



Solicitud de ayuda para

Proyectos de Investigación Científica y Desarrollo Tecnológico (Modalidades P1, P2 y P3)

SEÑALE ÁREA, PROGRAMA Y MODALIDAD PARA LOS QUE SOLICITA LA AYUDA, ASÍ COMO TIPO DE ENTIDAD SOLICITANTE

ÁREA DE INVESTIGACIÓN BÁSICA NO ORIENTADA

- Programa Nacional de Promoción General del Conocimiento:
- Física y Matemáticas
 - Biología de Organismos y Sistemas
 - CC de la Tierra y del Espacio
 - CC Jurídicas
 - Química
 - Bioquímica, Biología Molecular y Celular
 - CC Sociales
 - Filología y Filosofía
 - Fisiología
 - CC Económicas
 - Historia y Arte
- Programa Nacional de Astronomía y Astrofísica
- Programa Nacional de Física de Partículas y Grandes Aceleradores
- Programa Nacional de Fusión Termonuclear

ÁREAS CIENTÍFICO-TECNOLÓGICAS

- Programa Nacional de Biotecnología
- Programa Nacional de Diseño y Producción Industrial
- Programa Nacional de Materiales
- Programa Nacional de Procesos y Productos Químicos
- Programa Nacional de Recursos Naturales
- Programa Nacional de Recursos y Tecnologías Agroalimentarias
- Programa Nacional de Tecnologías de la Información y las Comunicaciones
- Programa Nacional de Socioeconomía

MODALIDAD: P1 P2 P3

TIPO DE ENTIDAD SOLICITANTE:

- Centro Público de I+D
- Centro privado de I+D sin ánimo de lucro
- Centro Tecnológico

Plan Nacional de I+D+I (2000-2003)

1. SOLICITUD

1.1. DATOS DEL PROYECTO

Título: **Técnicas Híbridas Estadístico-Numéricas de Predicción Meteorológica en diferentes Escalas Espacio-Temporales**

Acrónimo: METHYFOR (METeorological HYbrid FORecasting)

Área temática (*para PGC, modalidad P1*):

Número del objetivo científico-tecnológico (*para mod. P2, P3*): 1 (*atmósfera y clima*).

Clasificación UNESCO: 2509,1209

Duración (en años): 3

Número de investigadores: 5

1.2. PALABRAS CLAVE PARA LA IDENTIFICACIÓN DEL PROYECTO (relacionadas con el tema, tecnologías empleadas y aplicaciones del proyecto)

Predicción meteorológica y climatológica. Predicción probabilística. Downscaling. Redes neuronales. Redes probabilísticas. Métodos híbridos. Cálculo de analogías. Componentes principales no lineales.

1.3. DATOS DE LA ENTIDAD SOLICITANTE

Centro Público de I+D Centro privado de I+D sin ánimo de lucro

Centro Tecnológico (número de Registro oficial como CIT):

Nombre: Universidad de Cantabria

Acrónimo: UNICAN

C.I.F.: Q-3918001-C

Nombre del representante legal: D. José Jorda Catalá

Cargo: Excmo. Sr. Vicerrector de Investigación y Desarrollo

Teléfono: (ext.) Telefax:

Correo electrónico:

Dirección postal completa: Pabellón de Gobierno, Universidad de Cantabria, Avda. de los Castros s/n Santander (39005)

1.4. DATOS DEL INVESTIGADOR RESPONSABLE

Apellidos: Gutiérrez Llorente

Nombre: José Manuel

Entidad: Universidad de Cantabria

Centro: E.T.S.I. Caminos, Canales y Puertos

Departamento: Matemática Aplicada y Ciencias de la Computación

Teléfono: 942 201723 (ext.) Telefax: 942 201703

Correo electrónico: gutierjm@ccaix3.unican.es

Dirección postal completa: E.T.S.I. Caminos, Avda. de los Castros, s/n (39005) Santander

Es Doctor: SI NO

3.1. RESUMEN DE LA PROPUESTA

RESUMEN (debe ser breve y preciso, con los aspectos más relevantes y los objetivos propuestos)

Tradicionalmente se han venido utilizando técnicas estadísticas tales como "MOS", "Perfect Prog", filtros de Kalman, o técnicas basadas en análogos, para adaptar las salidas directas de los modelos numéricos de predicción meteorológica a las características locales y las singularidades geográficas de los lugares donde ha de precisarse la predicción. A medida que aumenta la resolución de los modelos, las salidas directas de los mismos compiten mejor con estas técnicas estadísticas, aunque para determinados elementos meteorológicos (especialmente los relacionados con la humedad y precipitación) queda un amplio margen para el perfeccionamiento de su predicción mediante la aplicación de técnicas estadísticas. Por otra parte, los modelos numéricos que se utilizan en predicción estacional y en la obtención de escenarios de cambio climático, e incluso en la preparación de los miembros de una predicción por conjuntos en el medio plazo, utilizan resoluciones próximas a 120 Km. Por tanto, el uso de técnicas de disminución de escala (downscaling) mediante métodos dinámicos o estadísticos se hace necesario en estos casos.

El primer objetivo de este proyecto es desarrollar métodos híbridos para la mejora de resolución de la predicción meteorológica en distintas escalas temporales, que abarcan desde el corto plazo hasta la predicción climática, en ámbitos regionales y locales. Estos métodos combinarán las salidas (las predicciones) de los modelos numéricos con la información diaria y con la información estadística contenida en los registros históricos disponibles en el área de interés y que caracterizan la meteorología y la climatología de esa área.

Para ello, el punto de partida será generar un sistema capaz de identificar cuantitativamente situaciones atmosféricas análogas, para lo cual se desarrollarán técnicas basadas en redes neuronales y funcionales. Este punto precisará de un análisis de componentes principales no lineales para sintetizar la información de forma que los algoritmos desarrollados tengan un uso operativo en tiempo real. El conjunto de situaciones atmosféricas análogas a otra dada servirá para inferir los pronósticos locales relacionando los campos análogos obtenidos con los registros correspondientes de los meteoros de interés en áreas locales concretas

Por otra parte, la necesidad de estimar con anticipación la incertidumbre de las predicciones ha motivado el desarrollo de distintas técnicas de predicción probabilística por conjuntos que ya están operativas para la predicción a plazo medio, y comienzan a experimentarse en la predicción a corto plazo, a la vez que al extenderse su uso a la predicción estacional y a la predicción climática se exploran nuevos procedimientos para preparar los miembros de estos conjuntos.

Como segundo objetivo, para la predicción en el corto plazo se ensayarán técnicas perturbativas estándar para el cálculo de los miembros del conjunto (vectores singulares, breeding fast growing modes, método LAF "Lagged Average Forecasting", cambios en las parametrizaciones físicas de los modelos, etc.) y se compararán con métodos computacionalmente más económicos basados en análogos, a partir de las evoluciones de un conjunto de configuraciones análogas a una configuración inicial.

Para la predicción estacional también se ensayarán técnicas basadas en la localización de análogos de temperaturas del agua del mar, donde cada miembro del conjunto será el promedio trimestral de los valores de los elementos meteorológicos asociados a situaciones sucesivas, para cada situación análoga encontrada. Este estudio permitirá contrastar la validez, utilidad y valor económico de los modelos basados en técnicas perturbativas.

El objetivo final del proyecto será el desarrollo de una aplicación que incorpore los resultados obtenidos en el transcurso del proyecto para su instalación, validación y uso operativo en el Instituto Nacional de Meteorología (INM), con el consiguiente beneficio social derivado de su aplicación.

TITLE OF THE PROJECT:

Hybrid Statistical-Numerical Techniques for Weather Forecasting in several Spatio-Temporal Scales.

SUMMARY:

Traditionally, techniques such as "MOS", "Perfect Prog", Kalman filters, or techniques based on analogies, have been used for adapting the outputs of numerical weather forecasting models to the local and regional characteristics and geographic singularities of places where the predictions are made. Although the resolution of short- and mid-range forecasting models have been substantially increased, there is still some room for these statistical downscaling techniques, specially for climatologic elements such as humidity and rain. On the other hand, long-range seasonal prediction models and models for climatic change scenario preparation, and even the preparation of ensemble members for medium-range ensemble prediction, use models with resolutions of approximately 120 Km, which require some kind of downscaling technique.

