

Non-linear modelling of monthly mean vorticity time changes: an application to the western Mediterranean

M. Finizio, S. Palmieri

University of Rome “La Sapienza”, Dep. of Physics, Piazzale A.Moro 2, I-00185 Rome Italy

Received: 14 January 1997 / Revised: 30 May 1997 / Accepted: 1 July 1997

Abstract. Starting from a number of observables in the form of time-series of meteorological elements in various areas of the northern hemisphere, a model capable of fitting past records and predicting monthly vorticity time changes in the western Mediterranean is implemented. A new powerful statistical methodology is introduced (MARS) in order to capture the non-linear dynamics of time-series representing the available 40-year history of the hemispheric circulation. The developed model is tested on a suitable independent data set. An ensemble forecast exercise is also carried out to check model stability in reference to the uncertainty of input quantities.

Key words. Meteorology and atmospheric dynamics · General circulation ocean-atmosphere interactions · Synoptic-scale meteorology

1 Introduction

Reliable meteorological outlooks of a few months, in spite of their unavoidable limitations lead to the provision of information representing mean conditions over rather large areas and are likely to have a huge impact on human activities, with special reference to agriculture. When large-scale patterns and extended time-ranges are considered, a considerable part of atmospheric variability is linked to long-period fluctuations. Predictability depends on slow oscillations or trends in low-layer forcing factors, possibly due to sea-surface-temperature (SST) anomalies and atmosphere-ocean interactions. When this type of slow change is considered, fast fluctuations lose their importance.

While in tropical regions, in the light of the present knowledge, these forcing factors appear as leading elements which may affect climate dynamics at seasonal scale, in extratropical areas fluid dynamical instabilities, together with non-linear interactions, lead to a much more complex variability framework (Palmer and Anderson, 1994; Livezey *et al.*, 1996). This variability is weaker in summer, allowing a higher predictability, but becomes larger in winter when tropical influences would be hindered by a more active action of planetary-scale waves. In this regard the best prediction results may be obtained in spring, as shown by Brankovic *et al.* (1994).

Quite a few investigations were also carried out to find teleconnections between northern-hemisphere atmospheric phenomena and low-level forcing factors. In the case of Europe correlations are not so high as in other areas (Palmer and Anderson, 1994). Moreover, Fraedrich and Müller (1993), analysing fields of pressure, temperature and rainfall (from 1880) over Europe, coupled with “El Niño-Southern Oscillation” (ENSO) events, stressed how the observed anomalies are about one-third of the climatological variability characterising this area. This kind of teleconnection is particularly evident in the months of February to May, following, with one years delay, the cold phase of ENSO (Halpert and Ropelewski, 1992). Nevertheless, apart from the SO effects, the role of SST is still not entirely clear.

Following a line originally proposed by Palmieri *et al.* (1995), the aim of the present work is to carry out a preliminary investigation of the role of various factors which may affect long-range predictability. It was necessary, for the sake of simplicity (required by the limited extension of time-series available), to adopt a scheme in which some important elements related to the boundary conditions (ocean-atmosphere interactions) are represented, such as:

1. low-level horizontal temperature gradients over Pacific and Atlantic Oceans;
2. the ENSO index.

These quantities are likely to contain the signal of fluctuations occurring in the range from monthly to

annual scale (10–1000 days) defined as “short-term climate variability” by Shukla (1991). On the other hand, it is well known that the interaction between boundary conditions (SST, geographical distribution of lands and oceans, surface hydrological cycle and albedo) are non-linear, due to a certain amount of feedback. In the described framework, with the intention of investigating an alternative approach in which model structure is derived by observables, an advanced statistical methodology able to capture non-linear effects, namely a recent version of the Multivariate Adaptive Regression Splines (MARS), initially introduced by Friedman (1991), is applied. This technique allows to seek non-linear relationships between atmospheric elements far away from one another both in space and time (Corte-Real *et al.*, 1995; Lewis and Stevens, 1991; Taliani *et al.*, 1996). The research goal is that of detecting possible teleconnections between some significant thermal patterns and relative vorticity at a given site. Starting from a number of observables in the form of time-series of meteorological elements, a model is implemented capable of fitting past records and predicting monthly vorticity time changes in the western Mediterranean.

2 Data set and atmospheric parameters

2.1 Data set

For this purpose a data set of northern-hemisphere monthly meteorological data (octagonal grid), provided by the USA National Meteorological Center (NMC – now NCEP), is used. The data set covers the period January 1946–June 1989. The considered atmospheric parameters are derived on the basis of these data.

2.2 Atmospheric parameters

Predictand: relative geostrophic vorticity at 500 hPa (V_o) at Balearic Islands (38°N, 3°E). This is a well-known key parameter representing the general synoptic-scale situation in the area: weather conditions over Spain, southern France, Italy and north-western Africa are closely linked to it.

Predictors: in order to allow a physical and dynamical interpretation of results, it was thought advisable to use quantities inspired by space derivatives of fields. No attempt was made to use statistical techniques, such as Principal Component Analysis, which might be effective in condensing the huge amount of information available but would on the other hand hamper the possible explanation of processes.

The following factors are considered:

1. North-Atlantic Oscillation (NAO), defined as the difference between the sea-level pressure (slp) anomalies over Azores and over Iceland.
2. Azores semi-permanent anticyclone (location and intensity), described by two parameters, the trigono-

metric tangent of its latitude (Tg) and the geostrophic vorticity at 500 mb (Azz).

