Systematic Metastable Atmospheric Regime Identification in an AGCM

CHRISTIAN FRANZKE
British Antarctic Survey, Natural Environment Research Council, Cambridge, United Kingdom

ILLIA HORENKO
Institut für Mathematik, Freie Universität Berlin, Berlin, Germany

ANDREW J. MAJDA
Department of Mathematics and Climate, Atmosphere, Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, New York

RUPERT KLEIN
Institut für Mathematik, Freie Universität Berlin, Berlin, Germany

(Manuscript received 2 September 2008, in final form 28 November 2008)

ABSTRACT

In this study the authors apply a recently developed clustering method for the systematic identification of metastable atmospheric regimes in high-dimensional datasets generated by atmospheric models. The novelty of this approach is that it decomposes the phase space in, possibly, overlapping clusters and simultaneously estimates the most likely switching sequence among the clusters. The parameters of the clustering and switching are estimated by a finite element approach. The switching among the clusters can be described by a Markov transition matrix. Possible metastable regime behavior is assessed by inspecting the eigenspectrum of the associated transition probability matrix.

The recently introduced metastable data-analysis method is applied to high-dimensional datasets produced by a barotropic model and a comprehensive atmospheric general circulation model (GCM). Significant and dynamically relevant metastable regimes are successfully identified in both models. The metastable regimes in the barotropic model correspond to blocked and zonal states. Similar regime states were already previously identified in highly reduced phase spaces of just one and two dimensions in the same model.

Next, the clustering method is applied to a comprehensive atmospheric GCM in which seven significant flow regimes are identified. The spatial structures of the regimes correspond to, among others, both phases of the Northern Annular Mode and Pacific blocking. The regimes are maintained predominantly by transient eddy fluxes of low-pass-filtered anomalies. It is demonstrated how the dynamical description of the slow process switching between the regimes can be acquired from the analysis results, and an investigation of the resulting simplified dynamical model with respect to predictability is performed.

A predictability study shows that a simple Markov model is able to predict the regimes up to six days ahead, comparable to the ability of high-resolution state-of-the-art numerical weather prediction models to accurately predict the onset and decay of blockings. The implications of the results for derivation of reduced models for extended-range predictability are discussed.

1. Introduction

The last couple of decades have seen a continuing search for preferred and recurrent atmospheric flow patterns or regimes. Synopticians were the first to notice the existence of recurring atmospheric flow patterns, with the blocking phenomenon the most pronounced example (Baur 1951; Rex 1950). The existence of atmospheric
regimes offers the potential of skillful extended-range predictions. In extended-range predictions one is not necessarily interested in forecasting the precise location of a cyclone on a particular day; rather, one wants to forecast the overall weather situation for relatively long periods of time, which are likely determined by regimes. For example, in a blocking situation the normal east–west propagation of migratory cyclones and anticyclones is obstructed by a quasi-stationary anticyclone so that they are forced on either a more northward or southward path. Since blockings can last for about a week or longer, this will certainly influence regional weather over periods longer then the typical life span of synoptic-scale cyclones. Current numerical weather forecasting models used by operational centers have problems in accurately predicting the onset and decay of blockings (Tibaldi and Molteni 1990; Pelly and Hoskins 2003; Palmer et al. 2008). Although the high-resolution ensemble prediction system of the European Centre for Medium-Range Weather Forecasts (ECMWF) can predict Euro–Atlantic block onsets up to 6 days and block decay up to 10 days in advance, its limits in the central Pacific region for block onset and decay are 4 and 8 days and 4 and 6 days in the west Pacific region (Pelly and Hoskins 2003). Thus, there is a strong need to understand better the dynamical origins of regimes and to predict their onset and decay. Of special interest are low-order models with predictive skill in forecasting the onset and decay of regime states like blocking because such models are computationally much cheaper than numerical weather prediction models.

In a pioneering study Charney and DeVore (1979) tried to explain the origin of flow regimes in dynamical terms. They used a highly truncated barotropic model for flow over topography and found two stable fixed-point solutions. One corresponds to a zonal flow configuration and the other to a blocking situation. Similar results have been found by Wiin-Nielsen (1979), Charney and Straus (1980), and Legras and Ghil (1985). On the other hand, studies by Reinhold and Pierrehumbert (1982), Tung and Rosenthal (1985), and Majda et al. (2006) find that the distinct flow regimes are not close to the fixed points of the planetary waves.

Whereas the above studies try to explain regime behavior in dynamical terms, there is also a huge body of work whose main focus is on the statistical identification of recurring and/or persistent flow patterns in datasets from either models or reanalysis datasets. The methods of choice include cluster analysis (Mo and Ghil 1988; Cheng and Wallace 1993; Michelangeli et al. 1995; Smyth et al. 1999), nonlinear principal component analysis (Monahan et al. 2000), and looking for quasi-stationary patterns (Vautard 1990). Michelangeli et al. (1995) find that recurrent patterns identified by cluster analysis do not necessarily correspond to persistent (quasi-steady) flow regimes.

