TIBSHIRANI, J.R. (1993): *An Introduction to the Bootstrap*. Chapman & Hall, 436 pp.
- JOHANSSON, A., BARNSTON, A.G., SURANJANA, S. y VAN DER DOOL, H. (1998): On the level and origin of the forecast skill in Northern Europe. *J. Climate*, 55, pp. 103-127.
- PENLAND, C. y MAGORIAN, T. (1993): Prediction of el Niño3 sea surface temperature using Linear Inverse Modeling. *J. Climate*, 9, pp. 1824-1839.
- PENLAND, C. y MATROSOVA, L.M. (1998): Prediction of Tropical Atlantic sea surface temperature using Linear Inverse Modeling. *J. Climate*, 11, pp. 483-496.
- RUIZ DE ELVIRA, A., CABOS NARVÁEZ, W.D. y ORTIZ BEVIÁ, M.J. (2000): Empirical forecasts of Tropical Atlantic sea surface temperature anomalies. *Quart. J. Roy. Meteor. Soc.*, 126, pp. 2199-2210.
- SÁNCHEZ GÓMEZ, E., GUTIÉRREZ CEBALLOS, C. y ORTIZ BEVIÁ, M.J. (1999): Empirical seasonal forecasts of North Atlantic atmospheric anomalies. *Proceedings of the XXI IUGG Meeting*, Birmingham, 49 pp.
- SÁNCHEZ GÓMEZ, E., ÁLVAREZ GARCÍA, F.J. y ORTIZ BEVIÁ, M.J. (2001): Empirical prediction of 850 hPa North Atlantic air temperature anomalies. *Quart. J. Roy. Meteor. Soc.*, 127, pp. 2761-2786.
- SÁNCHEZ GÓMEZ, E., CABOS NARVÁEZ, W. y ORTIZ BEVIÁ, M.J. (2002): Sea ice anomalies as long range predictors of atmospheric conditions. *Tellus A*, 45, 245-259.
- SHABBAR, A. y BARNSTON, A. (1996): Skill of seasonal climate forecast in Canada using Canonical Correlation Analysis. *Mon. Wea. Rev.*, 124, 2370-2385.
- VAUTARD, R., PLAUT, G., WANG, R. y BRUNET, G. (1999): Seasonal prediction of North American surface air temperature using space-time Principal Components. *J. Climate*, 12, pp. 380-394.