

REDES NEURONALES Y PATRONES DE ANALOGÍAS APLICADOS AL DOWNSCALING EN MODELOS CLIMÁTICOS

Rafael CANO TRUEBA*, José M. GUTIÉRREZ**,
Antonio COFIÑO** y Miguel A. RODRÍGUEZ***

**Instituto Nacional de Meteorología, CMT/CAS, Santander*

*** Dpto. de Matemática Aplicada. Universidad de Cantabria*

****Instituto de Física de Cantabria, CSIC*

RESUMEN

Este artículo describe un sistema experto para la simulación climática local utilizando las salidas de los modelos climáticos sobre un área limitada supra-peninsular y buscando analogías en las bases de datos de cada modelo. Éste conjunto de analogías entrena a una red neuronal sobre los datos locales de cualquier observatorio resultando un sistema objetivo para interpretar localmente las salidas de los modelos climáticos.

Palabras clave: Predicción estacional, redes neuronales, cálculo de analogías, downscaling, simulación.

ABSTRACT

This article describes an expert system for the climatic simulation using the low resolution outputs of the climatic models over a supra-peninsular limited area and looking for similar configurations in the data bases of each model. These allow to train a neural net using the local data of any observatory as output. It provides an objective system in order to interpret the outputs of the climatic models locally.

Keywords: Seasonal prediction, neural nets, calculation of analogies, downscaling, simulation.

1. INTRODUCCIÓN

Los modelos numéricos globales de simulación atmosférica actuales se basan en la integración numérica de complejos sistemas de ecuaciones diferenciales (predicción a medio y corto plazo) o sistemas acoplados atmósfera-océano-tierra de circulación general (predicción climática a largo plazo: GCM's). El tiempo de cómputo necesario para resolver estos sistemas de ecuaciones diferenciales hace que, hasta la fecha, el tamaño de rejilla sea tan grande que incluye zonas de gran variedad climática.

Sabemos, sin embargo, que estas singularidades climáticas están registradas en las series de datos que han sido tomados diariamente durante décadas en los numerosos observatorios del Instituto Nacional de Meteorología (I.N.M.). Por ejemplo, la Figura 1 muestra los observatorios del I.N.M. en la Zona del Cantábrico, donde se dispone de registros históricos completos (presión, viento, temperaturas, precipitación, evaporación, humedad, insolación, meteoros, ...) en siete estaciones (indicadas en el mapa con +) y registros de temperatura, precipitación y meteoros en numerosas estaciones (indicadas en el mapa con un triángulo en la Fig. 1).

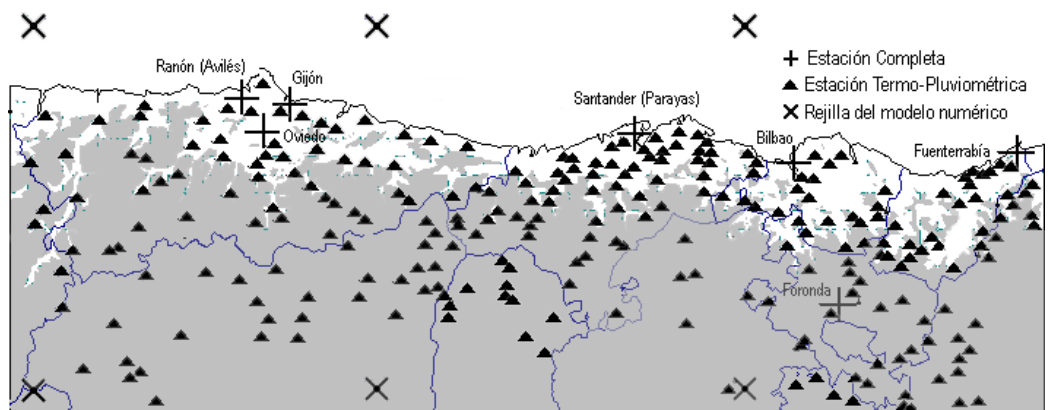
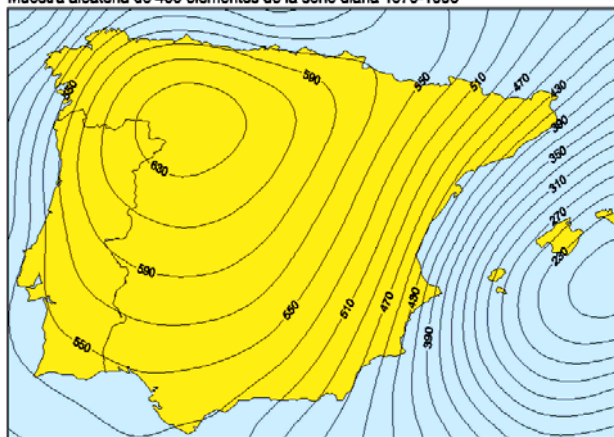