One of the objectives of this project is the development of hybrid methods for weather forecasting downscaling in several spatio-temporal scales, from the short- and mid-term, to seasonal forecasting in local and regional scales. These models will combine the outputs (forecasting) of numerical models with diary observations and statistical information contained in the historical registers available in the area of interest, which characterize the climatology of the area.

To this aim, we shall start with the development of a model for identifying similar atmospheric configurations, using neural and functional networks. This work will require a study of nonlinear principal components of the data for reducing and condensing the information, so the developed algorithms will be efficient. Hence, local forecasting will be performed by using a set of atmospheric configurations similar to a given one, by relating the selected fields with the climatologic elements of interest in local areas.

On the other hand, the convenience of estimating in advance the uncertainty associated with the predictions have motivated the development of several probabilistic "ensemble prediction" techniques, which are already operative for mid-range probabilistic forecasting. These techniques are currently being analyzed and adapted to short-range and seasonal forecasting.

We shall run several experiments for short-range probabilistic forecasting using perturbative standard methods for obtaining the members of the ensemble (singular vectors, breeding fast growing modes, Lagged Average Forecasting, changing physical parametrizations, etc.); these standard method will be compared with computationally economic methods based on analogies. In this case, the ensemble will be formed with a set of similar configurations to the initial one, studying their evolutions.

For seasonal forecasting we shall also use a method based on analogies of sea temperature, where each member of the set will be the average of the values of the evolutions of each of the found configurations similar to the initial one. This study will help to check the utility and economic value of models based on perturbative techniques.

A main objective of this project is the development of a program including the results and algorithms achieved in the project. This program will be installed in the Instituto Nacional de Meteorología to be validated and applied for daily forecasting.

3.1. INTRODUCCIÓN

Antecedentes y estado actual de los aspectos científico-técnicos, incluyendo la bibliografía más relevante.

Posibles coincidencias con actividades de otro(s) grupo(s) o Entidad(es) pública(s) y privada(s) en España, así como posible vinculación del tema propuesto con temáticas semejantes desarrolladas en ámbitos internacionales, preferentemente europeos.

- En las **modalidades P1 y P2**, destaque la originalidad y relevancia científica del tema propuesto.

El esfuerzo investigador llevado a cabo en las tres últimas décadas ha permitido un gran avance en el desarrollo de modelos de circulación atmosférica que incorporan las parametrizaciones de los fenómenos físicos relevantes, adecuados a las escalas espaciales a que dichos modelos se aplican. Por otra parte la aplicación operativa de estos modelos ha sido posible gracias a la disponibilidad de mejores y más complejos sistemas de observación (subsistema de observación terrestre: SYNOP, SHIP, TEMP, AIREP, DRIBU, etc. y subsistema espacial: SATEM, SATOB, etc., disponibles varias veces al día) y procesamiento de la información, así como a la gran atención prestada al problema de la asimilación de estos datos observacionales mediante el desarrollo y uso sucesivo de métodos de interpolación óptima y métodos de asimilación variacionales 3D y 4D, puestos a punto con el fin de establecer con la menor incertidumbre posible las condiciones iniciales a partir de las cuales se integran los modelos de predicción.

Como resultado de este esfuerzo, se dispone en la actualidad de eficientes modelos atmosféricos que se utilizan para la elaboración de un amplio abanico de predicciones con distintos alcances temporales: Por una parte, para la predicción a corto y medio plazo se utilizan modelos numéricos de área limitada y alta resolución; por otra parte los modelos acoplados océano-atmósfera de circulación general se aplican para la predicción de anomalías climáticas en la predicción estacional y para la preparación de escenarios climatológicos en función de diversos supuestos de forzamiento radiativo.

Así, las distintas predicciones que diariamente se preparan para el público y otros usuarios especializados en el Instituto Nacional de Meteorología y en otros centros meteorológicos europeos utilizan las salidas de diferentes modelos numéricos. Por una parte, para la predicción a corto y medio plazo se utilizan las salidas del modelo atmosférico global del Centro Europeo de Predicciones a Plazo Medio (CEPPM) y del modelo HIRLAM de área limitada (con una resolución espacial de 0.2° en comparación a los 0.6° del modelo del CEPPM). Por otra parte, el CEPPM integra diariamente y hasta 180 días un modelo acoplado océano-atmósfera para la elaboración de predicciones de anomalías climáticas trimestrales (predicción estacional). Finalmente, dentro del ámbito de la predicción climática, varios centros climáticos europeos (Centro Hadley en el Reino Unido, Instituto Max Plank en Alemania, Centro Nacional de Investigación de Méteo-France) llevan a cabo integraciones hasta 100 ó más años bajo diversas hipótesis de forzamiento radiativo, a las cuales tiene acceso el INM (a petición) para la realización de estudios de escenarios climáticos a escala regional.

Dos problemas comunes a toda esta gama de predicciones son:

- 3 La adaptación de las predicciones de los modelos numéricos a las características regionales y locales, para lo que se vienen aplicando distintas técnicas de disminución de escala (downscaling), y
- 4 La necesidad de estimar con anticipación la incertidumbre de las predicciones, para lo que se han desarrollado distintas técnicas de predicción por conjuntos (ensemble prediction).

En cuanto a la adaptación regional y local de las predicciones, se han venido utilizando tradicionalmente técnicas estadísticas como "MOS", "Perfect Prog" y filtros de Kalman sobre las salidas directas de los modelos numéricos (ver, por ejemplo, [Ay94], [Pet89]). A medida que aumenta la resolución espacial horizontal de los modelos, las salidas directas de los mismos compiten mejor con estas técnicas estadísticas, aunque para determinados elementos meteorológicos (especialmente los relacionados con la humedad y precipitación) queda un

amplio margen para su perfeccionamiento; este margen es más amplio cuanto más largos son los plazos de predicción, al disminuir con ellos la resolución de los modelos (120 Km. para el cálculo de los miembros de los conjuntos en predicción a plazo medio; alrededor de 200 Km. en la preparación de predicciones estacionales; 300 Km. o más para los modelos que se utilizan en predicción climática).

Otras líneas de actuación para desarrollar modelos de predicción local en el muy corto plazo se han basado únicamente en las series temporales de los registros históricos locales en los puntos de interés, a partir de los cuales se han intentado obtener modelos con capacidad predictiva. Sin embargo, los distintos análisis realizados utilizando modelos autoregresivos [BDSG96], redes neuronales [MAV94], o técnicas de embedding para modelización de sistemas dinámicos no lineales [BKPP93], están limitados desde el punto de vista de su capacidad predictiva operativa, y todavía no pueden considerarse alternativas eficientes en el marco de la predicción, a pesar de las expectativas suscitadas en algunos trabajos preliminares (ver [NN84] [TE88]).

Otro tipo de técnicas, basadas en el método de análogos, se han aplicado con éxito en algunos problemas concretos como, por ejemplo, la predicción de precipitaciones en ciertas regiones pluviométricas a nivel peninsular ([Can96], [RB95]) y en Cataluña [GLI99]. Estas técnicas están basadas en la idea de que estados análogos de la atmósfera están también asociados a fenómenos meteorológicos análogos [Lor69]. Bajo esta hipótesis, para un día concreto, es posible obtener de forma natural una predicción probabilística para un meteoro en un área local (temperatura, precipitación, etc.) a partir de la distribución que presentan los datos históricos del meteoro en un conjunto de días análogos extraído del registro de observaciones representativo de la climatología local correspondiente a esta situación atmosférica específica.

Los resultados obtenidos con esta técnica en problemas concretos muestran buenas expectativas en cuanto a sus posibilidades predictivas [GLI99] [RBA98]; más aún teniendo en cuenta que los métodos que se han utilizado hasta la fecha están basados en técnicas simples; por ejemplo, para el problema de la selección de análogos de un campo dado, se han venido utilizado técnicas estándar de agrupamiento basadas en la búsqueda directa de campos que presentan mínima distancia, o máxima correlación, con el campo dado. Actualmente se dispone de técnicas más eficientes que pueden ser utilizadas en este tipo de problemas, como las redes neuronales y funcionales [Bis97] [GD98] [CCGP99], que pueden autoorganizarse mostrando estructuras emergentes del conjunto de datos ayudando a una mejor agrupamiento y comprensión de los mismos, tal como se ha mostrado en recientes estudios preliminares [VVDT96] [ME97] [GCCR99].