The search for deviations from Gaussianity and especially multiple extrema in the phase space of the leading empirical orthogonal functions (EOFs) or planetary waves is another way of identifying flow regimes (Hansen and Sutera 1986; Kimoto and Ghil 1993; Corti et al. 1999; Monahan et al. 2000; Christiansen 2005; Berner and Branstator 2007). In this regime view, deviations from Gaussianity are the imprint of preferred and/or recurrent flow patterns due to nonlinear dynamics. Thus, the Gaussian part of the PDF is thought of consisting of only noise and/or linear dynamics that cannot create regime behavior. The statistical significance and existence of multiple extrema in the PDFs of observed data is a topic of debate in the recent literature (Nitsche et al. 1994; Stephenson et al. 2004; Ambaum 2008). Furthermore, recent studies reveal that the PDFs of the leading few EOFs of long integrations with a three-level quasigeostrophic model (Franzke and Majda 2006) and an atmospheric GCM (Berner and Branstator 2007) exhibit only slight, although significant, deviations from Gaussianity. Nonetheless, these PDFs are unimodal to first order. The deviations from Gaussianity take the form of some weak skewness and ridges with enhanced density in the case of joint PDFs (Berner and Branstator 2007; Franzke et al. 2005, 2007; Franzke and Majda 2006).

Recent studies introduced the concept of metastability for the identification of flow regimes (Majda et al. 2006; Franzke et al. 2008; Horenko et al. 2008, 2009b). In these studies a hidden Markov model (HMM) is fitted to a geophysical circulation model or reanalysis data. A hidden Markov model is designed to describe the situation in which part of the information of the system is unknown or hidden and another part is observed. In this context, the “observed” variable can be a representative variable of low-frequency variability whose dynamics may depend crucially on the overall flow configuration (e.g., zonal or blocked flow), which is unknown. The behavior of this latter “hidden” variable is described by a Markov transition matrix. If the eigenvalue spectrum of the Markov transition matrix exhibits a gap, then the corresponding hidden states can be associated with flow regimes. Thus, this method uses simultaneously both geometrical information of clustering in the phase space and temporal information of the hidden state evolution. This is a major advantage over most methods described above. The HMM approach focuses on the persistence rather the recurrence property of regimes and can also identify regimes in nearly Gaussian or unimodal data.
Once atmospheric regimes have been found, it is important to examine the skill in predicting the onset, decay, and transitions of the regime states. A simple approach for such forecast experiments is to fit a Markov chain to the data (e.g., Crommelin 2004; Deloncle et al. 2007; Kondrashov et al. 2008). In the HMM framework the Markov transition matrix of the regime transitions is already part of the regime identification algorithm (Majda et al. 2006; Horenko et al. 2008a; Franzke et al. 2008) and all data points get systematically assigned to a cluster based on the determination of the most likely hidden state sequence. Such approaches show useful forecasting skill.

In this study we apply a recently developed method (Horenko 2008, 2009a) that is especially suited for the investigation of high-dimensional datasets. This method allows for identification of hidden metastable regimes based on the differences in local EOF characteristics for each of the regimes; that is, it allows simultaneous regime identification and dimension reduction for time series of very high dimensionality. A major advantage of the method, compared to the available HMM-based approaches described in the context of meteorological applications (Majda et al. 2006; Franzke et al. 2008; Horenko 2008, 2009a), is its independence of a priori assumptions about the probability models (like the Markov and Gauss assumptions in context of the HMM–Gauss method). In the present paper we demonstrate how such assumptions can be verified a posteriori for realistic applications described in the context of the most likely hidden state transitions between the clusters. In this clustering approach we assume that the dynamical system consists of two variables: one variable \( Y \) that can directly be observed, but this observed variable depends on a “hidden” variable \( X \) (also referred to as cluster) that cannot be directly observed. In the present context the hidden variable is associated with flow regimes that strongly influence the observed variable.

Here we use the clustering approach FEM-K-EOFs recently developed by Horenko (2009b). This method is based on the minimization of the averaged clustering functional of a given time series \( x(t) \). The FEC method then tries to characterize the time series by \( K \) clusters (or hidden states) and the most likely cluster sequence \( \gamma_i(t) \), where the index \( i \) indicates the \( i \)th cluster and \( \gamma_i(t) \) denotes the probability that the state of the system at time \( t \) belongs to cluster \( i \). Thus, for an a priori given number \( K \) of clusters the FEC method tries to minimize the distance of the actual trajectory (in an appropriate metric) to one of the \( K \) clusters at time \( t \). This means that we are looking simultaneously for the cluster locations and the time evolution of the system in the space spanned by the clusters.

