Tellus (1997), 49A, 108–118
Printed in UK – all rights reserved

TELLUS
ISSN 0280–6495

Markov processes, dynamic entropies and the statistical prediction of mesoscale weather regimes

C. NICOLIS*[,1], W. EBELING[2] and C. BARALDI[1], [1]*Institut Royal Météorologique de Belgique, Avenue Circulaire 3, B-1180 Bruxelles, Belgium,* [2]*Institute of Physics, Humboldt-University, Invalidenstr. 110, D-10115 Berlin, Germany*

(Manuscript received 12 January 1996; in final form 18 April 1996)

ABSTRACT

The data pertaining to transitions between weather regimes grouped in 3 clusters are analyzed. Evidence that the dynamics of transitions between the regimes is not a first order Markov process is presented on the basis of the properties of residence time distributions. This conclusion is corroborated further by an entropy analysis revealing that the process is characterized by long range temporal correlations. Simple models of this behavior are developed and the repercussions on the problem of prediction are discussed.

1. Introduction

The atmosphere tends to organize in relatively persistent and recurrent circulation patterns. These weather regimes are being identified and classified since the early part of the century according to the space and time scales of interest and to the degree of refinement and detail of the pattern considered (see, for instance, Barry and Perry, 1973). An analysis of this type based on the phenomenology of synoptic processes led to the well-known 21 Grosswetterlagen of the European North Atlantic region. A less refined classification was obtained later by Hess and Brezowsky (1977) by considering a lumping into 3 types of atmospheric circulation: zonal, meridional and mixed. More recent methods are based on the phase space description in terms of a prescribed number of empirical orthogonal functions, where a cluster analysis is carried out to identify distinct flow configurations (Molteni and Tibaldi, 1990; Cheng and Wallace, 1993; Toth, 1993).

One of the motivations behind the pattern recognition and classification of meteorological fields into a finite number of 'symbols' is the possibility to carry out statistical predictions beyond the predictability time of the fine scale atmospheric variables. A classical way of tackling this problem is, given a sufficiently long string of weather types, to estimate the mean residence time of the atmosphere in the vicinity of a given type as well as the probability of transition between clusters. In this way a two-step transition probability matrix is constructed and this information together with the assumption that the underlying coarse grained dynamics defines a Markov process lead to well defined statistical predictions (Fraedrich and Müller, 1982; Spekat et al., 1983). On the other hand, it has been shown recently that the lumping of an initially continuous multivariate system (the atmosphere) into a finite number of heuristically chosen states (the clusters) gives rise, in general, to a non-Markovian dynamics (Nicolis and Nicolis, 1988; Nicolis, 1990). The very possibility to carry out long term predictions is, therefore, highly dependent on the character of the underlying dynamics.

The purpose of the present paper is to reconstruct some aspects of the dynamics underlying a set of data pertaining to the daily succession of weather regimes. In Section 2 the data set is presented and a preliminary statistical analysis is carried out.

* Corresponding author.

Evidence that this set is descriptive of a more intricate dynamics than a first order Markov process is presented and addressed further in Section 3. The particular approach used for this purpose is entropy analysis, which was recently applied successfully in the analysis of strings of symbols generated by a variety of dynamical systems and by certain natural information carrying sequences (Ebeling et al., 1990). Applied to our data set the method provides further evidence of long range temporal correlations, thereby suggesting that the underlying process may be non-Markovian. Accordingly, Section 4 is devoted to the setting up of a qualitative model of this type of behavior based on a discrete time deterministic mapping as well as of a simple non-Markovian probabilistic model accounting in a satisfactory manner for most of the data available. Some final comments on the possibility of carrying out long term predictions for weather clusters along with the main conclusions are compiled in the final Section 5.

2. The data set

The data set consists of daily weather regimes compiled by the Swiss Meteorological Institute from 1 January 1945 to 31 December 1989. The area concerned covers a radius of 222 km centered somewhere in the Alps and including Switzerland as well as part of Austria. The classification scheme consists of 3 main clusters, subdivided in 8 classes and a total number of 40 subclasses (Schüepp, 1979). For statistical reasons we limit the analysis in this paper to the main clusters which are defined according to the pattern of the isobars at a height of 5 km, the wind speed and its direction. They are known as the convective, advective and mixed type clusters respectively.

Fig. 1 displays the time sequence generated by the data (1a) along with the probability histogram of the main weather clusters as obtained from the 16436 daily values (1b). We notice that the time evolution presents a markedly intermittent character. Furthermore, the convective type is the most frequent ($P^c \sim 0.50$) followed closely by the advective one ($P^a \sim 0.44$). The mixed type appears to be approximately an order of magnitude sparser ($P^m \sim 0.064$). Nevertheless, it cannot be neglected altogether nor relegated to a mere transition between the convective and advective regimes.