Por otra parte, se dispone en la actualidad de amplios archivos de campos meteorológicos obtenidos a partir de las salidas diarias homogéneas de modelos de asimilación (por ejemplo el 'ECMWF Re-Analysis' ERA-15, que contiene un reanálisis de las salidas del modelo del CEPPM en el período 1979-1993; y el reanálisis de 40 años realizado por el National Weather Service de USA), lo que permite plantear técnicas basadas en análogos con garantías de obtener una buena aproximación estadística al problema. La carencia de este tipo de bases de datos hasta fechas muy recientes ha sido una de las mayores limitaciones para realizar estudios con este tipo de técnicas en este ámbito.

El segundo de los problemas antes mencionados, la estimación de la incertidumbre de las salidas de los modelos (englobado en el problema más general de la predecibilidad en meteorología), está relacionado con la estructura no lineal de los modelos atmosféricos y la sensibilidad que éstos presentan a los errores asociados con la especificación de sus condiciones iniciales y de contorno (ver [Tho57] y [Lor63]). Este problema obliga a pensar en la predicción del estado futuro de la atmósfera, en todos los plazos, en términos probabilísticos, teniendo en cuenta la sensibilidad de la predicción respecto a los errores en la especificación del estado inicial (ver [Pal00] para más detalles). Por ello, las condiciones iniciales habrán de representarse mediante distribuciones iniciales de probabilidad para las variables del modelo, y el problema de la predecibilidad consistirá en estimar la nueva distribución de las variables en un instante posterior.

Para abordar este problema, se han desarrollado diversos "métodos de predicción por conjuntos" que tratan de determinar las distribuciones de las variables del modelo en un instante dado a partir de la estadística contenida en un conjunto de predicciones obtenidas perturbando las condiciones iniciales (el método de vectores singulares [GB98], y el método de breeding [TK97] son dos de los métodos más populares desarrollados para este propósito). Estos métodos están ya en uso para las predicciones a medio plazo y actualmente se está estudiando su aplicación para otros plazos de predicción (corto plazo, estacional y preparación de escenarios climáticos).

Sin embargo, existe un fuerte debate sobre la validez de estas técnicas [EL99], especialmente sobre la representatividad que tiene la distribución inicial de probabilidades asignada a los miembros del conjunto con respecto de la incertidumbre inicial existente en el análisis, mostrando que se todavía trata de un campo en desarrollo donde se pueden aportar resultados.

Ésta será otra de las líneas de actuación en este proyecto, donde nos centraremos en la predicción por conjuntos en el corto plazo y en áreas regionales. Se ensayarán y compararán técnicas estándar para el cálculo de perturbaciones como los vectores singulares, 'breeding fast growing modes', método LAF ("Lagged Average Forecasting"), así como los cambios en las parametrizaciones físicas de los modelos. Se estudiarán también las limitaciones y las posibilidades de métodos basados en análogos para la selección de los miembros del conjunto que, en este caso, podrán suponerse equiprobables.

Se experimentará también la utilidad de un método computacionalmente económico para la predicción estacional basado en la localización de análogos de temperaturas del agua del mar en los Océanos Atlántico y Pacífico, donde cada miembro del conjunto será el promedio trimestral de los valores de las variables meteorológicas asociados a situaciones sucesivas de cada situación análoga encontrada. De esta forma se podrá contrastar la utilidad y valor económico de los modelos basados en técnicas perturbativas.

Referencias:

[Ay94] J.J. Ayuso, "Predicción Estadística Operativa en el INM". Monografías del Instituto Nacional de Meteorología, volumen B-34, 1994.

[BDSG96] J. Billet, M. DeLis, B.G. Smith, C. Gates "Use of Regression Techniques to Predict Hail Size and the Probability of Large Hail", Weather and Forecasting, Vol 12, 154-164, 1996.

[Bis97] Bishop, B. "Neural Networks for Pattern Recognition". Clarendon Press, 1997.

[BKPP93] T. Bountis, L. Karakatsanis, G. Papioannou, G. Pavlos "Determinism and Noise in Surface Temperature Series", Ann. Geophysicae, 11, 947-959, 1993.

[Can96] R. Cano "Sistema experto en interpretación de salidas CEPPM y HIRLAM". Proceedings del IV Simposio Nacional de Predicción del INM, 299-307, 1996.

[CGCR99] R. Cano, J.M. Gutiérrez, A.S. Cofiño y M.A. Rodríguez "Redes Neuronales y Patrones de Analogías Aplicados al Downscaling en Modelos Climáticos", en La Climatología Española en los Albores del Siglo XXI (Actas del I Congreso de la Asociación Española de Climatología). J.M. Raso y J. Martín-Vide editores, 113-121, 1999.

[CCGP99] Castillo, E., Cobo, A., Gutiérrez, J.M. and Pruneda, E. "An Introduction to Functional Networks with Applications". Kluwer Academic Publishers, 1999.

[EL99] R.M. Errico, R. Langland "Notes on the Appropriateness of 'Bred Modes' for Generating initial Perturbations used in Ensemble Prediction", Tellus, 51A, 431-451, 1999 (including a comment and a reply).

[GB98] R. Gelaro, R. Buizza, T.N. Palmer and E. Klinker "Sensitivity Analysis of Forecast Errors and the Construction of Optimal Perturbations Using Singular Vectors", Journal of the Atmospheric Sciences, Vol. 55, 1012-1037, 1998.

- [Gill82] A.E. Gill, "Atmosphere-Ocean Dynamics", International Geophysics Letters, Volume 30, 1982.
- [GCCR99] J.M. Gutiérrez, R. Cano, A.S. Cofiño and M. A. Rodríguez "PROMETEO: Un Sistema Experto para el Pronóstico Meteorológico Local basado en Redes Neuronales y Cálculo de Analogías", Actas de III Jornadas de Transferencia Tecnológica de Inteligencia Artificial, 11-19, 1999.
- [GD98] M.W. Gardner and S.R. Dorling "Artificial Neural Networks (the Multilayer Perceptron) - A Review of Applications in the Atmospheric Sciences", Journal of Applied Meteorology, Vol 39, 147-159, 1998.
- [GLI009] J. Gibergans y M.C. Llasat "Aplicación del Método de los Análogos a Cataluña", en La Climatología Española en los Albores del Siglo XXI (Actas del I Congreso de la Asociación Española de Climatología). J.M. Raso y J. Martín-Vide editores, 209-216, 1999.
- [Lor63] E.N. Lorenz "Deterministic Nonperiodic Flow", Journal of the Atmospheric sciences, 20, 130-141, 1963.
- [Lor69] E.N. Lorenz "Atmospheric Predictability as Revealed by Naturally Occuring Analogues", Journal of the Atmospheric sciences, 26, 636-646, 1969.
- [MAV94] Q.I. Moro, L. Alonso, C.E. Vivaracho, "Application of Neural Networks to Weather Forecasting with Local Data", Proceedings of the 12 IASTED International Conference on Applied Informatics, Vol 12, 68-70, 1994.
- [ME97] S.W. Miller and W.J. Emery, "An Automated Neural Network Cloud Classifier for Use over Land and Ocean Surfaces", Journal of Applied Meteorology, vol. 36, 1346-1362, 1997.
- [NN84] C. Nicolis and G. Nicolis, Nature, 311, 529-532, 1984.
- [Pal00] T.N. Palmer "Predicting Uncertainty in Forecasts of Weather and Climate", to appear in Reports on Progress in Physics, 2000.
- [Pet89] A.O. Persson, "Kalman filtering. A new approach to adaptive statistical interpretation of numerical meteorological forecasts". ECMWF Newsletter Num. 46, 16-20, 1989.
- [RBA98] B.K. Reichert, L. Bengtsson, and O. Akesson "A Statistical-Dynamical Modeling Approach for the Simulation of Local paleo Proxy Records Using GCM Output", Max-Planck Institut für Meteorologie, Tech. Report No. 274. 1998.
- [RB95] J. Ribalaigua y R Boren "Clasificación de patrones estacionales de precipitación diaria sobre España Peninsular y Baleares". Informe nº 3 del SAIC. INM 1995.
- [TE88] A.A. Tsonis and J.B. Elsner, "The Wheeler Attractor over Very Short Time Scales", Nature, 333, 545-547, 1988.
- [Tho57] P.D. Thompson, "Uncertainty of Initial State as a Factor in the Predictability of Large Scale Atmospheric Flow Patterns", Tellus, 9, 275-52, 1957.
- [TK97] Z. Toth and E. Kalnay "Ensemble Forecasting at NCEP and the Breeding Method", Mon. Weather Rev., Vol 125, 3297-3319, 1997.
- [VVDT96] M. Verdecchia, G. Visconti, F. D'Andrea, and S. Tibaldi "A neural network approach for blocking recognition" Geophysics Research Letters, 23(16), 2081-2084, 1996.