3. North Pacific ENSO defined as the difference between slp anomalies at the western coast of Mexico (21°N, 117°W) and Manila (16°N, 120°E), denoted as $Paco$. This is in line with Philander (1983) in which the correlation field of monthly mean surface pressure over Pacific area is presented.
4. Meridional thermal horizontal gradients, represented by the 850-hPa temperature difference between high (60°N) and low (30°N) latitude over the Pacific (longitude 170°W) and North Atlantic area (longitude 80°W), indicated as Tmp and Tma , respectively.
5. Zonal flow intensities Flp and Fla (same latitudes and longitudes as thermal meridional gradients).

Since our aim is the development of a model capable of describing and possibly predicting monthly time changes in the West Mediterranean relative geostrophic vorticity, our attention is concentrated on the central month anomalies of each season. Input observational data go through the following steps: firstly, each time-series is standardised; then a 3 month running average is applied (weight: 0.25, 0.50 and 0.25).

2.3 Parameter sampling

The total number of monthly values of both predictand and predictors is 518 (about 44 years). The first 470 are used to create and calibrate the model, while the last 48, about 10% of the whole set, are used as an independent sample for a preliminary test of the predicting skill of the model.

2.4 Fluctuation range of parameters

In Table 1, the fluctuation ranges during the observational period of parameters used in the model are presented. DVo indicates the time changes in relative geostrophic vorticity at 500 hPa over the western Mediterranean (see Sect. 4). Its minimum and maximum are obtained in June and October, respectively. In winter the meridional temperature gradient anomalies over the North Pacific and North Atlantic areas (Tmp and Tma) are negative, according to the largest temperature difference between equator and pole in this season. Conversely, in summer the values tend to be positive. Noteworthy is the existence of an annual cycle of the meridional temperature gradient, Tma . Nevertheless, a careful analysis indicates the existence of other meaningful cycles, such as that of 24–26 months.

Table 1. Fluctuation range of parameters during the observational period

par.	min.	max.	par.	min.	max.
Dvo	−2.190	2.250	$Paco$	−2.192	2.873
Vo	−1.813	2.120	Tma	−2.166	1.681
Tg	−0.974	2.904	Tmp	−1.842	1.910

3 Statistical methodology: multivariate adaptive regression splines (MARS)

3.1 General remarks

The MARS methodology was introduced by Friedman in 1991 and has been applied in a wide range of disciplines. It allows to create non-linear regression models describing the relationship between the dependent variable (predictand) and the independent ones (predictors). In the next sections some basic concepts of this methodology, together with the innovations proposed by Friedman, are outlined. For an exhaustive description reference should be made to Friedman (1991) and Lewis and Stevens (1991).

3.2 Basic concepts

Let y represent a single response variable which depends on a vector of p predictor variables $\mathbf{x} \equiv (x_1, \dots, x_p)$. Suppose that N samples of y and \mathbf{x} , $\{y_i, \mathbf{x}_i\}_{i=1}^N$, are available, and that y may be described by means of a regression model, $y = f(\mathbf{x}) + \varepsilon$ over some domain $D \subset R^p$, where $f(\mathbf{x})$ describes the unknown relationship whose existence is assumed, and ε represents the random additive error which we assume without loss of generality to have zero mean and variance σ_ε^2 . The aim is to develop a function $\tilde{f}(\mathbf{x})$ that represents an acceptable approximation of this relationship. MARS is based on *recursive partitioning* (RP), using *spline* functions to fit $f(\mathbf{x})$. The RP procedure consists of two phases: a *forward step* and a *backward step*. In the first phase (Fig. 1) the domain D is subsequently divided into regions (parents) until the terminal division in S subregions (daughters) is obtained.

So let $\{R_j\}_{j=1}^S$ be a terminal set of S disjoint subregions such that $D = \bigcup_{j=1}^S R_j$. Then RP estimates the unknown function $f(\mathbf{x})$ in \mathbf{x} by means of

$$\tilde{f}(\mathbf{x}) = \sum_{j=1}^S c_j(\mathbf{x})B_j(\mathbf{x}),$$

where $B_j(\mathbf{x})$ is a function assuming the value 1 when $\mathbf{x} \in R_j$ and 0 otherwise. $c_j(\mathbf{x})$ are, usually, coefficients determined within the subregions. The backward step

trims the subregions in excess ($S - M$) which do not contribute remarkably to a good estimate of $f(\mathbf{x})$. At the end, a good set of non-overlapping subregions $\{R_j\}_{j=1}^M$ together with the functions $c_j(\mathbf{x})$ are obtained, providing a reasonable fit of $f(\mathbf{x})$ over the domain D .

3.3 Friedman's innovations

The RP methodology presents some difficulties. The subregions being disjoint, created models may produce discontinuities at subregion boundaries. Moreover, because the subregions are obtained by an iterative division of the original domain D , these models show evident difficulties to estimate additive and linear functions. Further, in increasing the dimension P beyond 2, the so-called "curse of dimensionality" may appear: in order densely to populate high-dimensional spaces, to a linear increase in p , a corresponding exponential increase in sample size N would be necessary.

To overcome these difficulties Friedman proposed to retain parent regions in the process in order to obtain overlapping subregions of domain. Moreover, to eliminate the discontinuity introduced by the step functions $B_j(\mathbf{x})$, they were replaced by linear regression splines in the form of left and right truncated splines:

$$B_j(\mathbf{x}) = \prod_{i=1}^{s_j} \left(s_{ij} \left(x_{k(i,j)} - x_{k(i,j)}^o \right) \right)_+.$$