This method considers the clustering of possibly nonstationary multidimensional data \( x_t \in \mathbb{R}^d \) as a minimization problem,

\[
L(\Theta, \Gamma) = \sum_{i=1}^{K} \gamma_i(t) g(x(t), \theta_i) \to \min_{\Gamma(t), \Theta}, \tag{1}
\]

subject to constraints

\[
\sum_{i=1}^{K} \gamma_i(t) = 1, \quad \forall t \in [0, T] \tag{2a}
\]

and

\[
\gamma_i(t) \geq 0, \quad \forall t \in [0, T], \quad i = 1, \ldots, K, \tag{2b}
\]

where we want to minimize the object \( L \). The object \( \gamma_i(t) \) can be interpreted as the probability that the system at time \( t \) belongs to the \( i \)th cluster and \( \theta_i \) denotes the parameters and location of the \( i \)th cluster. The corresponding cluster distance functional characterizes the quality of describing given observation \( x_t \) at time \( t \) by a certain model \( i \) with parameters \( \theta_i \) (see Horenko 2009b for more details). As demonstrated in Horenko (2009b), one can incorporate additional information into the optimization, such as some smoothness assumptions of
functions in space $[\Gamma(t)]$, and then apply a finite Galerkin time discretization of this infinite-dimensional Hilbert space. For example, one can impose the weak differentiability of functions $\gamma_i$; that is,

$$|\gamma_i|_{H^1(0,T)} = ||\partial_t \gamma_i||_{L^2(0,T)} = \int_0^T [\partial_t \gamma_i(t)]^2 dt \leq C_i < +\infty, \quad i = 1, \ldots, K. \quad (3)$$

For a given observation time series, the above constraint limits the total number of transitions between the clusters and is connected to the metastability of the hidden process $\Gamma(t)$ (Horenko 2008).

The minimization functional now takes the form

$$L'(\Theta, \Gamma', \epsilon^2) = L(\Theta, \Gamma) + \epsilon^2 \sum_{i=1}^K \int_0^T [\partial_t \gamma_i(t)]^2 dt \to \min_{\Gamma', \Theta}, \quad (4)$$

where $\epsilon^2$ is a Lagrange multiplier; $\Gamma(t) = [\gamma_1(t), \ldots, \gamma_K(t)]$ denotes a vector of cluster weights and $\Theta = (\theta_1, \ldots, \theta_K)$ the corresponding orthogonal projectors $\theta_i = T_i \in \mathbb{R}^{n \times m}$; and the model distance function is given by $g = |x(t) - T_i^T x(t)|^2$, where $T_i$ is an $n \times m$ dimensional orthogonal projection matrix ($m \ll n$), where $n$ is the dimension of the state vector $x$ and $m$ is the dimension of the EOF manifold used for the projection. The optimization problem is now solved by a finite element approach (see Horenko 2009a,b for more information and a detailed description of the algorithm).

### b. Metastable regimes and model reduction

The studies by Majda et al. (2006), Franzke et al. (2008), and Horenko et al. (2008a) introduced the concept of metastable atmospheric flow regimes. Here we want to briefly review this concept. In this study we refer to metastable regimes if the eigenvalue spectrum of the Markov transition matrix, describing the switching between the clusters, has a statistically significant gap. Such a gap indicates that the state space $S$ of $X$ can be decomposed in two or more sets with relatively infrequent transitions between those sets; thus, the Markov chain is said to be metastable [see Majda et al. (2006), Franzke et al. (2008), and Horenko et al. (2008a) for more details] and the reduced Markov chain describes only the effective, slow transitions between metastable sets. The aim of this study is to identify the states where the system stays for long times before it eventually makes a transition to another regime state. Thus, here we look only for a small set of cluster states that have only infrequent transitions among them.

The concept of model reduction is best described by an example. As an example, we take the $4 \times 4$ transition matrix from Franzke et al. (2008):

$$A = \begin{pmatrix} 0.49 & 0.47 & 0.03 & 0.01 \\ 0.47 & 0.49 & 0.01 & 0.03 \\ 0.02 & 0.01 & 0.53 & 0.44 \\ 0.01 & 0.01 & 0.44 & 0.54 \end{pmatrix} \quad (5)$$

which has the eigenvalues $\lambda_1(A) = 1.0$, $\lambda_2(A) = 0.935$, $\lambda_3(A) = 0.096$, and $\lambda_4(A) = 0.018$. The slow dynamics (or low-frequency behavior) of this Markov chain can be approximated by a $2 \times 2$ transition matrix by lumping together each of the two submatrices $[1, 2]$ and $[3, 4]$. The lumping together is done in this example by summing over the elements $A_{11}$, $A_{12}$, $A_{21}$, and $A_{22}$ and dividing by 2 to give the element $A_{11}$ and by summing over $A_{13}$, $A_{14}$, $A_{23}$, and $A_{24}$ and dividing by 2 to give the element $A_{12}$ and so on for the submatrix $[3, 4]$. By doing so we get the reduced matrix