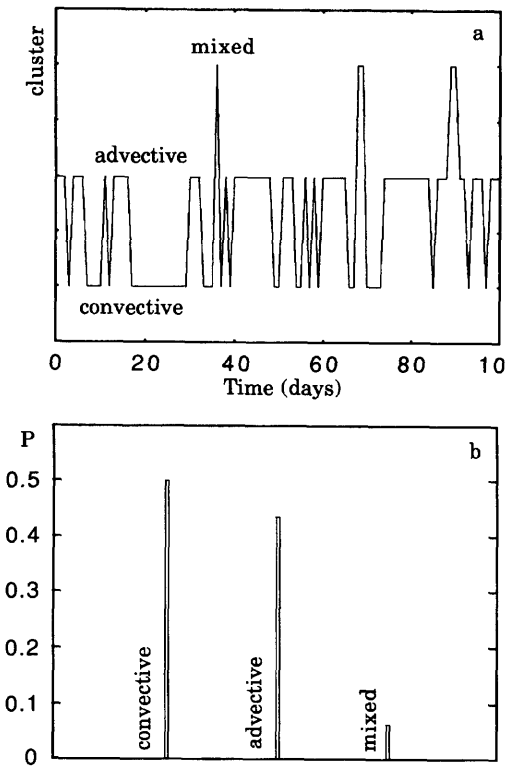


Fig. 1. Typical time sequence of the convective, advective and mixed weather regimes (a); and their respective probability distribution, (b).

Indeed, as seen from Fig. 1a advective-advective, convective-convective and advective-convective transitions may occur directly without involving the mixed regime.

For later use we also construct from the data the conditional probability matrix, giving the transition probabilities to the various states available starting initially from one particular state of the set. A straightforward counting process leads to

$$W = \begin{pmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{pmatrix} = \begin{pmatrix} 0.66 & 0.29 & 0.05 \\ 0.34 & 0.60 & 0.06 \\ 0.35 & 0.48 & 0.47 \end{pmatrix} \quad (1)$$

One may likewise construct from the time series the histogram of the *exit time distributions* from each state, measuring the persistence of this state if the evolution is started in its immediate vicinity. Fig. 2 depicts the results for the three weather clusters of our data set. One is struck by a rather slow decay of these distributions, a feature that is also reflected by the fact that at least as far as Figs. 2a and 2b are concerned, a fitting by an exponential turns out to be a rather poor one giving rise to deviations that are larger than 10%. Now, exponential decay of exit time distributions is one of the characteristic signatures of a first order Markov process (Gillespie, 1992). The results of Fig. 2 cast therefore doubt as to the Markovian character of the succession of weather regimes.

The presence of non-Markovian dynamics and, concomitantly, the existence of long range correlations suggested by the above result is further confirmed by the study of the actual frequencies of finding strings consisting of n symbols (individual clusters) in a prescribed order. Table 1 summarizes the observed frequencies of some of these sequences along with the frequencies that would prevail had the process been a first order Markov. We observe substantial differences ranging from a factor two to two orders of magnitude. Fig. 3 depicts the *rank order* distribution of sequences long of 8-days as obtained from the data. We find more than 800 sequences of the type 11111111 (instead of 440 predicted by the Markov model), the next most frequent sequence being 22222222. Notice that the vast majority of 'words' appear with low probability, whose values are comparable for all these sparse sequences.

In Section 3, further evidence of non-Markovian dynamics is produced based on the properties of entropy-like quantities.

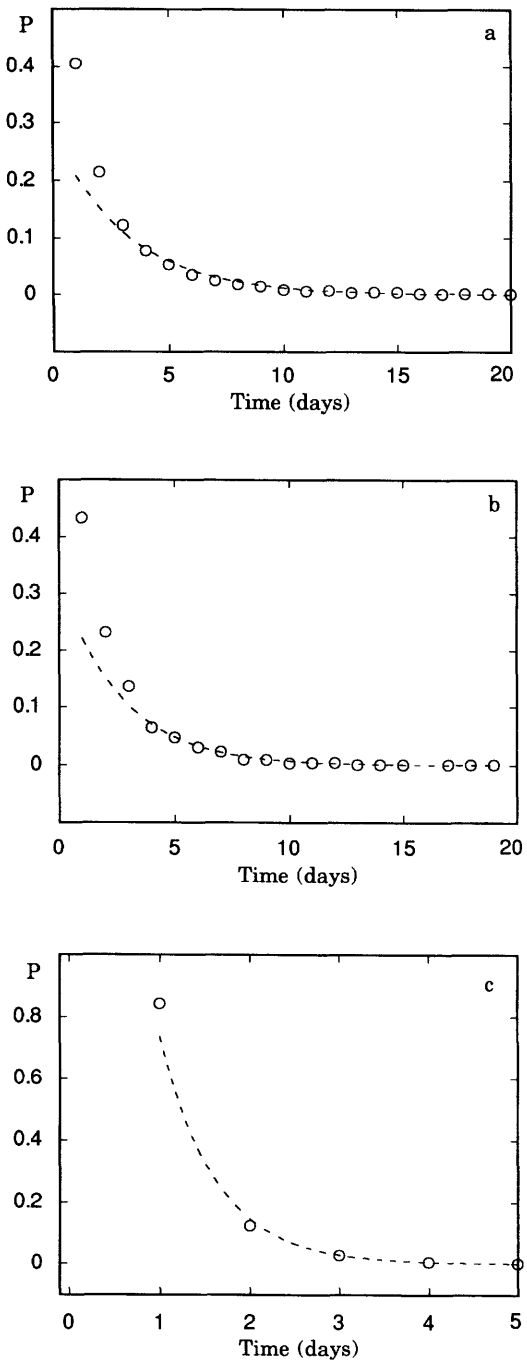


Fig. 2. Exit time probability distributions of the convective, (a); advective, (b) and mixed cluster, (c), as obtained from 16436 daily values. Dashed line represents best fit with an exponential function.