Here, s_j is the number of divisions of domain D that realises the j -th subregion, $x_{k(i,j)}^o$ is the knot referred to the particular predictor $x_{k(i,j)}$ and s_{ij} can assume the value +1 or -1 when the subregion to the right or to the left of the knot is chosen. Finally, the plus sign indicates that this function assumes its true value only when its argument is positive, otherwise it gives no contribution. Each term of the model, referring to a particular subregion, is formed by a coefficient, together with a product of these unidimensional truncated spline functions. For example, referring to Fig. 1 and considering the subregion 4, the function will be:

$$B_4(\mathbf{x}) = (x_1^o - x_1)_+ \cdot (x_2 - x_2^o)_+.$$

In the forward step, the MARS algorithm looks for the set of subregions in order to determine the basis functions. This research of partition points (knots) and coefficients is necessary in order to find the model which best approximates $f(\mathbf{x})$. In the backward step, MARS carries out a trimming procedure to remove the terms of the model which do not remarkably contribute to the quality of the fit. In this study, in order to select the model and assess its accuracy, the Schwarz-Rissanen criterion (SC) was used:

$$SC(S) = \ln(\hat{\sigma}_\varepsilon^2) + \frac{\ln(N)C(S)^*}{N},$$

where $C(S)$ is a function of the model complexity and S is the number of subregions determined in the forward step. This formulation provides better results in a form more amenable to interpretation (Lewis and Stevens, 1991).

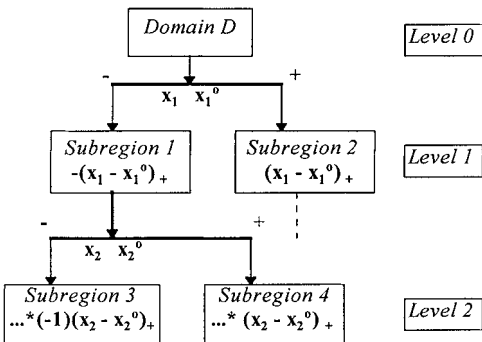


Fig. 1. Schematic reproduction of the formal step in RP methodology (x_k predictor, x_k^o Knot)

4 The proposed model

The model development is preceded by a preliminary phase with the intention of searching for correlations and finding existing cycles within the time-series analysed. On the basis of this information, some indications about potential predictors and their interactions are derived.

As already mentioned, the time change of relative geostrophic vorticity at 500 hPa over the West Mediterranean (DVo) is selected (difference between the value at month t and that at time $t - 2$): $DVo_t = Vo_t - Vo_{t-2}$. A 2-month backwards time differencing is introduced to damp a possible bias introduced by the smoothing process; for this reason predictors with a delay of 1 month have not been considered.

As potential predictors, Tma , Tg , $Paco$ and Tmp , defined in Sect. 2, are assumed. MARS internal parameters are chosen as follows:

1. model selection criteria, SC (Schwarz-Rissanen);
2. MI parameter is set to 3 (to limit the degree of interactions among predictors);
3. 80 basis points ($NK = 80$) are allowed (in order to reduce the number of useless subregions in excess created by the “forward-step” phase);
4. a value of 25 for the MS parameter (a good compromise between the possibility of developing a reliable model and the need to reduce the computational time).

The model equation is:

$$DVo_t = Vo_t - Vo_{t-2} = -1.807 + 0.794 \cdot (Vo_{t-4} + 1.813)_+ + \dots \quad (1)$$

$$- 0.125 \cdot (Vo_{t-4} + 1.813)_+ \cdot (Tma_{t-12} + 2.166)_+ + \dots \quad (2)$$

$$- 1.518 \cdot (Vo_{t-4} + 1.813)_+ \cdot (Tg_{t-2} + 0.263)_+ \cdot (Paco_{t-2} - 0.540)_+ + \dots \quad (3)$$

Table 2. Relative importance of terms

term	std. dev.	SC	predictors
1	0.6038	-0.7858	Vo_{t-4}
2	0.3349	-0.9255	Vo_{t-4}, Tma_{t-12}
3	0.1054	-0.9693	$Vo_{t-4}, Tg_{t-2}, Paco_{t-2}$
4	0.6509	-0.5878	Vo_{t-3}, Tma_{t-6}
5	0.2534	-0.8814	Vo_{t-3}, Tma_{t-12}
6	0.1131	-0.9699	Vo_{t-3}, Tg_{t-3}
7	0.2010	-0.9357	Tmp_{t-12}
8	0.1352	-0.9559	Tmp_{t-12}, Tma_{t-2}

$$+ 0.360 \cdot (0.853 - Vo_{t-3})_+ \cdot (Tma_{t-6} + 2.166)_+ + \dots \quad (4)$$

$$+ 0.284 \cdot (0.853 - Vo_{t-3})_+ \cdot (Tma_{t-12} + 0.941)_+ + \dots \quad (5)$$

$$+ 0.743 \cdot (0.853 - Vo_{t-3})_+ \cdot (-0.316 - Tg_{t-3})_+ + \dots \quad (6)$$

$$- 0.375 \cdot (Tmp_{t-12} + 0.119)_+ + \dots \quad (7)$$

$$+ 0.921 \cdot (Tmp_{t-12} + 0.119)_+ \cdot (Tma_{t-2} - 0.735)_+ \quad (8)$$

It includes one constant and eight terms, two of these with interaction level 1, five with level 2 and one with level 3. The knots or threshold values are 9: one for each of the following variables: $Vo_{t-3}, Vo_{t-4}, Tma_{t-2}, Tma_{t-6}, Tmp_{t-12}, Tg_{t-2}, Tg_{t-3}, Paco_{t-2}$, and two for Tma_{t-12} . The value obtained for SCs is -0.9711 , while the standard error of residuals is 0.025. An initial idea of the relative importance of each element may be obtained by considering the standard deviation, which represents the weight of model terms (Table 2, second column). A deeper interpretation involves the “deteriorated” SC values (Table 2, third column): they indicate to what extent the model deteriorates (as compared with the “global” SC value) when the corresponding term is eliminated. This type of estimate shows that the first, forth and fifth terms stand out as the most important