$$\tilde{A} = \begin{pmatrix} 0.96 & 0.04 \\ 0.025 & 0.975 \end{pmatrix}. \quad (6)$$

which has the same leading eigenvalue structure, $\lambda_1(\tilde{A}) = 1.0$ and $\lambda_2(\tilde{A}) = 0.935$, as the original transition matrix $A$. The transition matrix $\tilde{A}$ has, therefore, the same low-frequency behavior as $A$. This reduction is important in the present context because we will not know the exact number of regimes in atmospheric datasets a priori (if they exist at all). Thus, the examination of the eigenvalue spectrum provides an objective means to identify the number of regime states. The lumping together of submatrices can be easily done for rather low-dimensional transition matrices but is almost impossible for high dimensional matrices. For such matrices with large dimension we have to inspect the eigenvalue spectrum so as to derive a lower dimensional transition matrix. In case the eigenvalue spectrum exhibits a gap, we can utilize the number of eigenvalues above the gap—those close to 1. This number then constitutes the number of states for a new estimation of the transition matrix.

### c. Metastability, Markovianity, and embedding of hidden state sequence

In this study we want to quantify metastability by looking for the existence of a significant gap in the eigenvalue spectrum of a Markov transition matrix of the hidden state sequence. We have to be careful when inspecting the hidden state sequence $\Gamma(t)$ identified by the FEC method because it is not necessarily Markovian. This is partly due to the state space reduction of the full dynamics to only a few hidden states (clusters). To examine if the hidden state sequence is Markovian and to
fit a Markov transition matrix we use the recently developed method for estimation of a Markov transition matrix by fitting a generator to the hidden state sequence (Metzner et al. 2007). In case we are unable to fit a generator, this means the hidden state sequence is not Markovian. To make the hidden state sequence Markovian, we choose to embed the hidden state sequence time series (Broomhead and King 1986; Horenko 2008). To embed, we take the one-dimensional hidden state sequence $\Gamma(t)$ and create a two-dimensional time series:

$$
\hat{\Gamma}(t+1) = \begin{bmatrix} \Gamma(t) \\ \Gamma(t+1) \end{bmatrix}, \quad \hat{\Gamma}(t+2) = \begin{bmatrix} \Gamma(t+1) \\ \Gamma(t+2) \end{bmatrix},
$$

(7)

$$
\hat{\Gamma}(t+3) = \begin{bmatrix} \Gamma(t+2) \\ \Gamma(t+3) \end{bmatrix}, \ldots
$$

We now repeat the generator estimation with this new time series. If this time series is still not Markovian, we increase the embedding dimension,

$$
\hat{\Gamma}(t+2) = \begin{bmatrix} \Gamma(t) \\ \Gamma(t+1) \\ \Gamma(t+2) \end{bmatrix}, \ldots
$$

(8)

and so on until the embedded time series $\hat{\Gamma}$ is Markovian and we can successfully fit a generator. If the analyzed time-discrete process $\Gamma(t)$ has a finite memory depth $d$, the procedure described above will need $d-1$ iterations to construct a Markov process (see Horenko 2008). Note that by doing so we built in the Markov assumption of the hidden state sequence a posteriori and do not need to make a priori any assumption regarding Markovianity and conditional independence of the observed time series. This is a major advantage over the previous approaches based on HMM (Majda et al. 2006; Horenko et al. 2008a; Franzke et al. 2008).

If the Markov hypothesis is justified, the Markov transition probability matrix can be estimated directly from the time series $\hat{\Gamma}(t)$:

$$
P_{ij} = \frac{N_{ij}}{\sum_{k=1}^N N_{ik}},
$$

(10)

where $N_{ij}$ is the number of observed direct transitions between the states $s_i$ and $s_j$ along the analyzed series and $\sum_{k=1}^N N_{ik} \neq 0$ (see Horenko 2009c). If we define a discrete probability density function for the observed Markovian process $\Gamma(t)$ as $\pi(t) = [\pi_1(t), \ldots, \pi_N(t)]$, where $\pi_k(t) = P[\Gamma(t) = s_k]$, then the temporal evolution of this quantity can be described with the following deterministic equation:

$$
\pi(t_0 + \tau) = \pi(t_0)\mathbf{P}^\tau,
$$

(11)

where $\tau$ is some positive integer (e.g., a number of days). If the matrix $\mathbf{P}$ is estimated from the given embedded data $\Gamma$ according to (10), then defining

$$
\pi_k(t_0 + \tau) = \begin{cases} 1, & s_k = \Gamma(t_0) \\ 0, & \text{otherwise} \end{cases}
$$

(12)

one can calculate the prediction as a Markovian probability distribution at some later time $t_0 + \tau$ from (11). In the following, we will demonstrate the application of this procedure in context of atmospheric regime predictions.