3. Entropy analysis

3.1. Theoretical background

Let $P(X_1 \ldots X_n)$ be the probability to find a succession of daily weather regimes $(X_1 \ldots X_n)$ in a sequence of length n belonging to a data set of length N. The *block entropy* of the n-tuple variable $\boldsymbol{X}=(X_1 \ldots X_n)$ is defined as

$$H_n = -\sum_{X_1 \ldots X_n} P(X_1 \ldots X_n) \log P(X_1 \ldots X_n). \quad (2)$$

Table 1. *Observed relative frequencies of* n*-day clusters in comparison to the first order Markov model*

Sequence	Observed relative frequencies	Relative frequencies inferred from a 1st order Markov model
1.....1 (8 times)	$4.3 \cdot 10^{-2}$	$2.7 \cdot 10^{-2}$
2.....2 (8 times)	$3.5 \cdot 10^{-2}$	$1.2 \cdot 10^{-2}$
1.....1 (16 times)	$6.5 \cdot 10^{-3}$	$9.7 \cdot 10^{-4}$
2.....2 (16 times)	$1.0 \cdot 10^{-2}$	$2.0 \cdot 10^{-4}$
1.....1 (24 times)	$1.5 \cdot 10^{-3}$	$3.5 \cdot 10^{-5}$

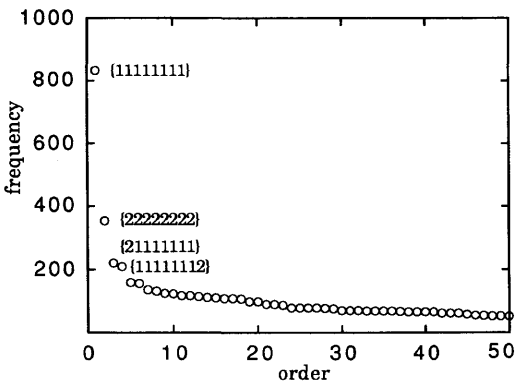


Fig. 3. Rank-ordered distribution of (absolute) frequencies of 8-day successive clusters as obtained from the data of Fig. 1. The most frequent sequences are at least a factor of 2 more frequent than in a first order Markov model.

Since P is normalized to unity relation (2) is positive unless all Ps except one are zero, in which case $H_n = 0$. Notice that H_n represents therefore the mean amount of information required by an observer monitoring the outcome of a particular realization $X_1 \ldots X_n$.

The *entropy excess* per unit time associated with the addition of one outcome in an initial sequence of length n is

$$h_n = H_{n+1} - H_n. \tag{3}$$

Obviously, h_n is the information needed to specify the $(n+1)$th occurrence, provided all n previous ones are known. Alternatively, h_n is the mean uncertainty of predicting the next step following an observed sequence of n states.

We note that the uncertainty may be written as an average of the logarithm of the conditional probability:

$$h_n = \langle \log(1/P(X_{n+1} | X_1 \ldots X_n)) \rangle \tag{4}$$

This formula may easily be generalized to the uncertainty of a state which is s steps into the future

$$h_n^{(s)} = \langle \log(1/P(X_{n+s} | X_1 \ldots X_n)) \rangle. \tag{5}$$

Applying this to our meteorological time series $h_n = h_n^{(1)}$ denotes the uncertainty of predicting the weather of the next day after an observation of n days. $h_n^{(2)}$ is likewise the uncertainty of predicting the weather of after tomorrow and so on.

Let us consider now sequences generated by a deterministic dynamical system (Ebeling and Nicolis, 1992). The simplest example is a periodic dynamics with period p. One can easily show that in this case the $(p+1)$th outcome is completely determined by the foregoing p ones such that

$$H_{p+k} = H_p,$$
$$h_{p+k}^{(s)} = 0 \qquad \text{with} \quad k \geq 1, \qquad s \geq 1. \tag{6}$$

In other words, beyond a sequence of p events the information conveyed becomes trivial and the predictability problem is automatically solved. The above argument can be generalized straightforwardly for multiperiodic discrete sequences.

Consider next the opposite extreme of purely random Bernoulli processes for which all possible sequences have equal probability. The number N_{nm} of sequences of m symbols of length n is $N_{nm} = m^n$, and

$$P(X_1 \ldots X_n) = \frac{1}{N_{nm}} = m^{-n}. \tag{7}$$

Under these conditions, we find

$$H_n = n \log m,$$
$$h_n^{(s)} = \log m, \qquad s \geq 1, \tag{8}$$

entailing that H_n is a linear function of the length

of the sequence considered and consequently that the information needed to specify one more step in the sequence is independent of n. The observation of an arbitrarily long sequence of states will thus not decrease the uncertainty of predicting the next one. Relation (8) suggests to use log m as the unit for H_n or h_n, which is equivalent to choosing m as the basis of the logarithms involved.

The above results carry through to a Markov chain of finite order λ such that:

$$P(X_{\lambda+1}|X_0 \ldots X_\lambda)=P(X_{\lambda+1}|X_1 \ldots X_\lambda), \qquad (9)$$

where $P(X_{\lambda+1}|X_0 \ldots X_\lambda)$ is the conditional probability for finding cluster $X_{\lambda+1}$ after the sequence of clusters $X_0X_1 \ldots X_\lambda$. Specifically, one finds straightforwardly that beyond the correlation time λ

$$H_n=e+nh, \qquad (10)$$

the only difference with the Bernoulli case being that h is now smaller than log m.