Figura 1. Estaciones completas (+) y termo-pluviómetricas (▲) de la Red Nacional de Observatorios del I.N.M. del Centro Meteorológico Territorial de Cantabria y Asturias. También se muestran los correspondientes puntos de rejilla (X) del modelo climático del Centro Hadley.

La notable diferencia de patrones climáticos entre estaciones cercanas que escapan de la resolución de los modelos se pone de manifiesto en las Figuras 2a y 2b, donde se muestran las correlaciones de

Mapa de Correlación para TM_{máx} en Oviedo con 25 Análogos
 Predictor: TSup 12z Serie ECMWF Re-Análisis
 Muestra aleatoria de 430 elementos de la serie diaria 1979-1993



las temperaturas máximas diarias de los observatorios de Avilés (Ranón) y Oviedo con las temperaturas de otros puntos de la península.

Figura 2a. Correlación de la temperatura máxima en superficie a las 12 horas en Oviedo respecto a las temperaturas en 1000 mb de otros puntos de la península.

Mapa de Correlación para TM_{12h} en Ranón con 25 Análogos
 Predictor: TSup 12z Serie ECMWF Re-Análisis
 Muestra aleatoria de 430 elementos de la serie diaria 1979-1993

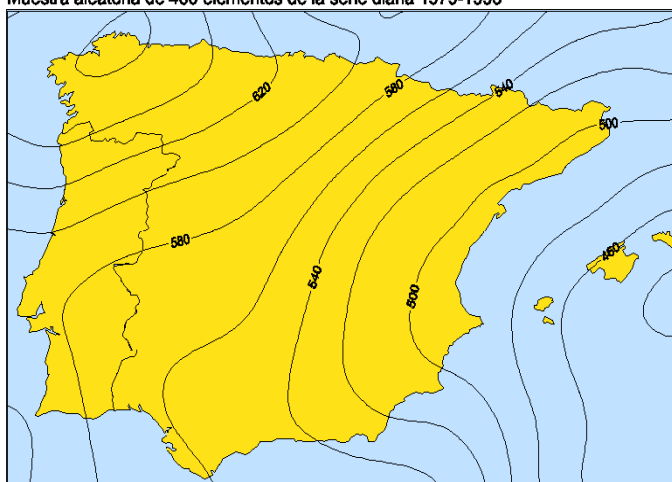


Figura 2b. Correlación de la temperatura máxima en superficie a las 12 horas en Ranón respecto a las temperaturas en 1000 mb de otros puntos de la península. Las temperaturas de ambas estaciones presentan patrones de correlación muy distintos, pese a distar sólo 25 Km.

Estos hechos ponen en evidencia la limitación de la resolución actual de los modelos climáticos a la hora de realizar pronósticos a nivel local, ó en áreas donde existan singularidades climáticas. En este artículo se presenta una nueva metodología para la mejora de la resolución mediante un downscaling estadístico sobre las salidas de estos modelos aprovechando los registros históricos de la red de observación del I.N.M. Esta nueva metodología se fundamenta en la selección automática de patrones análogos con los que se entrena una red neuronal para cada observatorio local donde se quieran extrapolar los campos propuestos por el modelo climático. Se ha constatado que esta técnica resulta más eficiente y objetiva que otros métodos lineales utilizados previamente.