### 3. Topographic stress model

The ideal barotropic quasigeostrophic equations with a large-scale zonal mean flow $U$ on a $2\pi \times 2\pi$ periodic domain (e.g., Carnevale and Frederiksen 1987; Grote et al. 1999; Majda et al. 1999, 2003, 2006, 2008; Franzke et al. 2008) are given by

$$
\begin{align*}
\frac{\partial q}{\partial t} + \mathbf{V} \cdot \nabla q + U \frac{\partial q}{\partial x} + \beta \frac{\partial \psi}{\partial x} &= 0 \\
\frac{dU}{dt} &= \frac{1}{4\pi^2} \int \frac{\partial \psi}{\partial x} dx dy,
\end{align*}
$$

(13)

with $q$ the potential vorticity, $U$ the large-scale zonal mean flow, $\psi$ the streamfunction, and $h$ the topography. In (13) the mean flow changes in time through the topographic stress; this effect is the direct analog for periodic geometry of the change in time of angular momentum due to mountain torque in spherical geometry.

The metastable regime characteristics already have been described extensively in Majda et al. (2006) and Franzke et al. (2008) based on the analysis of a reduced phase space of just one and two dimensions. Here we utilize this model to briefly compare the FEC method in the full phase space with the previously used HMM method in a highly reduced phase space. For this
purpose the whole dataset consisting of 57 Fourier modes and the large-scale mean flow $U$ is used. The length of the time series is 50,000 time units sampled every quarter time unit.

An examination of the hidden state sequence reveals that this sequence becomes Markovian for an embedding dimension of 5 (which corresponds to 1.25 time units). The eigenvalue analysis for the corresponding transition matrix reveals three metastable states (Fig. 1). There is a second gap after the fifth eigenvalue, although it is smaller than the gap after the third eigenvalue. Together with the eigenvalues we also display the 95% significance levels based on the sampling of probability matrix distributions (Noe 2008). This shows that the leading eigenvalues are very accurately estimated and that the gap is significant.

The corresponding flow patterns correspond to two zonal flow patterns with different magnitudes and one retrograde (blocking-like) pattern (not shown), which are very similar to the patterns in Franzke et al. (2008). Thus, this result is consistent with the HMM analyses of Majda et al. (2006) and Franzke et al. (2008) in which they only analyzed one- or two-dimensional time series consisting of the large-scale mean flow $U$ and the first Rossby wave mode. This confirms that in this model the regime dynamic is encoded in the large-scale mean flow $U$ and the first Rossby wave mode; furthermore, it also shows the ability of the FEC approach in identifying dynamically relevant metastable regimes in high-dimensional datasets of geophysical flow models.

4. Regimes in an atmospheric GCM

Now we apply the FEC method to the comprehensive National Center of Atmospheric Research (NCAR) GCM CCM0. This model is based on the primitive equations with a full physical parameterization package for processes like radiation and convection. It has nine vertical levels and its horizontal truncation is rhomboidal 15. The model has been used recently by Berner and Branstator (2007) and Branstator and Berner (2005). Franzke et al. (2008) have used the same model for a HMM-based regime analysis. Although this model cannot be considered state of the art, it still has a very realistic representation of low-frequency variability and planetary wave dynamics, which are the main focus of our study.

The dataset that we are using stems from a realization of a perpetual January integration of CCM0 with fixed solar insolation for January and constant sea surface temperature (SST) and has a length of 50,000 days with samples stored every 12 h. To arrive at a reduced system the AGCM states are represented in terms of 500-hPa geopotential height EOFs (in a standard squared metric).

The spatial pattern of the first EOF (Fig. 2a) bears similarity to the Northern Hemisphere annular mode (NAM; Thompson and Wallace 1998) and the second EOF pattern (Fig. 2b) is somewhat similar to the Pacific–North American (PNA) pattern, even though its Pacific center is phase shifted to the west. The other leading patterns are not similar to any well-known teleconnection patterns; they mainly represent wave trains around the Arctic (Figs. 2c and 2d). The explained variances of the EOFs do not show any gap; the explained variance decays rather monotonically, as is typical for such datasets (Fig. 3a). The first 100 EOFs explain about 97.5% of the total variance (Fig. 3b). Thus, we will first focus on this subset for the metastability analysis with the FEC method. As we will see below, this 100-dimensional subspace is more than sufficient to capture the essential metastable dynamics of CCM0.