It has been shown recently, that deterministic dynamical systems operating in the regime of fully developed chaos exhibit scaling properties of H_n similar to Markov chains, eq. (10). On the other hand, in the presence of weak chaos generated by intermittent systems or by systems operating at phase transition points the scaling of H_n with n turns out to be *sublinear* (Ebeling and Nicolis, 1992).

The above considerations suggest that estimates of the excess entropy per unit time from time series data may provide some hints on the character of the underlying dynamics. In particular the decay of the excess entropy (uncertainty) with the order n provides us with an information about the memory of the dynamics. For instance, if

$$h_n=\text{constant} \qquad \text{for } n\geqslant\lambda, \qquad (11)$$

then we can assert that the process considered has no memory extending beyond λ. If on the other hand the h_n are decaying up to an order λ,

$$h_n\leqslant h_{n-1} \qquad \text{for } n\leqslant\lambda, \qquad (12)$$

we may conclude that the memory is at least of order λ.

Finally, it is important to realize, particularly in the context of applications to meteorology, that a longer memory tends to decrease the uncertainty of predictions of future events.

3.2. *Application to the data set*

The entropy analysis of our data set is limited, at the outset, by the fact that the estimation of the block entropy (eq. (2)) is based on observed *relative frequencies* $f(X_1 \ldots X_n)$ rather than probabilities $P(X_1 \ldots X_n)$,

$$H_n^{\text{obs}}=-\sum_{X_1 \ldots X_n} f(X_1 \ldots X_n)\log f(X_1 \ldots X_n) \qquad (13)$$

Because of the limited number of data available, f is in general different from P and eq. (13) does not coincide with (2). As the length n of the string increases this difference becomes more and more pronounced. Actually one can show (Pöschel et al., 1995) that H_n is an upper bound for H_n^{obs},

$$H_n^{\text{obs}}\leqslant H_n \qquad (14)$$

for any finite n. In a graph in which H_n is plotted versus n this limitation will show up through the fact that, even for a random process, H_n will bend and tend to saturate beyond a moderate value of n rather than keep increasing according to eq. (10). At the level of excess entropy h_n this will result in a decrease with n for large n rather than a tendency toward a limit h as stipulated by eq. (10). This would compromize any clearcut distinction between random process and fully developed chaos on the one side (linear growth of H_n with n) and intermittent processes (sublinear growth of H_n with n) on the other side, on the sole basis of the data.

Recently, methods for remedying this deficiency have been proposed (Ebeling et al., 1995; Pöschel et al., 1995). The starting point is to construct from the data the rank-order distribution $P_n(k)$ of the observed frequencies of all words of a given length n as in Fig. 3. The Shannon entropy $-\Sigma_k P_n(k)\log P_n(k)$ of this distribution is just the block-entropy H_n one is interested in. Next, one guesses an analytic form of $P_n(k)$ involving a number of free parameters. For each parameter choice a sample of the same size as the experimental one is generated from this distribution and the frequencies of "words" are determined. The guessed distribution $P_n(k)$ is then optimized by modifying the parameters until the generated frequency distribution practically coincides with the observed one.

Summarizing, the following steps need to be

undertaken in applying the ideas of Subsection 3.1 to the analysis of our data set:

(i) Count the relative frequencies of symbols, pairs of symbols and of longer "words", up to a certain n_{max};

(ii) rank-order the observed frequencies in the sense of Fig. 3;

(iii) compute H_n^{obs} from the obtained relative frequencies, thereby providing a lower estimate of H_n;

(iv) perform the length corrections necessary to remove the spurious effects due to the finite length of the data set and compute the corrected value of the entropy.

The results of the above algorithm for our data set are shown in Figs 4. We see that the excess entropy h_n is decreasing with n even after the length corrections are performed. This implies a sublinear scaling of H_n with n indicative of a non-Markovian process. Evidence of memory effects in atmospheric dynamics has also been reported in Toth (1992). Notice that for $n \geqslant 8$ the length corrections are essential and yield much higher values for the entropies than H_n^{obs}.

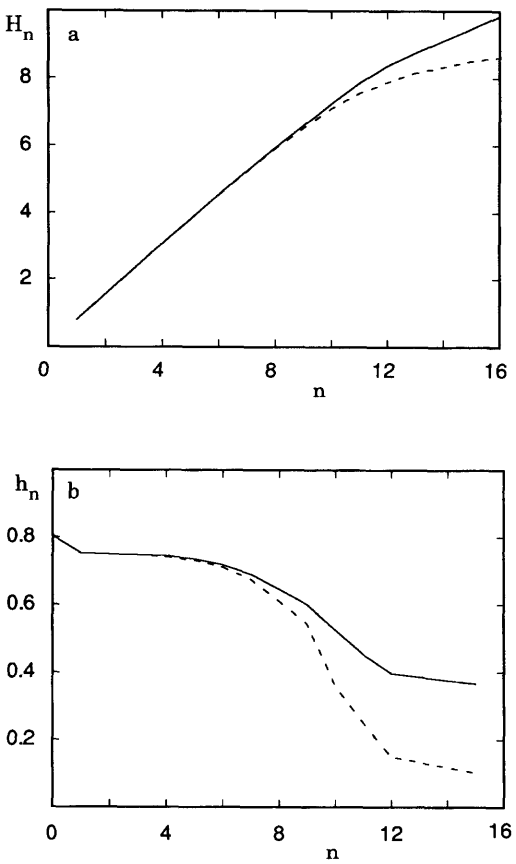


Fig. 4. (a) Block entropies for sequences of length n. The dashed line corresponds to the observed entropy based on the actual frequencies whereas the full line takes into account length corrections. (b) Excess entropies, uncertainties of the next outcome of a sequence consisting of n clusters (dashed line-observed, full line-corrected).