3.3. OBJETIVOS

Deben enumerarse y describirse con claridad y de manera realista (es decir, acorde con la duración prevista del proyecto) los objetivos concretos que se persiguen, los cuales deben adecuarse a las prioridades del Programa Nacional al que se adscribe el proyecto.

Se valorará la relevancia de los objetivos en el contexto científico-técnico correspondiente y de acuerdo con la modalidad elegida.

Los objetivos de este proyecto se han clasificado en tres grupos, acorde al planteamiento general de los mismos, a la metodología empleada y al personal científico encargado de ellos.

- 5 Por una parte se trabajará en la obtención de algoritmos eficientes para la obtención de campos meteorológicos análogos a partir de la base de datos con campos meteorológicos diarios ERA-15 o de la base de datos de NOAA. Se analizarán tanto las metodologías clásicas de clustering y las basadas en redes neuronales autoorganizativas, como las técnicas de cálculo de componentes principales no lineales para reducir el volumen de información, eliminando aquella redundante.
- 6 Posteriormente se utilizarán estos resultados para desarrollar modelos que relacionen los campos de salida de los modelos numéricos de predicción con la información histórica diaria procedente de los observatorios del INM y disponible en el área de interés para, de esta forma, poder predecir local y regionalmente los meteoros deseados. Los modelos desarrollados se implementarán en un Sistema Experto que se aplicará en la preparación de predicciones locales en diversas escalas temporales en aquellos observatorios de la red principal y termopluviométrica del INM en que se disponga de registros históricos diarios de las variables que desea predecir, siempre que dicha serie se extienda en el tiempo durante el mismo periodo en que están disponibles los campos meteorológicos de partida (entre 1979 y 1994 para los datos de ERA-15 y entre 1958 y 1998 para lo datos de NOAA).
- 7 Por otra parte, se analizará el problema de la predecibilidad en el corto plazo en áreas regionales mediante la aplicación de técnicas de “predicción por conjuntos” en modelos de mesoescala, comparando la eficiencia en este contexto de los diversos procedimientos actualmente en uso en predicción a plazo medio (breeding, vectores singulares, “lagged averaged method, LAF”). También se utilizará un “método del hombre pobre” basado en análogos como generador de perturbaciones aplicado a la predicción estacional. Sus resultados se compararán con los obtenidos en casos seleccionados del proyecto DEMETER, (Development of a European Multi-model Ensemble system for seasonal to inTERannual prediction), en el cual se estudia la capacidad de regionalizar en detalle la predicción estacional.

A continuación se describen con más detalle cada uno de los objetivos.

Ob1. Caracterización de Situaciones Atmosféricas Análogas.

El primer objetivo de este proyecto consiste en el desarrollo de modelos apropiados para caracterizar situaciones atmosféricas análogas a partir de uno o varios campos meteorológicos sobre la Península. En esencia se trata de un problema típico de agrupamiento (o clustering) pero hay que tener en cuenta el gran número de variables que componen cada campo y la gran redundancia existente. Por tanto, también será necesario prestar atención al problema de la reducción de variables (feature extraction) con la eliminación de variables redundantes. Se plantean dos objetivos:

Ob1.1. Selección de Campos Meteorológicos Análogos mediante Redes Neuronales.

Para el problema del agrupamiento de campos análogos (con una resolución espacial determinada), o más en particular para la obtención de un conjunto de campos análogos a uno dado, se analizarán las técnicas clásicas de clustering y también se analizarán las

posibilidades de las Redes Auto-Organizativas (SOM) como extensiones de las técnicas de k-means estándar. Para ello se estudiará el coste y la viabilidad de los distintos algoritmos de aprendizaje aplicados a las características de este problema concreto; se utilizarán los campos diarios disponibles en los recientes reanálisis del CEPPM y NWS, como patrones de entrenamiento.

Ob1.2. Componentes Principales no Lineales en Campos Meteorológicos.

De forma paralela al objetivo anterior, y con objeto de reducir el ingente número de variables que componen cada campo meteorológico perdiendo el mínimo de información posible, se analizarán la eficiencia de las distintas técnicas de reducción de dimensión en este problema. Se prestará especial atención al análisis de componentes principales no lineales, utilizando redes neuronales multicapa con “cuello de botella”, o redes funcionales con restricciones, que permiten obtener componentes no lineales óptimas bajo ciertas hipótesis.

Los métodos de reducción de dimensión obtenidos en Ob1.2 serán aplicados a los modelos desarrollados en Ob1.1 para desarrollar algoritmos eficientes para el cálculo de situaciones meteorológicas análogas a partir de las componentes principales obtenidas para los campos dados. Estos algoritmos supondrán una mejora objetiva de los métodos actualmente implementados en una versión preliminar de sistema experto que utiliza técnicas heurísticas para el cálculo de análogos.

Ob2. Mejora de la Predicción Local en Diferentes Escalas Temporales.

El segundo objetivo del proyecto es desarrollar métodos híbridos de downscaling que utilicen tanto las salidas de los modelos numéricos como los registros históricos procedentes de los observatorios y de los reanálisis de los modelos. Estos métodos serán implementados como parte de un sistema experto que pueda ser finalmente instalado y utilizado de forma operativa en el INM y que permita ganar detalle y precisión a las salidas de los modelos de predicción que el INM utiliza para la predicción a corto y medio plazo y para la predicción estacional de anomalías climáticas.

Ob2.1. Estudio de Métodos Híbridos Numérico-Estadísticos de Predicción.

Para desarrollar métodos eficaces para la mejora de resolución en las predicciones se utilizarán métodos híbridos que combinen las distintas técnicas que han sido utilizadas hasta la fecha para este propósito, principalmente, modelos numéricos, técnicas estadísticas y técnicas de análisis de series temporales (como se comentó en la introducción de este proyecto). En el modelo híbrido que se plantea en este proyecto un conjunto de campos análogos al campo previsto será el punto de partida (este conjunto de campos se obtendrá aplicando los algoritmos desarrollados en Ob1). El conjunto de días obtenidos servirá para aproximar la relación que presentan las variables de ese conjunto de campos con los valores horarios y diarios de las variables y la ocurrencia de determinados meteoros en el área de interés; en un trabajo preliminar se ha observado la conveniencia de considerar modelos no lineales para modelizar estas relaciones. En este problema las redes neuronales jugarán el papel de modelos genéricos de “ajuste”, mientras que las redes funcionales permitirán obtener modelos específicos basados en la información cualitativa disponible sobre las relaciones entre los campos y los meteoros de interés. Así mismo, un modelo autoregresivo (lineal o no lineal) de la evolución de los campos análogos permitirá aplicar de nuevo la técnica anterior para obtener una predicción local en una escala temporal superior. Estos modelos híbridos ampliarán y perfeccionarán los utilizados durante los dos últimos años en el marco de la colaboración INM Universidad de Cantabria.

Ob2.2. Desarrollo de Sistema Experto y Aplicación Operativa en Distintas Escalas Espacio-Temporales.