4.a. Eigenvalue spectrum

The FEC algorithm (Horenko 2008, 2009a) is used to identify hidden states in the subset of the first 100 EOFs of the 500-hPa geopotential height. An examination of the hidden state sequence reveals again that the sequence becomes Markovian for an embedding dimension of 5 (which corresponds to 2.5 days). Next, a Markov transition matrix is fitted to this hidden state sequence. The eigenvalue spectrum of the Markov transition matrix is displayed in Fig. 4 and reveals a gap after the seventh eigenvalue. In Fig. 4 are also displayed the 95% significance levels of the eigenvalue estimates based on the sampling of the Markovian transition probability matrices (Noe 2008). This shows that the first seven eigenvalues are estimated very accurately, whereas higher eigenvalues contain substantial uncertainty. Hence, the eigenvalue gap is significant.
Furthermore, all eigenvalues are real. Thus, the corresponding flow patterns of the hidden states correspond to standing planetary Rossby waves and not propagating Rossby waves. In order that the hidden states correspond to propagating Rossby waves, some of the eigenvalues need to have imaginary components. The fact that the hidden states correspond to standing waves indicates that they are persistent and quasi-stationary flow regimes. That the spatial structures of the hidden states in CCM0 resemble well-known patterns found in reanalysis data will be shown below.

Furthermore, we also examined the Markov transition matrix for preferred transitions between the hidden states. For this purpose we count only the transitions to a different hidden state. By doing so, the following transition matrix is obtained:

\[
A = \begin{pmatrix}
0 & 0.11 & 0.52 & 0.13 & 0 & 0 & 0.23 \\
0.04 & 0 & 0.23 & 0.12 & 0.16 & 0.33 & 0.12 \\
0.23 & 0.19 & 0 & 0.22 & 0 & 0.17 & 0.18 \\
0.13 & 0.19 & 0.09 & 0 & 0.26 & 0.07 & 0.24 \\
0 & 0.20 & 0 & 0.24 & 0 & 0.33 & 0.21 \\
0 & 0.18 & 0.18 & 0.13 & 0.35 & 0 & 0.13 \\
0.06 & 0.11 & 0.24 & 0.23 & 0.13 & 0.21 & 0
\end{pmatrix}
\]  

(14)
All transition probabilities in the above matrix are considered to belong to a preferred transition if their probability is significant larger than when all transitions would be equally likely. To find the preferred transitions we also estimate significance intervals for the transition probabilities according to Horenko et al. (2008a). All matrix elements in bold are significant at the 99.9% level and thus describe preferred transitions. As can be seen, each hidden state has at least one preferred transition to another state, while states 2, 3, 4, 5, and 7 have two preferred transitions to other states. These preferred transitions offer the possibility of skillful predictions, which will be investigated below.

b. Hidden state composites

The following discussion is based on composite fields based on the hidden state path. For example, to compute the composite of the conditional mean field of geopotential height of hidden state 1, we average the geopotential height fields over all times when the hidden state sequence belongs to hidden state 1, and so forth. A representative sequence of the hidden state path is displayed in Fig. 5. The conditional mean states corresponding to the hidden state sequence based on seven hidden states are displayed in Fig. 6. Hidden states 1 and 3 resemble the negative phase of the NAM and hidden states 5 and 6 the positive NAM phase. Hidden state 4 bears some resemblance to the PNA, whereas hidden states 2 and 7 are wave trains extending over the Pacific and Atlantic Oceans; both hidden states also bear some similarity to EOF3 and EOF4, respectively. The fact that hidden states 1 and 3 and also 5 and 6 have striking similarities does not indicate that those flow fields are just different phases of a propagating planetary wave. For this to be the case the eigenvalue spectrum would need to contain complex eigenvalues.
The hidden states 3 and 5 are also similar to the anomalous states identified by Branstator and Berner (2005) and Berner and Branstator (2007).

The projection of the hidden states onto the EOFs (Fig. 7a) shows that they mainly project on the leading five EOFs. Furthermore, the pattern correlation between the hidden state composites and the EOFs reveals the dominance of the first five EOFs for the hidden states (Fig. 7b). This suggests that a five-dimensional subspace is sufficient to capture the essential metastable behavior. The persistence characteristics show that hidden state 1 is much more persistent than the other six hidden states (Fig. 8). The second most persistent hidden state is hidden state 5, which projects strongly onto the positive NAM.

Another way to look at the geographical structure of the hidden states is to investigate the left eigenvectors of the Markov transition matrix (Fig. 9). We derive the

![Hidden State 1](image1)

![Hidden State 2](image2)

![Hidden State 3](image3)

![Hidden State 4](image4)

![Hidden State 5](image5)

![Hidden State 6](image6)

![Hidden State 7](image7)

**Fig. 6.** (a)–(g) Conditional mean states of (left) 500-hPa geopotential height composites (contour interval=100 gpm) and (right) the composites with the climatological mean subtracted (contour interval=10 gpm).
geographical structure of the left eigenvectors as a weighted mean of the hidden state composites according to the left eigenvector. Because the transition matrix is a stochastic matrix, the first eigenvalue is associated with the invariant measure and thus corresponds to the climatological mean state (not shown). The second eigenvector has its main centers of action over the North Pacific and the west coast of Canada. Its positive anomaly over the Pacific corresponds to a blocking-like situation. The remaining eigenvectors (3 through 7) project predominantly on both phases of the NAM and are dominated by their centers of action over the Arctic. Eigenvectors 4 and 7 are also similar to the anomalous states corresponding to the means of two-component Gaussian mixtures in the study by Berner and Branstator (2007). Furthermore, eigenvectors 3 and 7 correspond to the anomalous states found at the dynamical center of a quasi-linear region in the phase space of the first four EOFs by Branstator and Berner (2005).