4. Mathematical modeling

Drawing on the evidence of non-Markovian dynamics produced in the analysis of Sections 2 and 3, we develop, in this section, minimal models capable of accounting for the principal statistical properties of the data. We consider successively, a 1-dimensional recurrence giving rise to deterministic chaos and a 3-state stochastic model.

4.1. Deterministic model

Deterministic systems can generate discrete sequences involving a finite number of symbols by a process of coarse-graining. Specifically, the state space of the system is divided into m non-overlapping cells and the successive transitions of the trajectory between these cells are monitored as a function of time. In the presence of chaotic dynamics the transitions define a complex process. For 1-dimensional piecewise linear maps in the region of fully developed chaos the process turns out to be Markovian, provided that the partition is adequately chosen (Nicolis, 1990). Conversely, non Markovian dynamics can in principle be expected in nonlinear maps giving rise to weak chaos. In this subsection we illustrate this possibility on a modified cusp map. This dynamical system, known to produce intermittent behavior owing to the presence of a weakly unstable fixed point is defined by the following recurrence law (MacKernan, 1996):

$$X_{n+1} = \frac{a - \left[a^2 - 4\left(a - \frac{1}{2}\right) X_n\right]^{1/2}}{2\left(a - \frac{1}{2}\right)}, \qquad 0 \leqslant x \leqslant \tfrac{1}{2}, \tag{15a}$$

$$X_{n+1}=\frac{a-\left[a^2-4\left(a-\frac{1}{2}\right)(1-X_n)\right]^{1/2}}{2\left(a-\frac{1}{2}\right)},$$

$$\tfrac{1}{2}\leqslant x<1. \qquad (15b)$$

We notice that the two branches of the map are symmetric around $x=\frac{1}{2}$. For $a=1$, the system posesses a fixed point at $x=0$ on which the slope of the function is equal to unity. This entails that $x=0$ is a marginally stable fixed point, a feature that is responsible for the intermittent character of the dynamics. For $a<1$, $x=0$ remains a fixed point but the slope of the function at this point is now larger than unity (equal to $1/a$). In this region the fixed point is unstable. Nevertheless when parameter a is sufficiently close to 1 the dynamics will still exhibit a remnant of the intermittent behavior.

The invariant probability distribution of (15) is known to be

$$P(X)=2a+4\left(\frac{1}{2}-a\right)X. \qquad (16)$$

Model (15) can be transformed into a discrete version by partitioning X into 3 cells, C_1, C_2 and C_3 (the analogs of the weather clusters). We choose these cells as well as the value of the parameter a in a way to satisfy, as far as possible the data. From the discussion of Section 2 it is clear that we have at our disposal a first series of quantities pertaining to static properties, such as the asymptotic distribution of each weather regime. Identifying cells C_1, C_2 and C_3 with the convective, advective and mixed cluster, respectively and integrating eq. (16) over the partitions one gets

$$P^{c}=2\left[ac_1+\left(\frac{1}{2}-a\right)c_1^2\right],$$

$$P^{a}=2\left[a(c_2-c_1)+\left(\frac{1}{2}-a\right)\left(c_2^2-c_1^2\right)\right],$$

$$P^{m}=2\left[a(1-c_2)+\left(\frac{1}{2}-a\right)\left(1-c_2^2\right)\right], \qquad (17)$$

where c_1 and c_2 are the right boundaries of cells C_1 and C_2. For a given parameter a, eqs. (17) provide a partitioning of the phase space such that the probabilities of the discrete version of model (15) are identical to those of the data set. Specifically, with the choice of $a=0.954$ one finds,

$$C_1=\{0,\ 0.306\},\qquad C_2=\{0.306,\ 0.781\},$$

$$C_3=\{0.781,\ 1\}. \qquad (18)$$

As regards time dependent properties such as transition probabilities the only additional parameter one disposes is the monitoring time step Δn. With $\Delta n=3$, taken to represent the one day interval between successive values, one obtains the following conditional probability matrix,

$$\dot{W}=\begin{Bmatrix}0.63 & 0.31 & 0.05\\ 0.31 & 0.60 & 0.08\\ 0.72 & 0.28 & 0.00\end{Bmatrix}, \qquad (19)$$

again in good qualitative agreement with the data (see eq. (1)) especially for the most frequently occurring clusters C_1 and C_2. Fig. 5 depicts the 3-cell partition obtained with $a=0.954$ and a monitoring time $\Delta n=3$. Figs 6 and 7 depict respectively the exit time distribution for states C_1 and C_2 and the excess entropy associated to the symbolic dynamics between cells. We see very good agreement between data (circles) and theoretical model values (crosses). These results suggest that transitions between weather regimes might be modeled, at least qualitatively, by intermittent chaos. This view is further supported by the results of Fig. 1a pertaining to our data set. In this respect