Los métodos híbridos desarrollados en el apartado anterior serán implementados a modo de motor de inferencia de un Sistema Experto que también incluirá una base de conocimiento formada por las bases de datos de los observatorios del INM y de los campos diarios de los

modelos de los reanálisis utilizado. Posteriormente, para su aplicación operativa, el sistema será adaptado a las distintas características de los distintos modelos que se emplean en el INM para la predicción a distintas escalas temporales:

Ob2.2.1. *Uso del sistema experto en la predicción a corto plazo.*

Para la predicción a corto plazo el INM utiliza el modelo mesoescalar HIRLAM, modelo de área limitada. Dicho modelo se integra hasta 48 horas cuatro veces al día sobre una rejilla con resoluciones de 50 y 20 Km. Sus salidas se utilizan de manera directa, por ejemplo, para el cálculo de oleaje. Por otra parte, para adaptar las salidas a otros puntos que no estén sobre la rejilla se aplican correcciones basadas en filtros de Kalman o simples ajustes de elevación, con buenos resultados en la predicción de temperaturas. El sistema experto desarrollado podrá aplicarse en cualquier punto de la red termoplumiométrica del INM, que consta de varios miles de estaciones, aunque la información observada en ellos no esté disponible en tiempo real (el método de los filtros de Kalman sí tiene este requerimiento). Se investigará su posible aplicación a la predicción de la ocurrencia de meteoros como lluvia, nieve, niebla e incluso a la predicción de la cantidad de precipitación. Cabe volver a indicar aquí que el funcionamiento de un prototipo de sistema experto más elemental aplicado a las variables y meteoros antes mencionados para observatorios de la Vertiente Cantábrica dió resultados muy alentadores.

Ob2.2.2. *Uso del sistema experto en la predicción a medio plazo.*

El Centro Europeo de Predicciones a Plazo Medio produce diariamente una predicción determinista con un modelo operado con una resolución de 60 Km. y que se integra hasta diez días. Los campos producidos están disponibles en el INM unas ocho horas después del momento nominal en que se realizan las observaciones. Por tanto, el sistema expeto podrá aplicarse en este caso día a día igual que como se ha descrito en Ob2.2.1.

Por otra parte el CEPPM realiza diariamente una predicción por conjuntos, hasta diez días, con 50 miembros que se obtienen de la integración del mismo modelo utilizado en la predicción determinista (aunque con resolución de 120 Km.) a partir de un conjunto de condiciones iniciales obtenidas con el método de vectores singulares. Una selección de esta información está disponible en el plazo de ocho horas en el INM. El grueso de la información producida se mantiene en el CEPPM. En este caso, el sistema experto se aplicará a cada uno de los cincuenta miembros del conjunto con lo que para cada plazo de predicción (día tercero, cuarto,...) los análogos encontrados podrán ser varios cientos (con las naturales repeticiones). Para una explotación operativa posiblemente convenga que la aplicación se instale y opere, al menos parcialmente, en los ordenadores del CEPPM donde el INM dispone de tiempo disponible para ello. También en la fase de pruebas se requerirá instalar la aplicación en alguno de los equipos del CEPPM.

Ob2.2.3. *Uso del sistema experto en la predicción estacional de anomalías climáticas.*

También se aplicará el sistema experto en la predicción estacional, aplicado a cada uno de los treinta miembros del conjunto utilizados para la preparación de cada predicción del CEPPM. Cada uno de estos miembros corresponde actualmente a una integración diaria hasta 180 días de un modelo acoplado océano-atmósfera operada con resolución horizontal de 200 Km. Las anomalías trimestrales previstas sobre España se estimarán a partir de los resultados del “downscaling” realizado y por comparación con una climatología convenientemente adpatada. También para esta parte de la investigación se requerirá que la aplicación esté instalada en los equipos del CEPPM.

Ob2.2.4 *Uso potencial del sistema experto en la preparación de escenarios climáticos.*

Potencialmente, el método de downscaling propuesto es también de aplicación a las salidas de los modelos de Clima. Queda fuera del ámbito del proyecto realizar una aplicación sistemática del método a, por ejemplo, sendas ventanas de diez años correspondientes al clima actual y al clima de finales del siglo XXI, procedentes de integraciones realizadas en el Centro Hadley. No obstante, sí incluirá la puesta a punto

del sistema experto para su ensayo y posterior utilización por otro grupo del INM que trabaja en la preparación de escenarios climáticos a partir de otras técnicas de downscaling estadístico.

Ob3. Predecibilidad Atmosférica en Áreas Regionales.

El tercer objetivo del proyecto se centrará en el problema de la predicción por conjuntos en áreas regionales (modelos de mesoescala), analizando los métodos propuestos en la literatura y prestando especial atención a la representatividad que tiene el conjunto de miembros seleccionados con respecto a la incertidumbre inicial del análisis; también se estudiarán las posibilidades y las limitaciones de los métodos basados en análogos, analizándose también la posibilidad de aplicar las técnicas híbridas descritas en Ob2.1 en este campo, que nos permitirán desarrollar un nuevo modelo híbrido de predicción por conjuntos basado en la evolución de configuraciones iniciales análogas.

Ob3.1. Predicción por Conjuntos a Corto y Medio Plazo.

Nos centraremos en el problema de la predicción por conjuntos en áreas regionales, analizando, en una primera etapa, la eficiencia de los distintos métodos propuestos en la literatura (principalmente los métodos de breeding y vectores singulares), tratando de aportar algún resultado en el debate que existe actualmente sobre este tema. Para ello se trabajará, inicialmente, con modelos de baja dimensión para analizar los distintos aspectos teóricos que posteriormente serán aplicados a modelos operativos sencillos de mesoescala para mostrar su utilidad práctica. Finalmente, se analizarán las posibilidades de los métodos basados en análogos, valorando la calidad de los resultados obtenidos teniendo en cuenta su bajo coste operativo.

Ob3.2. Análogos de Temperaturas Oceánicas para la Predicción Estacional.

Dada la gran influencia que tienen las temperaturas de los océanos Atlántico y Pacífico en la caracterización de los ciclos estacionales, se plantea el desarrollo de un modelo de predicción por conjuntos basado en los análogos obtenidos a partir de la distribución de temperaturas en los océanos. Cada miembro del conjunto será el promedio trimestral de los valores de los elementos meteorológicos asociados a situaciones sucesivas, para cada situación análoga encontrada. Este estudio permitirá contrastar la validez, utilidad y valor económico de los modelos basados en técnicas perturbativas (las que se aplican en el marco del Proyecto DEMETER, que se desarrolla en el CEPPM y en el que el INM participa, y en el que se integran seis modelos océano-atmósfera durante 120 días sobre un período de treinta años y cuyas condiciones iniciales corresponden a nueve situaciones para cada una de las cuatro estaciones del año) en comparación con las de un modelo mucho más económico, antes esbozado y que también pretende introducir una incertidumbre inicial.

Por último, mencionar que los objetivos que se proponen en este proyecto, se encuadran naturalmente en el Programa Nacional de Recursos Naturales, dentro del objetivo prioritario "1. Atmosfera y Clima" que, por ejemplo, en su punto 1.2 prioriza "el desarrollo y validación de técnicas de predicción climática para áreas geográficas de interés nacional en distintas escalas temporales".

3.4. BENEFICIOS DEL PROYECTO

En las propuestas de las **modalidades P2 y P3** expóngase cuál puede ser su utilidad para los sectores productivos o las áreas socioeconómicas correspondientes. Y añádase cuál puede ser el beneficio según la cobertura (nacional, local, sectorial) del proyecto.

Si en el proyecto participan otras Entidades, destáquense los beneficios que presumiblemente se derivarán para cada uno de los participantes.

Este proyecto de investigación se dirige al importante sector medioambiental de la predicción meteorológica y climatológica, y se desarrolla conjuntamente por un grupo interdisciplinar de investigadores de la Universidad de Cantabria, el CSIC, y el INM. Por tanto, un primer beneficio del proyecto es el importante intercambio de conocimientos que se produciría entre investigadores de ámbitos afines, con el consiguiente enriquecimiento de cada uno de los centros públicos de investigación participantes (Universidad de Cantabria y CSIC) y del Instituto Nacional de Meteorología (INM).

Por otra parte, dada la originalidad de los objetivos propuestos en este proyecto, la consecución y divulgación de los mismos situaría al equipo investigador y, como consecuencia al INM, en una posición inicial privilegiada en el ámbito de las técnicas de predicción híbridas, tanto a nivel de investigación básica, como a nivel operativo (con la implantación y uso de las aplicaciones desarrolladas en el INM).