To explore some of the dynamical processes responsible for the regime behavior, we examine composites of the transient eddy forcing $[-\nabla^2 \varepsilon \cdot (u_\varepsilon)]$, where $u$ denotes the horizontal wind vector and $\varepsilon$ the relative vorticity; these are anomalies with respect to their climatological mean fields conditional on the respective hidden state. These composites are displayed in Fig. 10 for the respective hidden states and give an indication of the nonlinear forcing involved in maintaining the respective hidden state mean fields. The geographical structure of the transient eddy forcing fields suggests that the transient eddy forcing over the North Pacific plays a major role in the regime dynamics. The transient eddy forcing is mainly concentrated over the Pacific Ocean for hidden states 2 through 6 (Figs. 10b-f).
Transient eddy forcing mainly acts over both the Pacific Ocean and Asia for hidden state 1 (Fig. 10a) and over both the Pacific and Atlantic oceans for hidden state 7 (Fig. 10g). A comparison with the transient eddy forcing from low-pass filtered\(^1\) (periods longer than 10 days, superscript \(L\)) anomalies \(\frac{2}{C_1}(u_L z_L)\), high-pass filtered (periods less than 10 days; superscript \(H\)) anomalies \(\frac{2}{C_1}(u_H z_H)\) and the interaction between high- and low-pass filtered anomalies \(\frac{2}{C_1}(u_H z_L + u_L z_H)\) reveals a complex picture. Although there is a preference of both the low- and high-pass filtered transient eddy fluxes to play a role in the more annular-mode-like regimes (hidden states 1, 2, and 5; Fig. 10; note the smaller contour interval of the high-pass filtered transient eddy forcing), this is less the case for hidden state 6. Our analysis also reveals that the low-frequency transient eddy fluxes are dominant for the more wave-train-like hidden states 2 and 7 and also the PNA-like regime state (Fig. 10, middle column). The interactions between high- and low-pass filtered anomalies (not shown) are of only secondary importance. Overall, the low-pass filtered transient eddy fluxes are involved in the maintenance of all regime states, whereas the high-frequency transient eddy fluxes are involved in only the annular-mode-like regimes states. The involvement of the low-pass filtered transient eddy fluxes offers the possibility of successful predictions of the regime states, which will be investigated next. It also has to be mentioned that the correspondence between the transient eddy fluxes (Fig. 10) and the regime states (Fig. 6) is not perfect, indicating that other processes, for example, linear processes such as vorticity advection, also contribute to the regime maintenance.

**Predictability**

The successful identification of metastable regimes offers the prospect of successful long-range prediction of at least the regime states, although not necessarily the precise weather conditions on a particular day. In current weather and climate prediction models the skillful forecasting of the onset and decay of blocking situations is still a major problem (Tibaldi and Molteni 1990; D’Andrea et al. 1998; Pelly and Hoskins 2003; Palmer et al. 2008).

Since some of our metastable regime states are associated with blocking-like situations, this suggests that it is at least possible to generate skillful probabilistic forecasts of the onset and decay of blockings or other regime states by predicting the evolution of the hidden states. To test the predictive skill of the Markov matrix in predicting the hidden state sequence we split our dataset in half and use only the first half to estimate the Markov transition matrix of the hidden state and then utilize the second half of the dataset for forecast experiments. To calculate the Markovian prediction, we use the procedure based on (11) described above for different values of the embedding dimension. In Fig. 11 the percentage of successful predictions of the hidden state for two sets of experiments is displayed: One is based on the original hidden state sequence and the other uses an embedding of the hidden state sequence with embedding dimension 5. For this embedding the hidden state sequence becomes Markovian. Both settings show an almost exponential decay in prediction skill with an \(e\)-folding time-scale of about 6 days, and the short-term predictability is enhanced for up to 6-day

---

\(^1\) We use a digital Lanczos filter with 31 weights for the filtering.
FIG. 10. Conditional transient eddy forcing composites with climatological transient eddy forcing mean subtracted: (left) full transient eddy forcing (contour interval 5 m$^2$/s$^2$), (middle) transient eddy forcing from low-pass-filtered anomalies (periods $>10$ days, contour interval $=5$ m$^2$/s$^2$), and (right) transient eddy forcing from high-pass-filtered (periods $<10$ days, contour interval $=1$ m$^2$/s$^2$).
forecasts for the case with embedding. This offers the prospect of making successful predictions of the onset and decay of blockings and other regime states of up to 6 days in advance. As is well known, current weather forecasting models have problems accurately predicting the onset and decay of blockings. Thus, our empirical–dynamical ansatz of predicting the hidden-state sequence is a promising ansatz in improving prediction skill and might serve as a reference model in operational weather prediction.