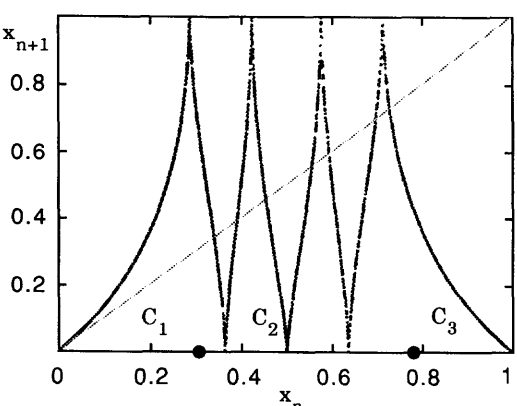
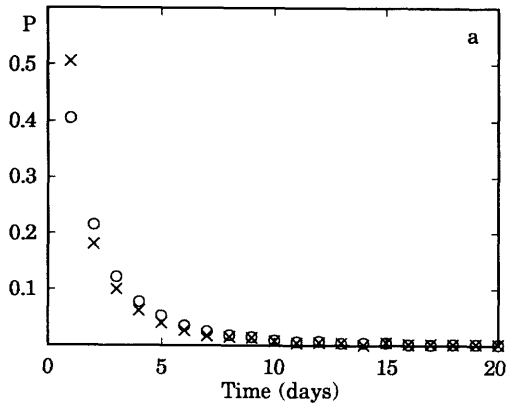


Fig. 5. Recurrence plot of model (15) as obtained from 5000 iterations with $a=0.954$ and $\Delta n=3$. Partitioning the phase space into cells C_1, C_2 and C_3 gives rise to a discrete version of the dynamics whose probability distribution is near to Fig. 1b.



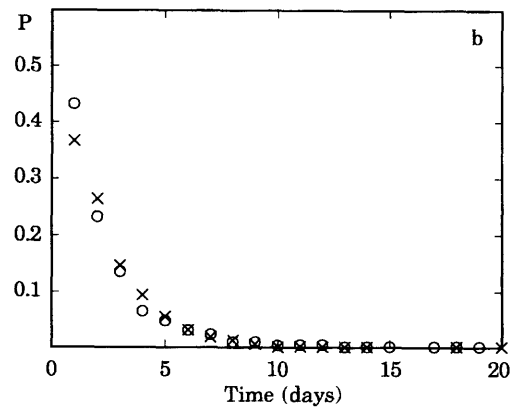


Fig. 6. Exit time probability distribution of states C_1, (a); and C_2, (b); of model (15) (crosses) as obtained from 16436 data points and of the clusters of Fig. 1 (circles). Parameter values as in Fig. 5.

it is also worth noting that intermittency is a ubiquitous feature of complex large scale atmospheric flows.

4.2. *Stochastic model*

In this subsection, we develop an alternative view, whereby the transitions between weather regimes are modeled as a three-state continuous time stochastic process. The evolution of the different realizations of such a process can be viewed as a succession of long halting periods during which the trajectory remains in one particular state, interrupted by rapid jumps between states. Accordingly, for a full specification of the process the following information needs to be

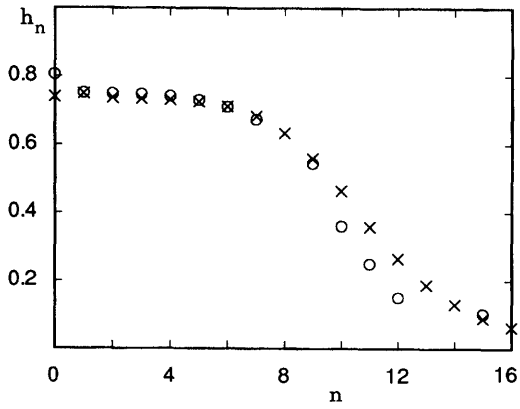


Fig. 7. Excess entropy, eq. (3), as a function of the sequence length, n, associated to the symbolic dynamics between cells, model (15), (crosses) compared to that of Fig. 4b (cicles). In both cases the number of data is 16400.

supplied: (i) the transition probability matrix between states; (ii) the halting time probabilities for each state.

If the process were Markovian, the time evolution of the instantaneous probabilities P_1, P_2, P_3 to be in cells C_1, C_2, C_3 respectively would be generated entirely in terms of the transition probability matrix $\dot{W}$, eq. (1), according to the *master equation* (Gillespie, 1992):

$$\frac{dP_1}{dt} = W_{21}P_2 + W_{31}P_3 - (W_{12}+W_{13})P_1\,,$$

$$\frac{dP_2}{dt} = W_{12}P_1 + W_{32}P_3 - (W_{21}+W_{23})P_2\,,$$

$$\frac{dP_3}{dt} = W_{13}P_1 + W_{23}P_2 - (W_{31}+W_{32})P_3\,, \qquad (20)$$

where we used the property that the row sums of $\dot{W}$ are equal to unity and that $\sum_{i=1}^{3} P_i = 1$.

The exit time distributions corresponding to this picture are obtained by ignoring, successively, in each equation, all terms but the diagonal ones. We thus obtain,

$$p_1(\tau) = \frac{1}{W_{12}+W_{13}}\, e^{-(W_{12}+W_{13})\tau}\,, \qquad (21)$$

and similarly for $p_2(\tau)$ and $p_3(\tau)$. This is in agreement with (and provides the justification of) the assertion made in Section 2 that exponentially distributed exit times are a characteristic signature of a first order Markovian process.