Este sistema fruto de la metodología expuesta anteriormente ha sido desarrollado en colaboración con la Universidad de Cantabria, para el área del Centro Meteorológico Territorial de Cantabria y Asturias y se halla actualmente en fase de experimentación.

2. MODELOS DE REGRESIÓN VS. REDES NEURONALES

En los últimos años se ha producido un importante avance en el desarrollo de nuevas técnicas inteligentes para el análisis de datos y el aprendizaje automático a partir de éstos. Estas nuevas técnicas (los sistemas expertos, CASTILLO, 1997; las redes neuronales, BISHOP, 1997; las redes funcionales, CASTILLO, 1999, etc.) extienden las posibilidades de los análisis estadísticos estándar permitiendo desarrollar modelos más potentes y flexibles. Por ejemplo, las redes neuronales multicapa son una generalización de los modelos autoregresivos lineales y logísticos. Estas nuevas técnicas tienen una sólida base teórica y práctica y están siendo aplicadas en diversos ámbitos de la meteorología (NEBEKER, 1995), mostrando ser más eficientes que las técnicas estadísticas clásicas de regresión y clasificación; este hecho se pone de manifiesto especialmente en problemas como en el

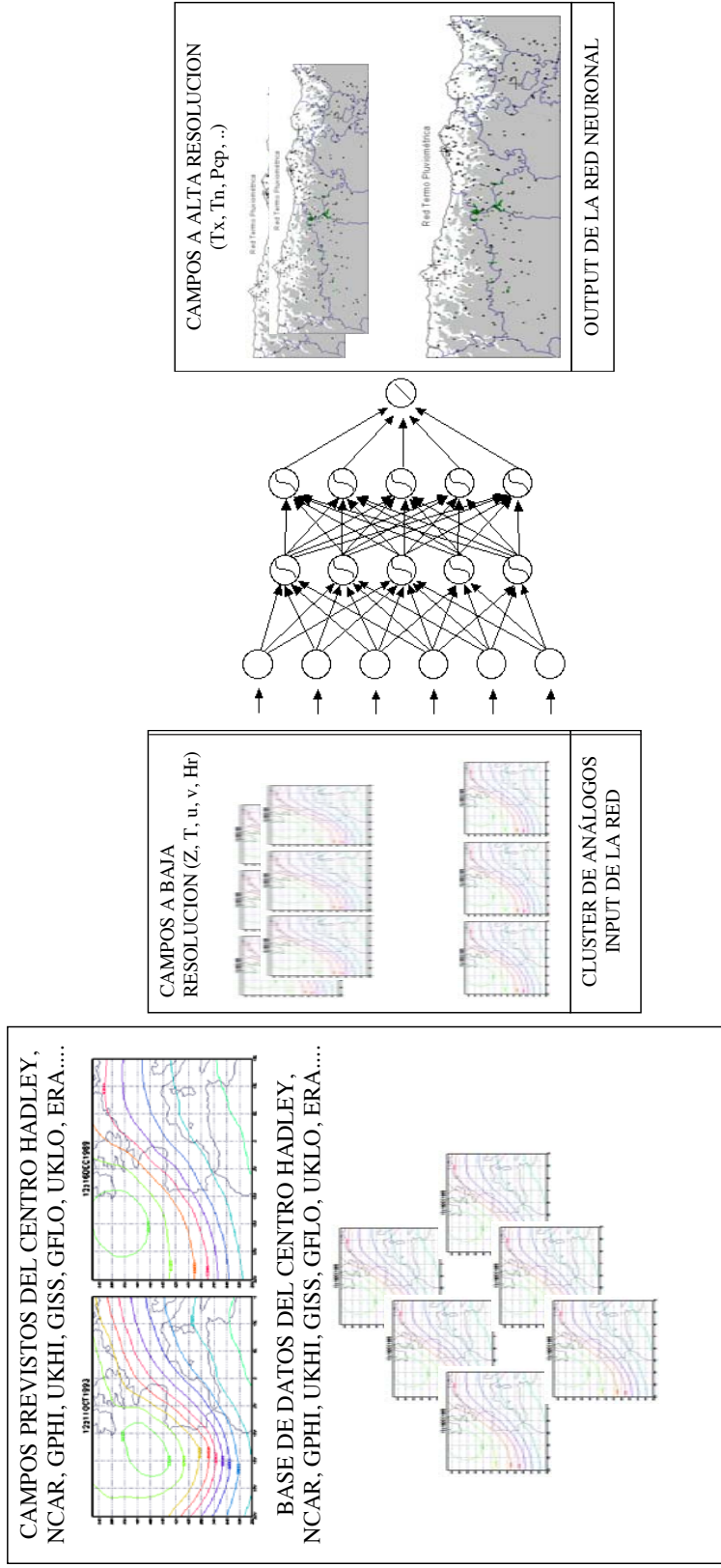


Figura 3. Ilustración esquemática del proceso de selección de análogos y posterior entrenamiento de la Red Neuronal. Aunque el sistema es válido para cualquier modelo, por el momento sólo disponemos de series de los modelos del Centro Hadley, del experimento PROVOST y del ECMWF: ERA. Concretamente el ejemplo es para la serie ERA.

que se aborda en este trabajo, donde se dispone de un registro histórico significativo (datos de aprendizaje) a partir del cual se quieren inferir pronósticos para situaciones concretas. El proceso de downscaling desarrollado está basado en dos etapas: En primer lugar, se realiza un cálculo de analogías entre las salidas del modelo climático y los registros disponibles en la base de datos correspondiente al modelo empleado, obteniéndose un conjunto de situaciones análogas. A continuación se infieren pronósticos a mayor resolución en base a los campos generados por el modelo y a los registros históricos locales para las situaciones análogas obtenidas. Se comparan los resultados obtenidos en el proceso de inferencia utilizando un modelo de regresión lineal clásico y una red neuronal que capturan en distinta medida la dinámica de evolución de las situaciones meteorológicas análogas.