Table 3. Simultaneous conditions to be fulfilled in order that a given term of the model be “switched on”. In columns 2, 3, 4 and 5, sign, minimum, mean and maximum value of each contribution are reported

term	Sign	min. (month)	mean	max. (month)	cond. I	cond. II	cond. III
1	+	Nov Dec Jan	+1.41	Jun Jul Aug	$Vo_{t-4} > -1.813$		
2	-	Jun Jul Aug	-0.52	Nov Dec Jan	$Vo_{t-4} > -1.813$	$Tma_{t-12} > -2.166$	
3	-	Mar Apr May-Dec	-0.21	Jul Aug Sep	$Vo_{t-4} > -1.813$	$Tg_{t-2} > -0.263$	$Paco_{t-2} > 0.540$
4	+	Jun Jul Aug	+0.86	Nov Dec Jan	$Vo_{t-3} < 0.853$	$Tma_{t-6} > -2.166$	
5	+	Dec Jan Feb	+0.34	Aug Sep Oct	$Vo_{t-3} < 0.853$	$Tma_{t-12} > -0.941$	
6	+	Dec Jan Feb	+0.16	Apr May Jun	$Vo_{t-3} < 0.853$	$Tg_{t-3} < -0.316$	
7	-	Jun Jul Aug	-0.27	Nov Dec Jan	$Tmp_{t-12} > -0.119$		
8	+	Jan-...-Jul	+0.23	Jul Aug Sep	$Tmp_{t-12} > -0.119$	$Tma_{t-2} > 0.735$	

ones. A scheme describing the interactions among terms is offered in Table 3, where the simultaneous conditions to be fulfilled are indicated in order that a given term of the model be “switched on”.

5 Analysis of results

5.1 Some statistical considerations

Model residual distribution is nearly normal; the distribution of results is somewhat narrower than that of observed data, indicating that the model tends to underestimate extreme events. In fact, the variance accounted for is $\sim 64\%$, with a correlation in the sample of 0.80. This result is practically the same in the prediction test (i.e. in the data sample not used to implement the model), a symptom of model stability (see Sect. 5.7). On the basis of a comparison of results for both “main data sample” (MDS) and “independent data sample” (IDS), the presented model is selected as the best within the set of a number of schemes derived by means of MARS methodology. Alternative MARS models often show better results in the calibration phase, but their output deteriorates in the prediction test.

5.2 Comparison with simpler methods

Although it is very likely that atmosphere modelling on time-scales longer than 10 days may require a non-linear statistical approach, it is interesting to compare the MARS methodology with more conventional linear autoregression schemes. For this purpose the following reference model was developed:

$$DVo_t = -1.300DVo_{t-3} + 0.995DVo_{t-4} - 1.800DVo_{t-5} + \\ + 1.210DVo_{t-6} - 1.500DVo_{t-7} + 0.519DVo_{t-8} + \\ - 0.730DVo_{t-9} - 0.180DVo_{t-11}.$$

A test of this scheme within the MDS gives a correlation of 0.73, while residuals still show a significant amount of autocorrelation. The comparison between the MARS non-linear scheme and the linear one indicates a gain in accounted variance of about 9%. Moreover, the MARS model seems to be more effective in catching up fast variations.

5.3 Analysis and interpretation of terms; first section [terms (1)–(3)]

The model mathematical expression is divided into three sections according to the initial domain subdivision to make easy the physical interpretation of terms given by Eqs. (1)–(8). For the first section [terms (1)–(3)], elements are characterised by the function $(Vo_{t-4} + 1.813)_+$. Term (1) is very important; the SC criterion value indicates a serious deterioration of results when it is omitted (Table 2). On the other hand, being the knot

(-1.813) coincident with the minimum value of Vo observed over the period 1946–1989, this term is always “on”, tending to balance the negative constant term and producing, at the same time, an oscillation in the predictand behaviour. It may be shown that this is equivalent to the annual cycle in geostrophic vorticity over the western Mediterranean actually observed in climatological studies. In the terms (2) and (3) this behaviour is coupled with that of other elements. The second term is also important: its “basis points” coincide with the minimum values assumed by Vo and Tma . This term is thus always “on”. Its largest contribution appears in summer months when the function $(Tma_{t-4} + 2.166)_+$ has large values due to a reduction in Atlantic meridional temperature gradients. In February the function $(Vo_{t-4} + 1.813)_+$ has a wide range of variability, while the largest fluctuations in the second term are present in June. This third term represents the interactions of Mediterranean vorticity Vo with the permanent subtropical anticyclone in Atlantic area and with Pacific oscillation index SO: although active, this term contributes in a small number of cases and has a relatively light weight (see also Table 2).

5.4 Second section [terms (4)–(6)]

A common feature of terms included in this part is the function $(0.853 - Vo_{t-3})_+$. When term (4) is removed the model capability seriously deteriorates (see Table 2). In this term the said function is coupled to an element (always “on”) containing Tma_{t-6} . It has the role of giving a contribution to the increase in cyclonic vorticity: the contribution is small in summertime when Tma_{t-6} assumes large negative values (in fact, 6 months before, namely in wintertime, the temperature difference pole minus equator is well below average), while it is appreciable in wintertime when Tma_{t-6} assumes large positive values (6 months before, in summer, the temperature difference pole minus equator is definitely above average). The term is active in all months and brings about, together with the first term, the structure of predictand.

Term (5) is characterised by an interesting link with the previous term (4). In this case, the Tma_{t-12} threshold value is effective: it kills the term in months such as January or February, when the probability of finding large pole-equator temperature difference anomalies (12 months before) is high.

The role of term (6), whose contribution depends on the function $(-0.316 - Tg_{t-3})_+$, may be better cleared up by discussing some actual cases.