As pointed out in Section 2 the exit time distributions corresponding to the data are poorly fitted by an exponential, entailing that the description based on eq. (21) is actually not valid for the problem in hand. In what follows we shall still use the transition probability matrix induced by the data set, eq. (1), and account for the non-Markovianity through *non-exponential* halting time probabilities (Tsironis and Van den Broeck, 1988). To be specific, we consider two generic forms corresponding to a stretched exponential and to a sum of two exponentials:

$$p(\tau) = K\, e^{-b\tau^{1/2}} \tag{22a}$$

$$= K\, e^{-b\tau^{2/3}} \tag{22b}$$

$$= K\,(e^{-b_1\tau} + e^{-b_2\tau}), \tag{22c}$$

where K is a normalization constant, and fit the parameters b such that an optimal agreement with the experimental distribution is obtained. Fig. 8 summarizes the comparison between the exit time probability of the convective cluster (circles) and the one provided by the model distribution (22a) (crosses). We observe that the agreement is manifestly better than in Fig. 2 especially in the short and long time regimes. Relation (22a) appropriately fitted for the advective and mixed cluster provides likewise a satisfactory agreement. It is worth noting that the remaining model distributions (22b) and (22c) give also rise to statistically acceptable distributions in the sense that deviations from the observations are significantly less than 10% as well.

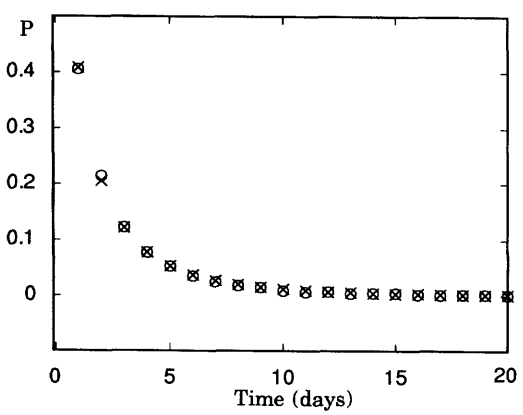


Fig. 8. As in Fig. 2a (circles) and a best fit, eq. (22a), (crosses) with parameters $K \sim 2.14$ and $b \sim 1.65$.

The time evolution of P_1, P_2, P_3 can no longer be cast in a simple form similar to eq. (20). Nevertheless, with the information concerning $\dot{W}$ and $p(\tau)$ at our disposal we are in the position to generate realizations of the underlying process. This simulation can be achieved by a Monte Carlo type simulation as follows (Gillespie, 1992): (i) start from a given initial state i (cluster); (ii) next, draw a random number τ_i subject to the distribution (22) appropriate to that state. The nearest integer $[\tau_i]$ to this number provides the time (days) that the system will stay in this particular cluster; (iii) at $t = [\tau_i]$, state i is left. The particular state that will occur after the transition is chosen according to the values of the transition probabilities $W_{ij}(j \neq i)$. Specifically, a uniformly distributed random number in the interval (0, 1) is drawn. If its value is between 0 and $W_{ij}/\Sigma_j W_{ij}$ a transition to state j is performed; once in state j, the above algorithm is repeated until a good statistics is realized.

Having the realization of the process at our disposal we are in the position to make a forecasting of the evolution starting with a given initial state. Fig. 9 depicts a time series generated by the above method. The simulation reproduces, clearly, the salient features of the data (Fig. 1), especially the intermittent character of the transitions between states. Moreover, by construction, the process generated from the algorithm gives

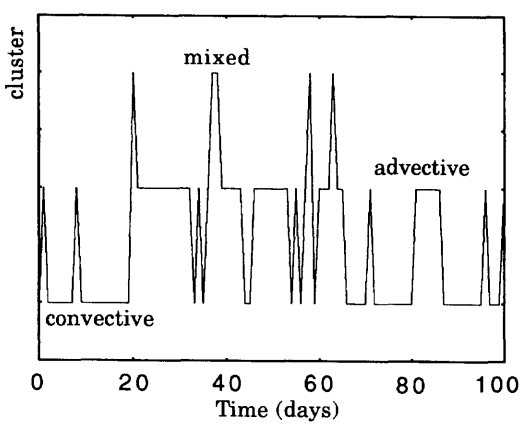


Fig. 9. Typical time sequence of the convective, advective and mixed weather regimes generated from Monte Carlo simulations by adjusting the probability of exit times (Fig. 2) by eq. (22a) as well as the transition probability matrix $\dot{W}$ from the observations.

rise to a transition probability matrix $\dot{W}$ (eq. (2)) as well as a residence time statistics around each regime identical with those provided by the data (Fig. 2).

Finally, in Fig. 10 the scaling of h_n with n for the three forms of $p(\tau)$ chosen, eq. (22), is reported (dotted lines). We observe a qualitative agreement with h_n constructed directly from the data (full line), especially between $n=1$ and $n=10$.

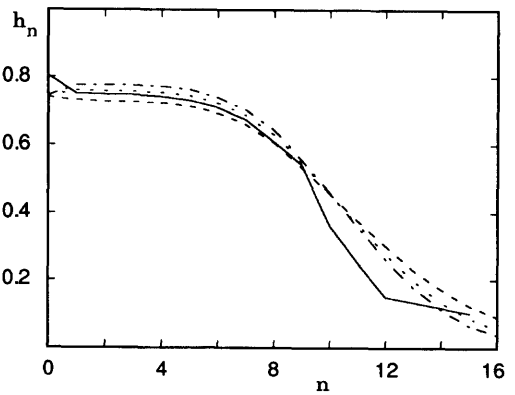


Fig. 10. As in Fig. 7 but for the non-Markovian stochastic models (22a), (dashed line); (22b), dotted line; (22c) (dashed dotted line) compared to that of Fig. 4b (full line).