3. SELECCION DE ANALOGOS

La primera etapa que realiza el sistema es seleccionar, de la base de datos, un conjunto de situaciones análogas (en nuestro área de trabajo) para cada salida del modelo climático. Este conjunto se obtiene comparando los estados de la atmósfera predichos por el modelo con los estados almacenados en su base de datos. El número de variables asociadas con los campos del grid del modelo es tan grande que es difícilmente manejable; para reducir la complejidad hemos empleado un criterio mixto de selección de variables que tiene en cuenta el comportamiento de la atmósfera a escala sinóptica en niveles medios y altos y a mesoescala en niveles medios y bajos (esta técnica es una variación de la de 'rejillas alternadas' empleada por el EPS en el ECMWF): para niveles medios y altos se considera un conjunto de puntos de grid uniformemente distribuidos sobre un área limitada de escala sinóptica y para la mesoescala se consideran los puntos de grid más cercanos a la estación donde se desee realizar el downscaling (ver Figura 1). Por tanto, en una primera aproximación, la discrepancia entre dos configuraciones atmosféricas dadas se obtendría en base a la distancia euclídea de los campos numéricos resultantes del proceso anterior. Más detalles sobre este modelo de Selección se describirán en un trabajo más extenso, donde se analizará este problema en mayor detalle considerando un modelo de red neuronal autoorganizativa.

Con objeto de ilustrar el funcionamiento del método, en la Figura 4 se muestra el cluster de analogías con 20 elementos obtenido a partir de una configuración dada por el modelo climático (marcada en la figura con '+').

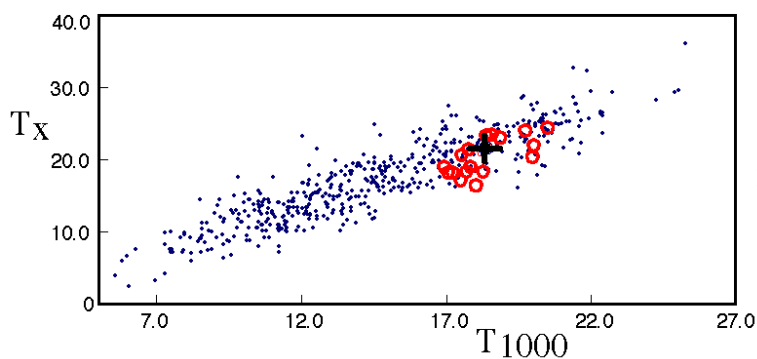


Figura 4. Cluster de analogías con veinte elementos (señalados con círculo) para un día problema dado.