5.4.1 The role of the Azores high-pressure cell: some cases. We consider first the case of October 1950, which has the following characteristics:

1. relative geostrophic vorticity 500 hPa over the Balearic Islands “3 months before” (July) is definitely anticyclonic (Fig. 2a);

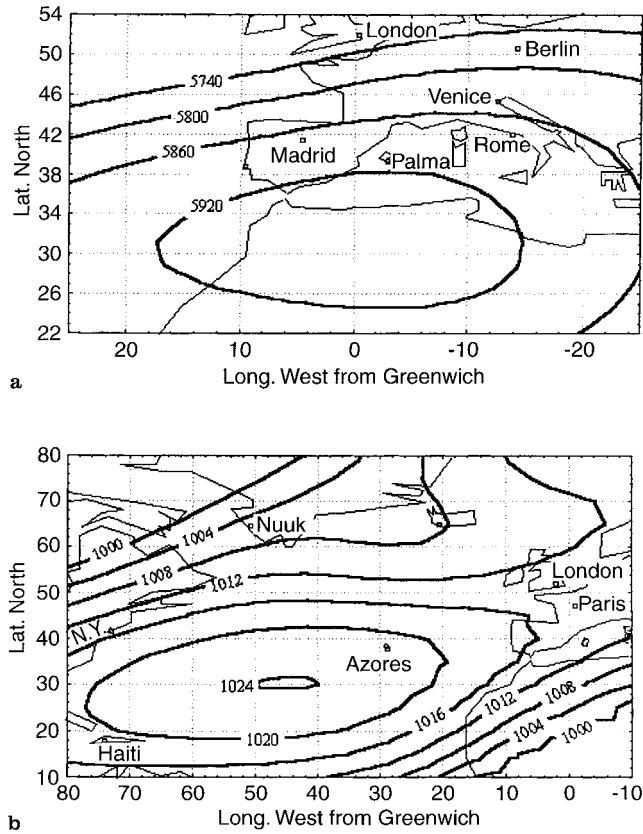


Fig. 2. a Upper: geopotential pattern over the western Mediterranean in July 1950. b Lower: sea-level pressure pattern over the North Atlantic area in July 1950

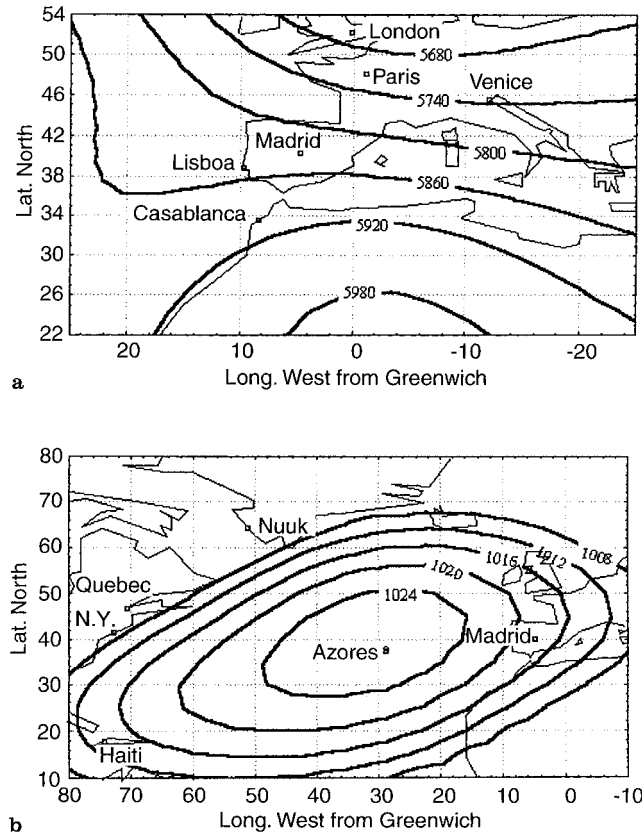


Fig. 3. a Upper: geopotential pattern over the western Mediterranean in July 1966. b Lower: sea-level pressure pattern over the North Atlantic area in July 1966

2. Atlantic high-pressure centre “3 months before” is relatively weak and is located south-west of Azores (Fig. 2b).

In the case of October 1966 we see:

1. relative geostrophic vorticity over the Balearic Islands “3 months before” is slightly anticyclonic or neutral (Fig. 3a);
2. the subtropical high, rather intense, is centered “3 months before” somewhat north of Azores extending a ridge over western Europe (Fig. 3b).

As a consequence of these differences (Table 4) an interesting role of the Azores high-pressure cell stands out. When the model is applied to compute the vorticity change from August and October, its southward migration in July 1950 gives rise to an appreciable contribution of the term (6). Conversely, in October 1966, the sixth term is “off” as a consequence of the north-east position of the Azores high-pressure cell in July (“3 months before”).

5.5 Third section [terms (7),(8)]

This part is characterised by the influence of the temperature pattern over the Pacific region: $(Tm_{t-12} + 0.119)_+$. Term (7) is particularly active in summer. In this season it is very effective in describing the vorticity strong decrease which is a feature of this period. A typical range of contribution is ~ -0.8 to ~ -0.3 . This variability is likely to be linked to the position and extension of the cold polar vortex. Term (8), which includes a “spline” containing Tm_{t-2} , has the role of modulating the effect of the previous term. When the Atlantic temperature-difference anomaly decreases, satisfying the condition $Tm_{t-2} > 0.735$, this function causes an increase in the vorticity contrasting the effect of term (7). These splines are most often active in summer.