5. Conclusions

It is by now well-established that small initial errors are amplified by atmospheric dynamics, entailing fundamental limitations in carrying out predictions of the state of the atmosphere beyond a time of the order of the week. The possibility to extend this limited horizon is a problem of obvious concern, whose solution is to be sought in alternative ways to tackle the problem of prediction.

One line of approach that has been intensely explored is to focus on the properties of variables or states which are coarser than the ones obeying to the full scale primitive equations. An interesting method for carrying out such a *coarse-graining*, or *lumping* is to classify atmospheric circulation into a limited number of distinct recurrent weather regimes and resort to a statistical description of the transitions between them. Now this task is faced, at the outset, with the challenge that the dynamics of these transitions must remain a reasonably faithful image of the exact dynamics. Ordinarily, this requirement is bypassed and one resorts to a statistical description in which the transition dynamics is heuristically assimilated to a first order Markov chain (De Swart and Grasman, 1987). On the other hand, a fundamental result of probability theory stipulates that the Markov property generally fails when an initial continuous, multivariate system is lumped into a discrete system. This suggests that the traditional description of the dynamics of weather clusters needs to be reassessed.

In the present work evidence has been produced that the transitions between weather regimes show unmistakable signs of non-Markovian dynamics. This possibility, first suggested by the analysis of 16436 daily observations grouped in three clusters, has been corroborated further by entropy analysis of the time sequence. Two minimal models of this behavior have been subsequently developed: a deterministic model giving rise to intermittent chaos and a stochastic model involving a non-exponential distribution of residence times around each cluster.

Our results suggest that the problem of prediction of mesoscale weather regimes is subtler than usually thought. Surprisingly, it turned out that the prediction of certain patterns can be pushed to longer times than those suggested by mere inspection of the transition probability matrix. We have introduced two new methodologies for implementing such long term predictions: The block entropy or the excess entropy scaling with n; and the direct simulation of the (non-Markovian) process generated by the transition probability matrix and the residence time distributions.

In future investigations it is intended to carry out a similar analysis using a finer division of data in clusters as well as the output of numerical forecasting models of atmospheric dynamics.

6. Acknowledgments

This work is supported, in part, by the Belgian Federal Office for Scientific Technical and Cultural Affairs (O.S.T.C.) under the Global Change Program and the Commission of the European Communities under the Environment Program.

REFERENCES

Barry, R. G. and Perry, A. H. 1973. *Synoptic climatology: methods and applications.* Methuen and Co. Ltd., 555 pp.

Cheng, X. and Wallace, J.M. 1993. Cluster analysis of the Northern Hemisphere wintertime 500 h Pa height field: Spatial patterns. J. Atmos. Sci. **50**, 2647–2696.

De Swart, H. E. and Grasman, J. 1987. Effect of stochastic perturbations on a low-order spectral model of atmospheric circulation. *Tellus* **39A**, 10–24.

Ebeling, W. and Nicolis, G. 1992. Word frequency and entropy of symbolic sequences: a dynamical Perspective. Chaos, Solitons and Fractals **2**, 635–650.

Ebeling, W., Engel, A. and Feistel, R. 1990. *Physik der Evolutionprozesse.* Akademie-Verlag, Berlin.

Ebeling, W., Pöschel, T. and Albrecht, K.-F. 1995. Entropy, transinformation and word distribution of information-carrying sequences. *Bifurcation and Chaos* **5**, 51–61.

Fraedrich, K. and Müller, K. 1982. On single station forecasting: sunshine and rainfall Markov chains. *Beitr. Phys. Atmosph.* **56**, 108–134.

Gillespie, D.T. 1992. *Markov processes.* Academic Press, San Diego.

Hess, P. and Brezowsky, H. 1977. Katalog der Grosswetterlagen. *Ber. Dtsch. Wetterdienstes.* **113**, 39 pp.

MacKernan, D. 1996. *Statistical aspects of deterministic chaos.* PhD dissertation, University of Brussels, Belgium.

Molteni, F. and Tibaldi, S. 1990. Regimes in the wintertime circulation over northern extratropics. Part II: Consequences for dynamical predictability. *Q. J. Meteorol. Soc.* **116**, 1263–1288.

Nicolis, C. 1990. Chaotic dynamics, Markov processes and climate predictability. *Tellus* **42A**, 401–412.

Nicolis, G. and Nicolis, C. 1988. Master equation approach to deterministic chaos. *Phys. Rev.* **A38**, 427–433.

Pöschel, T., Ebeling, W. and Rosé, H. 1995. Guessing probability distributions from small samples. *J. Stat. Phys.* **80**, 1443–1451.

Schüepp, M. 1979. *Witterungsklimatologie.* Beiheft zu der Annalen des Schweizerischen Meteorologischen Anstalt, Zürich.

Spekat, A., Heller-Schulze, B. and Lutz, M. 1983. Über Grosswetterlagen und Markov-Ketten. *Met. Tdsch.* **36**, 243–248.

Toth, Z. 1992. Quasi-stationary and transient periods in the northern hemisphere circulation series. *J. Climate* **5**, 1235–1247.

Toth, Z. 1993. Preferred and unpreferred circulation types in the northern hemisphere wintertime phase space. *J. Atmos. Sci.* **50**, 2868–2888.

Tsironis, G.P. and Van den Broeck, C. 1988. First-passage times for nonlinear evolution in the presence of dichotomous noise. *Phys. Rev.* **A38**, 4362–4364.