La Figura 5 muestra en detalle los campos de geopotencial en 1000 y 500mb previstos por el modelo (parte superior) junto con el mejor análogo obtenido (parte inferior). En estas figuras se puede observar la concordancia en ambos campos de geopotencial.

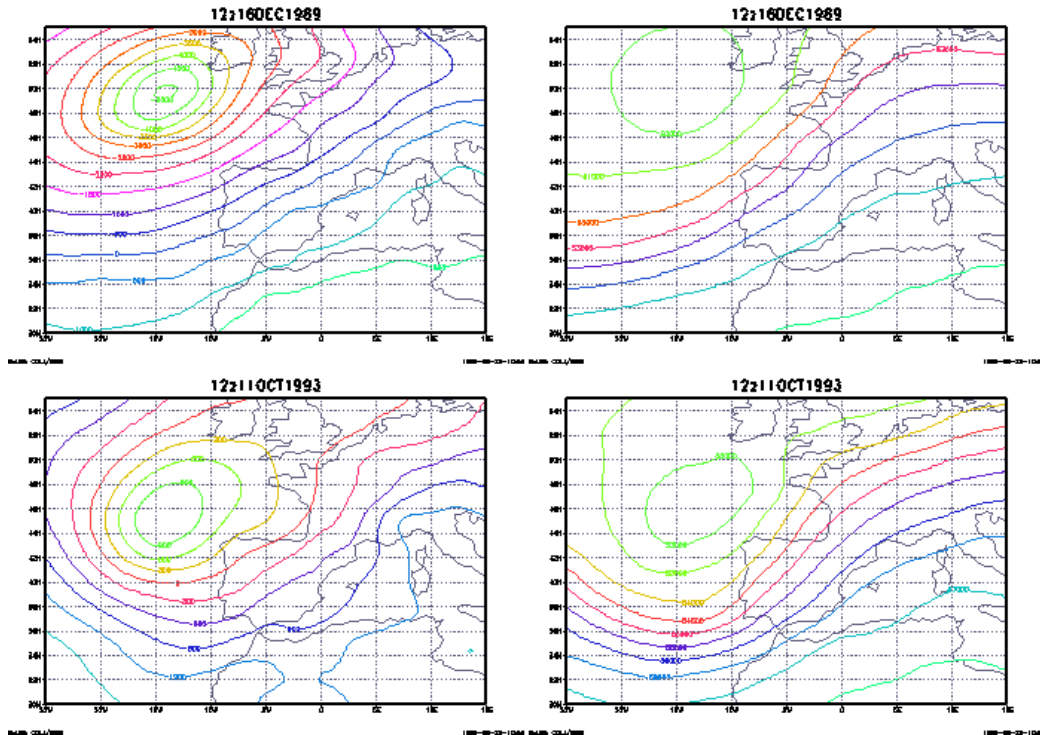


Figura 5. Búsqueda de análogos para una configuración dada (parte superior). El mejor análogo obtenido, para este caso, fue el 11 de Octubre de 1993 (parte inferior).

4. MODELOS DE REGRESION Y REDES NEURONALES

En base a los días análogos obtenidos se puede obtener una interpretación local en una zona concreta donde se encuentre un observatorio del INM. Para ello bastará extraer una relación entre los valores x del modelo climático (por ejemplo los campos de temperaturas, geopotenciales, etc., en los puntos de grid) y el valor a inferir y registrado en el observatorio (por ejemplo, las temperaturas) en los días análogos.

Para ello, la opción más sencilla es utilizar un modelo de regresión múltiple que establezca la relación lineal óptima entre las variables. La Figura 6 muestra un ejemplo sencillo donde se ilustra la conveniencia de utilizar un modelo no lineal más flexible (por ejemplo, una red neuronal). En este ejemplo, puramente ilustrativo, se infiere la temperatura máxima en Oviedo a partir de una única variable (por ejemplo la temperatura generada por el modelo en el punto de grid más cercano). Después de numerosas comparaciones en varias estaciones, se ha constatado heurísticamente que los mejores resultados se obtienen con conjuntos de análogos comprendidos entre 25 y 50 elementos.