5.5.1 The role of the polar vortex. In July 1957, in which the temperature difference pole – equator (12 months before) is above average, the term brings about

Table 4. A scheme describing the values assumed by parameters and the contributions of sixth term and of the single splines

period	model	term 6	$V_{0_{t-3}}$	$(0.853 - V_{0_{t-3}})_+$	Tg_{t-3}	$(-0.316 - Tg_{t-3})_+$
Oct. 1950	-1.354	+0.366	-1.425	2.278	-0.532	0.216
Oct. 1966	1.257	null	-0.699	1.552	0.272	null

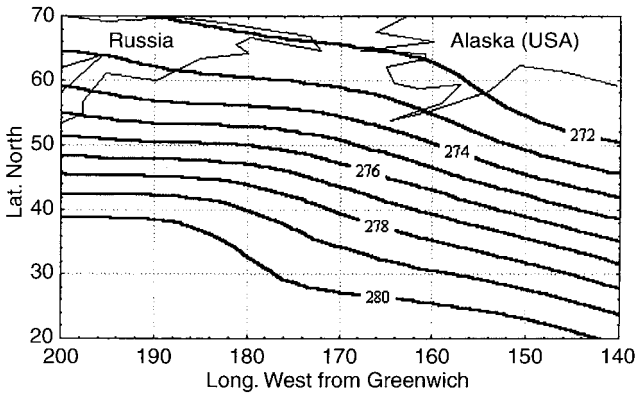


Fig. 4. Temperature pattern (in °K) in July 1956 over the North Pacific area. Characteristic situations, among those analysed, representing maximum contribution in this month for the seventh term

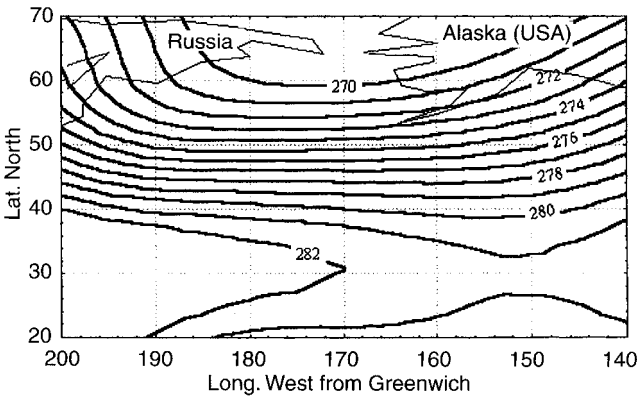


Fig. 5. Temperature pattern (in °K) in July 1971 over the North Pacific area. Characteristic situations, among those analysed, representing minimum contribution in this month for the seventh term

anticyclonic vorticity over the western Mediterranean (Fig. 4). An opposite situation is that of July 1972 (Fig. 5).

5.6 Teleconnections in the space-time domain

A different way of interpreting how the model works is that of considering a specific month, such as November (monthly rainfall maxima in the western Mediterranean area tend to concentrate in this month) and discuss the contribution of terms which are more meaningful when a prediction is computed of the vorticity time change between September and November. We therefore confine the discussion to their role (Table 5):

- Term (1): a marked anticyclonic vorticity 4 months before (July) in the western Mediterranean produces a large contribution in this term, increasing cyclonic vorticity.
- Term (2): in this term the role of the horizontal temperature gradient in the Atlantic area appears. This gradient will be large in November, 12 months before. As a consequence the anomaly becomes negative and the term (which has a negative sign) will be small. In a situation of this type the contribution to the decrease of vorticity will be reduced.
- Term (4): vorticity 3 months before (August) in the western Mediterranean shall be definitely anticyclonic. Simultaneously, the horizontal temperature gradient in Atlantic area 6 months before (May) shall be small. The concurrence of these two elements gives rise to an appreciable increase in cyclonic vorticity.
- Term (5): in addition to term (4), the occurrence of a marked horizontal temperature gradient in the Atlantic area 12 months before (November) gives rise to a positive contribution in order to increase cyclonic vorticity.
- Term (7): the horizontal temperature gradient in the Pacific area 12 months before (November of the previous year) will be large. As a consequence, the anomaly tends to be negative and term (7) (which has a negative sign) is “switched off”. Its contribution to the decrease in vorticity disappears.

According to Table 3, the terms (2), (4) and (7) give the maximum contribution in November. The other

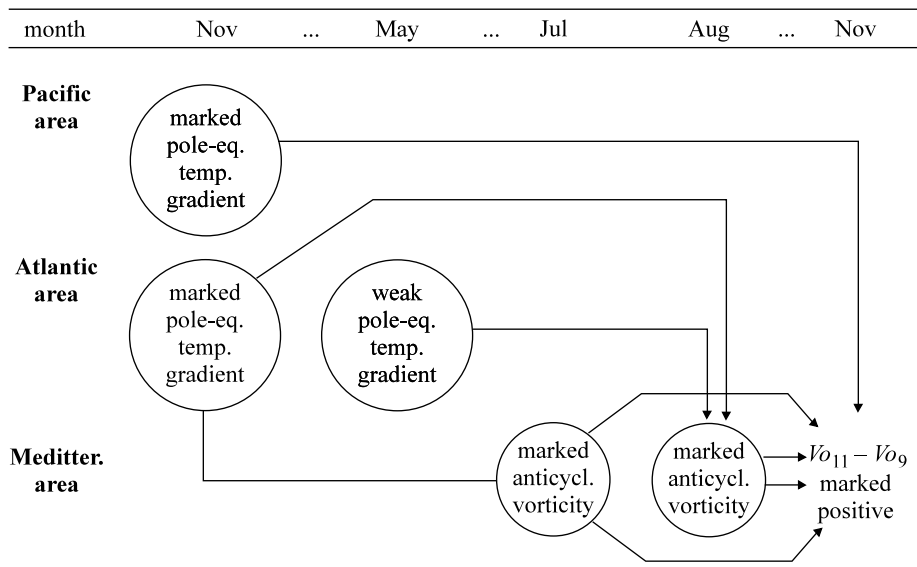


Table 5. Typical sequence of monthly features leading to a marked increase in cyclonic vorticity between September and November