Así mismo, dado que en este proyecto se aborda el tema de predicción local a medio y largo plazo, los resultados obtenidos pueden tener un claro beneficio para diversos sectores socioeconómicos que se podrían servir de pronósticos más localizados para planificar sus actuaciones ó para desarrollar políticas de prevención de riesgos naturales.

Finalmente, por su temática, objetivos y por el grupo de investigación involucrado (que incluye investigadores del departamento I+D del INM) este proyecto tiene el valor añadido de poder considerarse simultáneamente como:

- un proyecto de I+D, al profundizar en temas de investigación tratados en proyectos de investigación básica obtenidos en convocatorias públicas anteriores, aplicando los conceptos en problemas reales de gran interés;
- un proyecto de transferencia de resultados de I+D, al considerar como objetivo fundamental el desarrollo de una herramienta operativa, que podrá sumarse a las herramientas de predicción de que actualmente dispone el INM, y que podrá contribuir de forma importante en la elaboración de predicciones regionales y locales.

3.5. ACTIVIDAD DEL GRUPO SOLICITANTE

Este apartado, junto con el 2.4 anterior, tiene como finalidad determinar la capacidad del grupo en el tema y, en consecuencia, la viabilidad de la actividad propuesta. Por tanto, se deberán indicar con claridad los siguientes aspectos:

Si el proyecto es continuación de otro previamente financiado, los objetivos ya logrados y los resultados alcanzados.

Si el proyecto aborda una nueva temática, los antecedentes o resultados previos del grupo, que justifiquen su capacidad para llevar a cabo la nueva propuesta.

El equipo de investigadores que integra este proyecto está formado por investigadores de tres instituciones públicas distintas:

- José Manuel Gutiérrez y Antonio Galván, profesores numerarios de la Universidad de Cantabria, junto con Antonio S. Cofiño, becario predoctoral de Investigación en la Universidad de Cantabria.
- Miguel Angel Rodríguez, investigador del CSIC en el Instituto de Física de Cantabria (Instituto de investigación mixto CSIC-Universidad de Cantabria), y
- Rafael Cano y Juan José Ayuso, meteorólogos del Instituto Nacional de Meteorología (INM).

Por una parte, R. Cano está adscrito desde hace seis años al área de I+D del INM, trabajado durante los últimos años, entre otras cosas, en el desarrollo de un prototipo de sistema experto para el pronóstico meteorológico local a medio plazo utilizando técnicas heurísticas y lineales (ver [Can96]); este sistema constituye la base inicial para el desarrollo de las nuevas técnicas que se plantean en este proyecto. En distintos períodos del desarrollo de este trabajo contó con la colaboración del meteorólogo J.J. Ayuso, predictor numérico principal del Servicio de Modelización Numérica del tiempo, que pertenece a la subdirección de I+D del INM; por otra parte, J.J. Ayuso tiene una amplia experiencia en la aplicación de técnicas estadísticas en el ámbito de la meteorología operativa (ver, por ejemplo [Ay94]) siendo importante su participación en la implantación y verificación del sistema MOS de adaptación estadística de los salidas de los modelos de predicción numérica del tiempo (ver [AA91], [AR92] y [Ay92]); posteriormente se ha dedicado a la asimilación de la información en los modelos numéricos trabajando dentro del grupo europeo "High Resolution Limited Area Modelling (HIRLAM)" con especial dedicación a la asimilación de la humedad del suelo (ver [Ay95], [Ay96a], [Ay96b] y [Ay98]). Actualmente está iniciando en el INM los estudios sobre la predecibilidad del modelo HIRLAM, participando como investigador de estos temas en el plan científico HIRLAM5 para el período 2000-2003, también es uno de los investigadores propuesto por el INM para participar en el proyecto europeo "Limited-area Mesoscale Ensemble Prediction (LIMES)" presentado al V programa marco de la U.E.

Este trabajo preliminar sirvió de punto de partida para la colaboración establecida hace dos años con J.M. Gutiérrez y M.A. Rodríguez, en la Universidad de Cantabria, en el seno de un convenio marco establecido entre el INM y la Universidad de Cantabria, y mediante el cual esta última institución concedía un becario de investigación predoctoral, Antonio S. Cofiño, para trabajar en el marco de esta colaboración, desarrollando actualmente su tesis doctoral bajo la dirección conjunta de J.M. Gutiérrez y M.A. Rodríguez.

Fruto de dicha colaboración, han sido las distintas mejoras del prototipo de sistema experto inicial depurando las técnicas lineales clásicas implementadas. También se ha comprobado la utilidad del uso de técnicas no lineales, como las redes neuronales, en su mejora (ver [GCCR99], [CGCR99]), justificando así los objetivos planteados en este proyecto. La colaboración establecida no sólo ha tenido repercusión en el ámbito científico, sino que también se han divulgado los resultados obtenidos mediante la impartición de diversos cursos en distintos ámbitos. Por ejemplo, el curso "Introducción a la Inteligencia Artificial en Meteorología Operativa" fue impartido dentro de los Cursos de Formación del Instituto Nacional de Meteorología y el curso "Redes Neuronales y Probabilísticas en Modelización y Predicción" fue impartido dentro de los XV Cursos de Verano en Laredo de la Universidad de Cantabria.

El investigador principal del proyecto, J.M. Gutiérrez, es profesor titular de la Universidad de Cantabria y tiene una amplia experiencia en el desarrollo y aplicación práctica de sistemas expertos basados en redes probabilísticas, habiendo colaborado en distintos proyectos de

investigación en esta línea obtenidos en convocatorias públicas (PB94-1056, TIC96-0580). Desarrolló su tesis doctoral en este tema resolviendo algunos problemas relacionados con la especificación de modelos probabilísticos y con las técnicas de propagación simbólica, recopilándose los resultados obtenidos en el libro [CGH97], reconocido internacionalmente como un libro de referencia en este campo y citado en numerosas publicaciones científicas. Posteriormente, su interés se derivó al estudio de las redes neuronales como herramientas de modelización no lineal, desarrollando recientemente en colaboración con otros investigadores una generalización de las mismas mediante las denominadas "redes funcionales" que permiten la incorporación de conocimiento tanto cuantitativo, como cualitativo, en la estructura de la red (ver [CG98a,b], [CCGP99a,b], [CG00], [CCGP00]). Es éste uno de los temas que actualmente centran su interés y el que está directamente relacionado con este proyecto.

Otra de las líneas de investigación relacionadas con el proyecto en las que ha trabajado J.M. Gutiérrez colaborando, entre otros, con M.A. Rodríguez es en el ámbito de los sistemas dinámicos no lineales. Entre las distintas aportaciones realizadas en este campo, destacan por su relevancia con este proyecto la caracterización de medidas multifractales que permiten analizar la dinámica de fluidos en regímenes turbulentos (ver, por ejemplo, [GR99], [GIR96]), y el desarrollo de herramientas de análisis de sistemas dinámicos (ver, por ejemplo, [GIGM96], [GI98]).

Es también esta línea de trabajo la que ha ocupado la actividad de A. Galván (ver, por ejemplo [GG97]), especializado en el estudio de fenómenos caóticos y sus aplicaciones en mecánica y química.

M.A. Rodríguez es un investigador del CSIC de reconocido prestigio en el ámbito de la Física Estadística y los Sistemas No Lineales. Inicialmente, sus trabajos se centraron en el análisis de sistemas con ruido, realizando importantes contribuciones en el campo ([RP85], [VRP91], [HR90], [GIR93]), que incluyen la aplicación en la caracterización de fluctuaciones en reactores nucleares (ver, por ejemplo, [MR85], [RP87], [RSS87]) y en láseres ([NPR91], [VRP93]). Actualmente su interés se centra en distintos aspectos de los medios desordenados, incluyendo el crecimiento de interfases (ver [LRC97] y [LRR99]) y la modelización de sistemas caóticos y sus aplicaciones. M.A. Rodríguez ha participado en numerosos proyectos de investigación y ha sido investigador principal de los proyectos PB93-0054-C02-02 (financiado por la DGICYT) y PB96-0378-C02-02 (correspondiente a la DGES Networks activity of the training and mobility researches programme ERB4061PL95), abordando la dinámica y el desorden de estructuras espaciotemporales fuera del equilibrio.