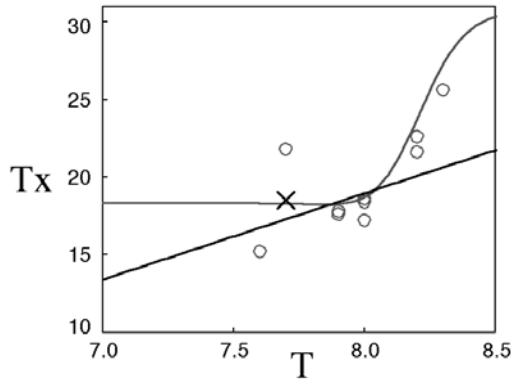


Figura 6. Comparación de un modelo de regresión lineal y una red neuronal para obtener la temperatura máxima en Oviedo T_x , a partir de la temperatura T simulada por el modelo en el nodo más cercano. El punto marcado con X corresponde al valor correcto.

5. SALIDAS DEL SISTEMA

En esta sección se ilustra la la tasa de fiabilidad del sistema en condiciones óptimas para predicción estacional, determinando su potencial de simulación real del clima local para periodos cortos de control, es decir obtenemos cuál sería su comportamiento en el caso de que el modelo fuese perfecto.

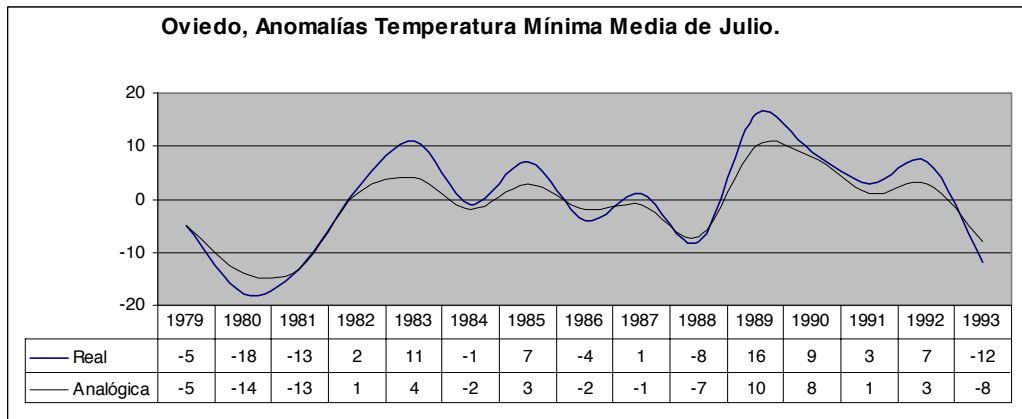


Figura 7a. Anomalías en la temperatura mínima media (en décimas) del mes de Julio en Oviedo. La Real es la que se observó y la Analógica es la interpretada por el sistema a partir de un modelo de simulación climática.

La intención última es su aplicación sobre modelos climáticos, en esta línea actualmente se están realizando las pruebas preliminares sobre el experimento PROVOST (PRediction Of climate Variations On Seasonal to interannual Timescales) promovido por el ECMWF para predicción estacional. Este experimento se corre sobre cuatro modelos diferentes: ECMWF, Meteo France, EDF y UKMO y consiste, muy brevemente, en aplicar los citados modelos a 120 días a partir de 9 días consecutivos con resolución T63L31.