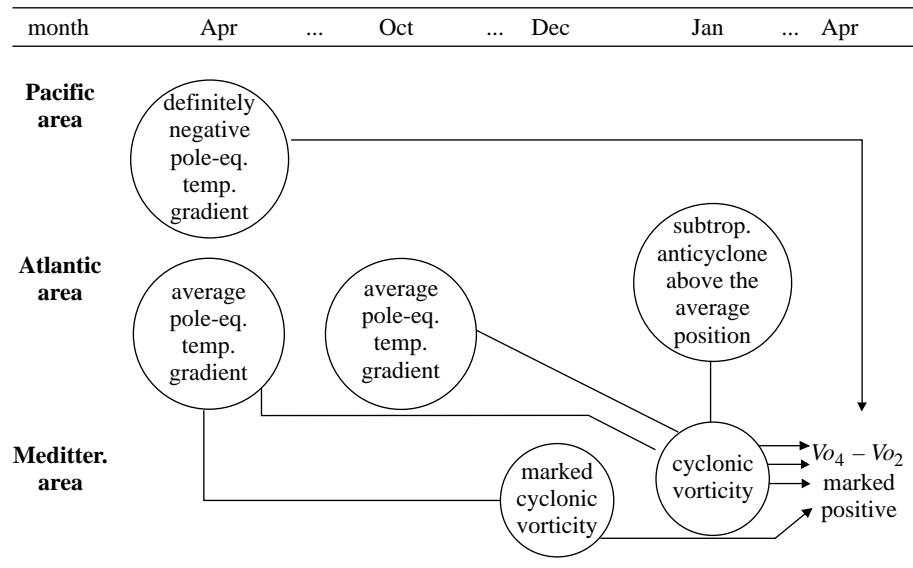


Table 6. Typical sequence of monthly features leading to a marked increase in cyclonic vorticity between February and April

terms [(3) and (6)], appear only in a very small number of cases, and their contributions do not seem appreciable. Applying similar consideration to April (the central month of a season in which a secondary rainfall maximum is present in the Mediterranean area), an identical scheme of the role of the terms can be produced (Table 6). However, some differences can be seen. When the vorticity time change between February and April is computed, a small increase in cyclonic vorticity appears. This behaviour seems to be due to the reduced contribution of terms (4) and (5) (which have a positive sign). In fact, *3 months before* (January) the vorticity in the western Mediterranean will be cyclonic. As a consequence the function $(0.853 - V_{0_{t-3}})_+$ assumes small values, reducing the contribution of the terms. The first term has an important role, having the greatest values in this month. At the end, the role of term (6) can be showed: *3 months before*, the concurrence of the Atlantic subtropical anticyclone somewhat northwards of its normal position and of the cyclonic vorticity in the western Mediterranean gives rise to an increase in vorticity in April.

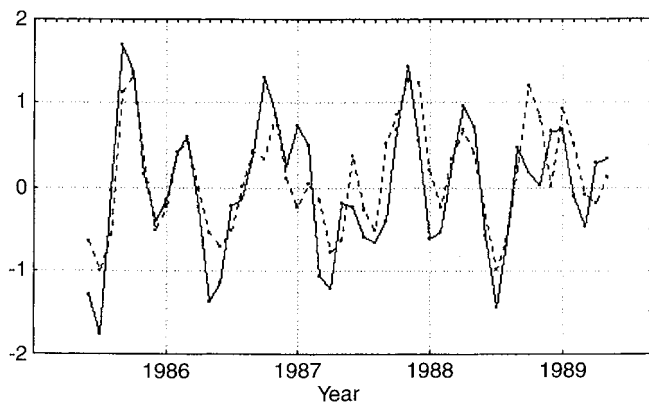


Fig. 6. Predictand pattern from June 1985 to May 1989. *Solid line* indicates the behaviour of observed data, *dashed line* indicates the behaviour of the model developed by MARS

5.7 Analysis of forecast

Model prediction skill is tested on the basis of a 4-year data set (IDS), starting in June 1985 and ending in May 1989 (Fig. 6). The model prediction results, using IDS data as input, are perfectly consistent with the model behaviour within MDS: the correlation between predicted and observed vorticity time changes over the western Mediterranean is 0.78 with a mean square error (MSE) of 0.255, while, within MDS, correlation is 0.80 and MSE is 0.279. Although a more exhaustive investigation is necessary to assess the model performance, this is a promising indication in support of the model stability.

5.8 Model stability

In order to investigate the dependence of model stability on the uncertainty of input data, an ensemble of model runs is carried out: input data have been perturbed by means of a Montecarlo technique, producing 19 time-series within the range of plus-minus the monthly standard deviation σ (reduced by a \sqrt{k} factor, where k is the number of the observed values available for each month). The ensemble of predictions is shown in Fig. 7. The set of correlations and mean square errors between the observed and predicted values is analysed applying a normality test (Shapiro-Wilk's test $W - Shapiro$ and Wilk, 1965). The hypothesis of a normal distribution is acceptable at a 5% significance level. Another element in support of the good model performance is the scatter of results: they are confined within the range of $\pm\sigma$.