En cuanto a las áreas de conocimiento involucradas en este proyecto, cabe destacar que, aparte de la modelización numérica de la circulación atmosférica (donde R. Cano y J.J. Ayuso cuentan con gran experiencia), algunos aspectos de la Inteligencia Artificial, en especial los modelos de redes neuronales y funcionales, así como el área de la dinámica no lineal serán de suma importancia para el desarrollo del proyecto, ya que en estas áreas del conocimiento residen los fundamentos para la caracterización de situaciones atmosféricas análogas y el desarrollo de los modelos híbridos de predicción (Ob1 y Ob2) y el estudio de la predecibilidad de los modelos de predicción atmosférica (Obj3), respectivamente; en este caso la experiencia previa, tanto teórica como aplicada, de los investigadores de este grupo permite afrontar el proyecto con suficientes garantías.

En cuanto al grado de cohesión existente en el equipo ya se ha mencionado que, aunque los investigadores pertenecen a tres organismos distintos, se pueden considerar agrupados en dos equipos de investigación: uno integrado por J.M. Gutiérrez, M.A. Rodríguez y A. Galván, que han colaborado en distintos proyectos y trabajos a lo largo de los últimos seis años relacionados con los sistemas no lineales y su modelización y con las técnicas de predicción basadas en redes neuronales; otro integrado por R. Cano y J.J. Ayuso, en el ámbito de la meteorología operativa, aplicando distintas técnicas estadísticas en el problema de la predicción meteorológica y climatológica. Ambos equipos de investigación han tenido la experiencia de trabajo en común durante los dos últimos años, trabajando en el ámbito de un acuerdo marco firmado entre la

Universidad de Cantabria y el INM. Por tanto, tanto la experiencia previa, como el grado de cohesión del equipo son óptimos para la adecuada realización de este proyecto.

Referencias:

- [AA91] R. Azcarraga, and J.J. Ayuso. "Statistical system for forecasting in Spain". In Lectures presented at the WMO training workshop on the interpretation of NWP products, 45-50 (1991).
- [AR92] J.J. Ayuso y P. Del Rio. "Adaptación estadística MOS para predicción del viento en superficie". *Jornadas Científicas A.M.E.*, XXIII:III, 157-162 (1992).
- [Ay92] J.J. Ayuso, "Método de Predicción Perfecta utilizando los Análisis del LAM del INM", *Proceedings del III Simposio Nac. de Predicción del INM*, 227-230 (1992).
- [Ay94] J.J. Ayuso, "Predicción Estadística Operativa en el INM". *Monografías del Instituto Nacional de Meteorología*, volumen B-34 (1994).
- [Ay95] J.J. Ayuso, "Soil moisture assimilation in HIRLAM". In *HIRLAM 3 Workshop on Soil Processes and Soil/Surface Data Assimilation*, 45-50 (1995).
- [Ay96a] J.J. Ayuso, "Asimilación de la humedad del suelo en modelos numéricos". *Proceedings del IV Simposio Nac. Predicción del INM*, 497-504 (1996).
- [Ay96b] J.J. Ayuso, "Some experiments with soil moisture assimilation". In *HIRLAM 3 Workshop on Physical Parametrizations*, 119-124 (1996).
- [Ay98] J.J. Ayuso, J. "Tests of isba scheme and soil moisture assimilation in a summer case". In *HIRLAM 4 Workshop on Physical Parametrizations*, 151-156 (1998).
- [Can96] R. Cano "Sistema experto en interpretación de salidas CEPPM y HIRLAM". *Proceedings del IV Simposio Nacional de Predicción del INM*, 299-307 (1996).
- [CG98a] E. Castillo and J.M. Gutiérrez, "A Comparison of Functional Networks and Neural Networks". *Proceedings of the IASTED International Conference on Artificial Intelligence* (Hamza, M. H. editor) IASTED/ACTA press. 439-442 (1998).
- [CG98b] E. Castillo and J.M. Gutiérrez, "Nonlinear Time Series Modeling and Prediction Using Functional Networks. Extracting Information Masked by Chaos", *Physics Letters A*, Vol. 244, 71-84 (1998).
- [CG00] E. Castillo and J.M. Gutiérrez, "A Minimax Method for Learning Functional Networks", *Neural Processing Letters*, Vol. 11(1), 39-49 (2000).
- [CGH97] E. Castillo, J.M. Gutiérrez and A.S. Hadi, "Expert Systems and Probabilistic Network Models", Springer-Verlag: New York, USA (1997).
- [CCGP99a] E. Castillo, A. Cobo, J.M. Gutiérrez and E. Pruneda. "An Introduction to Functional Networks with Applications", Kluwer Academic Publishers, Boston, USA (1999).
- [CCGP99b] E. Castillo, A. Cobo, J.M. Gutiérrez, and E. Pruneda "Working with Differential, Functional and Difference Equations using Functional Networks", *Applied Mathematical Modelling*. Vol 23, 89-107 (1999).
- [CCGP00] E. Castillo, A. Cobo, J.M. Gutiérrez and E. Pruneda "Functional Networks: A New Network-Based Methodology", *Computer-Aided Civil and Infrastructure Engineering*. Vol 15, 90-106 (2000).
- [CGCR99] R. Cano, J.M. Gutiérrez, A.S. Cofiño y M.A. Rodríguez "Redes Neuronales y Patrones de Analogías Aplicados al Downscaling en Modelos Climáticos", en *La Climatología Española en los Albores del Siglo XXI* (Actas del I Congreso de la Asociación Española de Climatología). J.M. Raso y J. Martín-Vide editores, 113-121 (1999).

- [GG97] F. Gascón and A. Galván, "Obtaining Fractals in the Fabry-Perot Interferometer", *Fractals*, Vol 5, 187-198 (1997).
- [GCCR99] J.M. Gutiérrez, R. Cano, A.S. Cofiño and M. A. Rodríguez "PROMETEO: Un Sistema Experto para el Pronóstico Meteorológico Local basado en Redes Neuronales y Cálculo de Analogías", *Actas de III Jornadas de Transferencia Tecnológica de Inteligencia Artificial*, 11-19 (1999).
- [GR99] J.M. Gutiérrez and M. A. Rodríguez "A New Exact Method for Obtaining the Multifractal Spectrum of Multinomial Measures and IFSP Measures", *Chaos, Solitons and Fractals*, Vol 11, 675-683 (2000).
- [GIR96] J.M. Gutiérrez, A. Iglesias, and M.A. Rodríguez "A Multifractal Analysis of IFSP Measures: Application to Fractal Image Generation", *Fractals*, 5(1), 17-27 (1996).
- [GI98] J.M. Gutiérrez and A. Iglesias, "A Mathematica Package for the Analysis and Control of Chaos in Nonlinear Systems", *Computers in Physics*, Vol. 12(6), 608-619 (1998).
- [GIR93] J.M. Gutiérrez, A. Iglesias y M.A. Rodríguez "Logistic Map Driven by Dichotomous Noise", *Physical Review E*, 48(4), 2507-2513 (1993).
- [GIGM96] J.M. Gutiérrez, A. Iglesias, J. Guémez and M.A. Matías, "Suppression of Chaos Through Changes in the System Variables and Numerical Rounding Errors", *Chaos, Solitons and Fractals*, Vol. 7(8), 1305-1316 (1996).
- [HR90] E. Hernández, M.A. Rodríguez, L. Pesquera, and M. San Miguel "Calculation of Transport Properties for One-dimensional Random Walk in presence of Weak and Strong disorder". *Physical Review B*, Vol. 42, 19653 (1985).
- [LRC97] J.M. López, M.A. Rodríguez, and R. Cuerno "Power Spectrum Scaling in Anomalous Kinetic Roughening of Surfaces". *Physica A*, Vol. 246, 329 (1997).
- [LRR99] J.M. López, J.J. Ramasco, and M.A. Rodríguez "Comment on Macroscopic Equation for the Roughness of Growing Interfaces in unched Disorder". *Physical Review Letters*. In press.
- [MR85] I. Martínez and M.A. Rodríguez, "Some Exact Results for the Mean Neutron Density Induced by an Absorber with Random Motion". *Annals of Nuclear Energy*, Vol. 12, 113 (1985).
- [NPR91] J.M. Noriega, L. Pesquera, and M.A. Rodríguez and "Intensity Correlation Functions of Dye Lasers with Gain Noise". *Physical Review A*, Vol. 43, 4008 (1991).
- [RP87] M.A. Rodríguez and L. Pesquera, "Stability Analysis of Linear Reactor Systems with Reactivity Fluctuations". *Journal of Nuclear Science and Technology*, Vol. 21, 797 (1984).
- [RP85] M.A. Rodríguez, L. Pesquera, M. San Miguel and J.M. Sancho, "Unified Description of Internal Fluctuations and External White Poisson Noise". *Physics Letters A*, Vol. 107, 443 (1985).
- [RSS87] M.A. Rodríguez, M. San Miguel and J.M. Sancho, "Theory of Fluctuations and Parametric Noise in a Point Nuclear Reactor Model". *Annals of Nuclear Energy*, Vol. 10, 263 (1983).
- [VRP91] A. Valle, M.A. Rodríguez, and L. Pesquera "Diffusion in a Continuous Medium with Space Correlated Disorder". *Physical Review A*, Vol. 43, 948 (1991).
- [VRP93] A. Valle, M.A. Rodríguez, and L. Pesquera "Statistical Properties of Pulses: Application to Modulated Gas Lasers". *Physical Review A*, Vol. 47, 4176 (1993).