5. Conclusions

In this study we applied the recently developed finite element clustering method (Horenko 2008, 2009a) in order to objectively identify metastable regimes in a high-dimensional dataset produced by a comprehensive atmospheric GCM. This method is designed to identify hidden structures in high-dimensional datasets. We first applied this method to barotropic flow over topography, where we confirmed earlier results on regime behavior with a HMM method in a reduced subspace. Next we applied the FEC method to a comprehensive GCM, the NCAR GCM CCM0. In a 100-dimensional phase space of the 500-hPa geopotential height, the FEC method is able to identify seven dynamically significant metastable regimes. Some of the regimes correspond to the positive and negative phase of the Northern Annular Mode, respectively. Others are very similar to the states that are at the centers of nonlinear features in the

![Maps](image_url)
mean tendency fields in the study by Branstator and Berner (2005; see their Fig. 12). The existence of hidden states is also consistent with the results by Berner and Branstator (2007), who investigated the same GCM dataset. In their study they also find evidence for the existence of two regimes when examining the phase space spanned by the first four EOFs (see their Fig. 16). These two regime states are very similar to some of our hidden states.

In a previous study by Franzke et al. (2008), no evidence of metastable regimes was found by investigating the same GCM dataset using the HMM method. In their study only one-dimensional time series of the leading four EOFs were separately investigated by HMM. In the present study we examined a very high-dimensional phase space using a more advanced method that relaxes some major assumptions of HMMs and simultaneously estimates the regime and dimension reduction. This shows that the regime behavior in CCM0 is only visible in a multidimensional phase space and not in projections on individual EOFs, thus suggesting that in this GCM the regime behavior is due to a nonlinear interaction of various EOFs and cannot be found by just examining individual EOFs separately.

One important question is which processes are responsible for maintaining the hidden states. Since this model is run in a perpetual January mode with fixed solar insolation and SST, all possible regime behavior is due to nonlinear atmospheric dynamics alone. Our results suggest that the transient eddy forcing plays a major role in maintaining the regimes (although the correspondence between the transient eddy forcing and the regime states is not perfect). This suggests that other processes also contribute to the maintenance of the regimes. Both low- and high-frequency eddies contribute to the transient eddy forcing with the low-frequency eddies playing the major role, although our results also show that the storm tracks have distinct differences among the hidden states. This suggests that mainly the nonlinear interaction among low-frequency planetary Rossby waves are involved in the regime dynamics of all hidden states, whereas the high-frequency eddies contribute to the maintenance of the more annular-mode-like regimes. It is also noteworthy that most of the transient eddy forcing takes place over the Pacific Ocean. This suggests that in CCM0 the Pacific region plays a major role for the regime dynamics. While this analysis offers some insight into the relevant regime dynamics, the exact dynamical processes leading to the switching between the atmospheric regimes is still an open question and will be further addressed in future studies.

The very long persistence property of the regime state corresponding to the negative NAM phase offers the possibility of skillful extended-range predictions if one is able to successfully predict the onset and decay of the hidden states and the switching among the hidden states. This particular hidden state also corresponds to a blocked circulation in the sense that this hidden state leads to a weakening and/or meridional displacement of the jet stream. Such a distortion of the jet stream is usually associated with a blocked flow. The question whether replacements of the jet stream on a hemispheric scale associated with the NAM leads to simultaneous blockings in the Pacific and Atlantic regions needs further study.

Our predictability study shows that a simple Markov model for the evolution of the hidden states has predictive skill for about six days in successfully predicting the hidden state. This is about the same skill as the ECMWF Ensemble Prediction System in T255 resolution has in predicting the onset and decay of blocking situations (Pelly and Hoskins 2003), but with a much lower computational cost. Thus, such a model could serve as a complementary model in operational ensemble weather prediction. These results have implications for the derivation of low-order models for skillful extended-range prediction. Such low-order models need to capture the essential dynamics responsible for the switching among the regime states. Thus, to predict this switching the low-order models need to accurately capture the leading eigenvalue structure of the Markov transition matrix, which determines the evolution of the low-frequency modes and more importantly also the regime switching. The prediction prospects of such models have been already alluded to in Majda et al. (2006), where it has been shown that a single stochastic differential equation is capable of capturing the metastable low-frequency regime
behavior compared with the fully turbulent 57-mode topographic stress model. Horenko et al. (2008a) report skillful short-term predictions of temperature for Europe with a model that combines hidden Markov models and linear stochastic differential equations. Hence, our results show the potential for practical use of low-dimensional models for extended-range predictions.

Acknowledgments. We thank Dr. Grant Branstator and Andy Mai for providing us with the CCM0 data set.

REFERENCES


Monahan, A. H., J. C. Fyfe, and G. M. Flato, 2000: A regime view of Northern Hemisphere atmospheric variability and