6 Conclusions

Starting from a number of observables in the form of time-series of meteorological elements, a model capable of fitting past records and predicting monthly vorticity

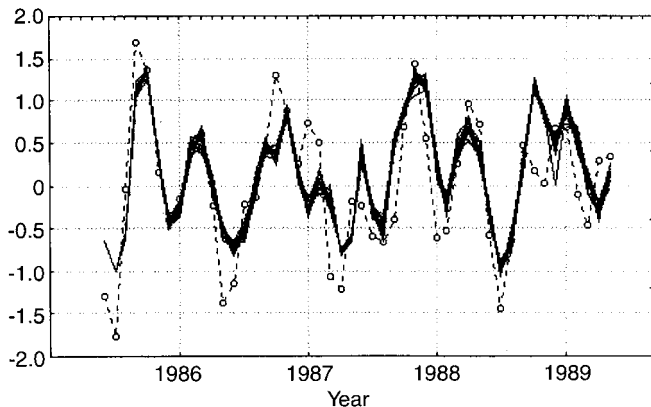


Fig. 7. Graphical representation of ensemble forecast: dashed line indicates the behaviour of observed data, solid lines indicate the behaviour of perturbed simulations

time changes in the western Mediterranean is implemented. The model is derived by means of a new powerful statistical methodology (MARS) which allows to capture the non-linear dynamics of time-series representing the available 40-year history of the northern-hemisphere circulation. After a short review of the MARS methodology the various terms appearing in the model equation were discussed and some possible physical and dynamical interpretations proposed.

When a prediction of the vorticity time changes between September and November is compared, a weak cyclonic or anticyclonic vorticity 3 months before (August) in the western Mediterranean, associated with a small horizontal temperature gradient in the Atlantic area 6 months before (May) is likely to give rise to an appreciable increase in cyclonic vorticity. Moreover, a large horizontal temperature gradient in the Pacific area 12 months before hampers the decrease in vorticity.

Applying similar considerations to April (the central month of a season in which a secondary rainfall maximum is present in the Mediterranean area), the Atlantic subtropical anticyclone somewhat northwards of its normal position and an above-normal value of the surface pressure difference between western coasts of Mexico and Manila (2 months before) are elements hampering the increase in vorticity in the western Mediterranean. On the other hand, a weak temperature gradient pole-equator in the Pacific area, 12 months before, coupled with a below-normal pole-equator temperature gradient in the Atlantic, 2 months earlier, would favour an April increase in vorticity.

The model shows interesting behaviour: the largest contribution of the second term of the model appears in summer months when a reduction in Atlantic meridional temperature gradients occurs. The largest fluctuations in the second term are present in June.

In summer and in early autumn, the fundamental role of meridional temperature gradients over major oceans emerges, as well as that of the position of subtropical anticyclonic cells.

The developed model is tested on a suitable independent data set with very encouraging results. The

comparison between the non-linear scheme and a linear one indicates that when MARS methodology is used a gain in accounted variance of about 9%. Moreover, the MARS model seems to be more effective in catching up fast variations. An ensemble forecast exercise, carried out to check model stability with reference to the uncertainty of input quantities, also supports the stable performance of the model.

In conclusion, the use of the MARS technique appears effective in describing the non-linear dynamics of time-series and, at the same time, in predicting with a 2-month lead the evolution of relative geostrophic vorticity over the western Mediterranean area. Interesting connections between atmospheric phenomena far apart in time and space, which could be hardly detected by other means, are also singled out, providing a base for further studies and a better understanding of atmospheric general circulation.

Acknowledgements. The Editor-in-Chief thanks a referee for his help in evaluating this paper.

We appreciate the constructive criticism and suggestions from both the anonymous referee and the Editor-in-Chief. The study was supported by the "Consiglio Nazionale delle Ricerche-Gruppo Nazionale Difesa Catastrofi Idrogeologiche".

References

- Brankovic, C., T. N. Palmer, and L. Ferranti, Predictability of seasonal atmospheric variations, *J. Clim.*, **7**, 217–237, 1994.
- Corte-Real, J., X. Zhang, and X. Wang, Downscaling GCM information to regional scales: a non-parametric multivariate regression approach, *Clim. Dyn.*, **11**, 413–424, 1995.
- Fraedrich, K., and K. Müller, Climatology of wavenumber-frequency spectra at the 500-mb height along 50°N during the El Niño/Southern Oscillation extremes, *Meteorol. Z.*, **2**, 80–84, 1993.
- Friedman, J. H., Multivariate adaptive regression splines, *Ann. Stat.*, **19**, 1–67, 1991.
- Halpert, M. S., and C. F. Ropelewski, Surface temperature patterns associated with the Southern Oscillation, *J. Clim.*, **5**, 577–593, 1992.
- Lewis, P. A. W., and J. G. Stevens, Non-linear modeling of time series using Multivariate Adaptive Regression Splines (MARS), *J. Am. Stat. Assoc.*, **86**, 864–877, 1991.
- Livezey, R. E., M. Masutani, and M. Ji, SST-forced seasonal simulation and prediction skill for versions of the NCEP/MRF model, *Bull. Am. Meteorol. Soc.*, **77**, 507–517, 1996.
- Palmer, T. N., and D. L. T. Anderson, The prospects for seasonal forecasting – a review paper, *Q. J. R. Meteorol. Soc.*, **120**, 755–793, 1994.
- Palmieri S., et al., Mediterranean climatic extremes (MECLEX), Project report, proposal to European Commission – Directorate-general XII for science, research and development – research and technical development programme in the field of environment and climate, 1995.
- Philander, S. G. H., El Niño Southern Oscillation phenomena, *Nature*, **302**, 296–297, 1983.
- Shapiro, S. S., and M. B. Wilk, An analysis of variance test for normality (complete samples), *Biometrika*, **52**, 591–611, 1965.
- Shukla, J., Short-term climate variability and predictions, in *Climate change: science, impacts and policy – proceedings of the second world climate conference*, Eds. J. Jäger and H. L. Ferguson, Cambridge University Press, Cambridge, pp. 203–210, 1991.
- Taliani, M., S. Palmieri, and A. M. Siani, Visibility: an investigation based on a multivariate adaptive regression splines techniques, *Meteorol. Appl.*, **3**, 353–358, 1996.