3.6. METODOLOGÍA Y PLAN DE TRABAJO

Se debe detallar y justificar con precisión la metodología que se propone y debe exponerse la planificación temporal de las actividades, incluyendo cronograma (se adjunta modelo).

Como ya se ha comentado anteriormente el equipo de investigación está integrado por los investigadores:

- *José Manuel Gutiérrez (JG)*,
- *Antonio Galván (AG)*,
- *Miguel A. Rodríguez (MR)*,
- *Rafael Cano (RC)*, y
- *Juan José Ayuso (JA)*

Este equipo ha tenido una experiencia de trabajo en común, durante los dos últimos años, trabajando en un proyecto de investigación dentro de un acuerdo marco entre la Universidad de Cantabria y el INM, que ha servido para comprobar la viabilidad del proyecto que aquí se presenta y las tareas que es necesario realizar para la consecución de los objetivos marcados. Fruto de esta colaboración se ha conseguido una beca ocupada por:

- *Antonio S. Cofiño (AC)*,

miembro integrante del equipo investigador que realiza su tesis doctoral en algunos de los temas abordados en este proyecto y cuya beca finaliza el año siguiente al comienzo del proyecto, participando por tanto el primer año del mismo (como se indica en el cronograma adjunto). Por otra parte, también incluimos dentro del equipo investigador el contratado solicitado para los tres años de duración del proyecto:

- *Licenciado contratado (CO)*

como se ve en el cronograma adjunto, CO estará dedicado a tiempo completo al desarrollo de algoritmos y de la aplicación informática (el sistema experto) que se persigue en este proyecto siendo, por tanto, pieza fundamental del mismo. La complejidad del software que se pretende desarrollar, junto con la necesidad de trabajar con distintos formatos estándar para el manejo de ficheros y bases de datos, hace imprescindible poder contar con una persona con experiencia que centralice toda la actividad.

Una vez presentados de los objetivos de este proyecto, es bastante inmediato establecer un plan de trabajo que lleve a la consecución de los mismos (ver cronograma adjunto para más detalles sobre la formación de los distintos grupos y de las tareas a desarrollar):

En la primera fase del proyecto (primer año y medio) se establecerán tres grupos de trabajo paralelos para los tareas: desarrollo de modelos basado en redes neuronales para la obtención de situaciones meteorológicas análogas (**T1**); análisis de componentes principales en campos meteorológicos (**T2**); y análisis de métodos de predicción por conjuntos en áreas regionales (**T7**).

Como ya se ha comentado en la presentación de los objetivos, las tareas T1 y T2 utilizarán principalmente modelos de redes neuronales y funcionales; la utilidad de estas técnicas en problemas similares ha sido sobradamente probada en la literatura. Por ejemplo, Lener y otros han publicado recientemente una comparación de distintas técnicas para caracterización y reducción de información en Pattern Recognition Letters 20, 1999, mostrando las ventajas de estos modelos. En este caso, se habrá de disponer de software con el que poder analizar modelos de redes neuronales en los problemas planteados. El gran volumen de datos que es necesario manejar hace imprescindible poder contar con software profesional, como se describe en la propuesta de gastos del proyecto.

También será necesario poder contar con dos ordenadores personales con gran capacidad de cálculo y almacenamiento para poder llevar a cabo las tareas programadas de forma apropiada.

Actualmente no se dispone de ordenadores personales para asignar a este proyecto por lo que también es imprescindible poder contar con ellos para la realización del mismo.

Las tareas T1, T2 y T7 requerirán un amplio estudio bibliográfico previo sobre cada uno de los temas, siendo esta parte del trabajo especialmente importante en el grupo T7 dada la variedad y complejidad de los métodos presentado en la literatura para la predicción por conjuntos. Será de gran utilidad poder contar con la visita de algún investigador especializado en estos temas para resolver cuestiones de especial importancia para este proyecto, así como poder realizar alguna estancia breve en algún centro extranjero para adquirir experiencia en alguno de estos temas. Por ello se ha solicitado una partida de gastos en esta línea en la propuesta de gastos del proyecto.

También de forma paralela el primer año se comenzará con el diseño y desarrollo del sistema experto (T3) que integrará los algoritmos y métodos desarrollados en las distintas etapas de este proyecto. Durante el primer año se analizarán otras aplicaciones meteorológicas operativas para decidir la conveniencia de utilizar los distintos formatos (predicciones y análisis de los modelos numéricos y registros históricos de estaciones) y las distintas herramientas de visualización científica que habrán de integrarse en la aplicación (Grads, Vis5D, etc.). En esta etapa también será necesario realizar alguna estancia breve en otros centros de investigación (principalmente en el Centro Europeo en Reading, UK) para resolver alguno de los conflictos que puedan surgir en el diseño de la aplicación. En este punto es imprescindible poder contar con una persona dedicada a tiempo completo a esta tarea desde el principio del proyecto y que centralice todo el trabajo de desarrollo informático del mismo. Para ello se ha solicitado un contrato por los tres años de duración de un profesional del área de la Informática.

En la segunda fase del proyecto, la actividad se centrará en torno a dos tareas. En la primera se tratarán de desarrollar métodos híbridos numérico-estadísticos basados en análogos para la predicción local (T5) que, posteriormente se implementará en el sistema experto (T6) que ya contará con los módulos de tratamiento de datos y de cálculo de análogos, este último desarrollado en la primera mitad del segundo año (T4), a partir de los resultados obtenidos por los dos primeros grupos de investigación.

Por otra parte, tras haber analizado las posibilidades de los métodos estándar de predicción por conjuntos, se procederá a aplicar las técnicas basadas en análogos en este problema (T8). Con estas técnicas se tratará de aprovechar la ventaja, respecto de los métodos estándar analizados, de la mayor representatividad que se tiene de la incertidumbre existente en la configuración inicial utilizando un conjunto de situaciones atmosféricas análogas.

Respecto a la financiación solicitada en este proyecto de investigación, dentro del capítulo de "Material inventariable" se contempla la adquisición de material bibliográfico (libros y revistas especializadas), la adquisición de material informático (dos ordenadores personales de gran capacidad de almacenamiento y cálculo y distintos periféricos (placas Ethernet, CDROM grabable, etc.) y la adquisición de diverso software estadístico y de modelización de redes neuronales, como se ha indicado anteriormente.

En el capítulo de "viajes y dietas", se contemplan los gastos para la interacción con investigadores de proyectos afines, que requiere la asistencia periódica a congresos y reuniones científicas. También, como ya se ha comentado, la realización de algunas tareas requerirán la estancia breve de algún investigador del grupo en un centro nacional o extranjero, incluyendo las estancias breves en el INM (Madrid), necesarias para una adecuada coordinación de los distintos aspectos del proyecto. Por otra parte, también es necesario poder contar con investigadores extranjeros especialistas en campos de interés.