Aunque la resolución espacial y temporal de estos modelos globales es inferior a la del empleado en el control, al tratarse de una técnica perfect-prog la diferencia de calidad

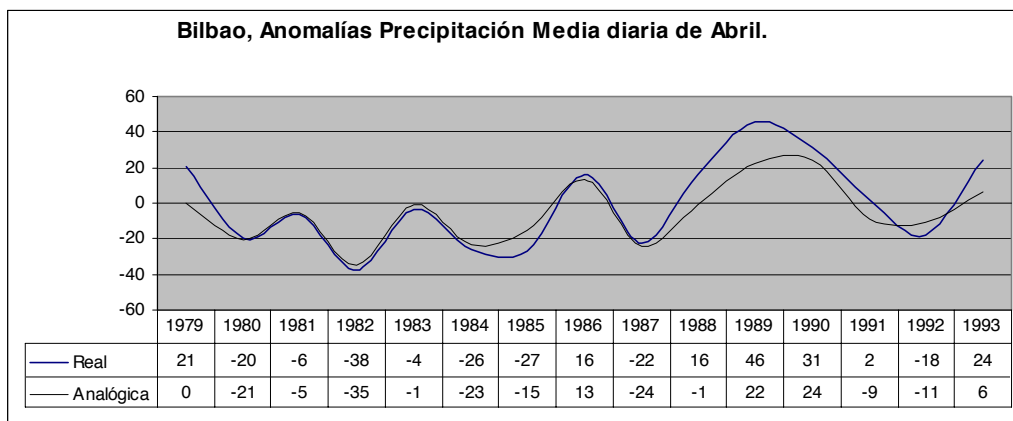


Figura 7b. Anomalías en la precipitación media diaria (en décimas) del mes de Abril en Bilbao. La Real es la que se observó y la Analógica es la interpretada por el sistema a partir de un modelo de simulación climática.

final entre el control y el experimento real se debe mayoritariamente al error cometido por el modelo climático. En este sentido la técnica desarrollada está más avanzada que el modelo sobre el que se aplica, por lo que podemos suponer que durante los próximos años, el sistema sea capaz de asimilar automáticamente todas las mejoras de resolución, y en definitiva, de calidad de los modelos climáticos, que ya vienen siendo muy notables respecto a modelos precedentes.

En no más de cinco años cabe esperar que las predicciones estacionales emitidas por los modelos climáticos se aproximen a la precisión y resolución que tienen actualmente los modelos de predicción a medio plazo ya que se están aclarando algunos de los puntos negros en la descripción de los elementos del sistema climático como el balance radiativo, la circulación oceánica, la química atmosférica y estratosférica y la componente biológica.

REFERENCIAS BIBLIOGRÁFICAS

BISHOP, B. (1997): *Neural Networks for Pattern Recognition*. Clarendon Press.

BRANKOVIC, C., PALMER T.N. (1998): *Seasonal skill and predictability of ECMWF PROVOST ensembles*. Research Department. Technical Memorandum N° 263, ECMWF.

CANO R. (1996): *Sistema experto en interpretación de salidas de Modelos Numéricos*. Nota Técnica 3.5, 1993, CMT CAS, I.N.M. IV Simposio Nacional de Predicción. I.N.M.

CARSON, D.J. (1998): *Seasonal Forecasting*. Q.J.R. Meteorol. Soc.

CASTILLO, E., GUTIERREZ, J.M. y HADI, A. (1997): *Expert Systems and Probabilistic Net-work Models*. Springer-Verlag.

CASTILLO, E. *et al.* (1999): *An Introduction to Functional Networks with Applications*. Kluwer Academic Publishers, Boston, USA.

GARDNER M.W., DORLING S.R. (1998): *Artificial Neural networks (The Multilayer Perceptron)- A review of applications in the atmospheric sciences*. Atmospheric environment, OXFORD, ENGLAND.

NEBEKER, F. (1995): *Calculating the Weather: Meteorology in the 20th Century*. Academic Press .

PALMER T.N. *et al.* (1998): *A Probability and Decision-Model Analysis of PROVOST Seasonal Multi-Model Ensemble Integrations*. Research Department. Technical Memorandum N° 265, ECMWF.

REICHERT B.K., BENGTSSON L., ÅKESSON O. (1998): *A Statistical-dynamical modeling approach for the simulation of local paleo proxy records using GCM output*. Max-Planck-Institut für Meteorologie. Report N° 274, 1998.

REPORT ECMWF (1999): *Seasonal Forecasting Users Meeting 17-18 June 1988*. ECMWF.

STOCKDALE T.N. *et al.* (1998): *Global seasonal rainfall forecasts using a coupled ocean-atmosphere model*. Nature, vol 392, 26 march.

