Project no. GOCE-CT-2003-505539

Project acronym: ENSEMBLES

Project title: ENSEMBLE-based Predictions of Climate Changes and their Impacts

Instrument: Integrated Project

Thematic Priority: Global Change and Ecosystems

M4.4.5 Establish the methodologies for identifying weather/climate regimes and the linkages between the stratosphere and troposphere for use in the remainder of the project

Due date of deliverable: February 2007
Actual submission date: March 2007

Start date of project: 1 September 2004 Duration: 60 Months

Istituto di Scienze dell' Atmosfera e del Clima – Consiglio Nazionale delle Ricerche
ISAC-CNR

Vn 1.0

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)

Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the Consortium (including the Commission Services)	

M4.4.5 Establish the methodologies for identifying weather/climate regimes and the linkages between the stratosphere and troposphere for use in the remainder of the project

M4.4.5 (a) Methodologies for identifying weather/climate regimes

Atmospheric flow regimes are usually defined as large-scale circulation patterns associated with statistical equilibria in phase space, in which the dynamical tendencies of the large-scale flow are balanced by tendencies due to non-linear interactions of high-frequency transients. The existence of states with such properties can be verified in a rigorous way in numerical simulations with simplified numerical models (as in the pioneering study of Reinhold and Pierrehumbert 1982, or in the experiments by Legras and Ghil 1985 and Vautard and Legras 1988). Since then, circulation regimes have been simulated and diagnosed using a variety of models, ranging from quasi-geostrophic (e.g. Marshall and Molteni 1993; D'Andrea and Vautard 2001, Kondrashov et al. 2004), to balance-equation models (Itoh and Kimoto, 1999), to primitive-equation general circulation models (GCMs) (e.g. Hansen and Sutera 1990; Haines and Hannachi 1995; Monahan et al. 2000). Although their identification in the actual observed record remains a non-trivial statistical problem (see the review by Molteni et al. 2006), recent studies have obtained results which are consistent across different methodologies (Alhamed et al. 2002; Robertson and Mechoso 2003; Cassou et al. 2004; Kravtsov et al. 2006).

Whatever the type of forcing and the time scale involved, it is now widely recognized that the variation of circulation regime properties is an important issue for climate predictions, from seasonal forecasts to future climate scenarios. However, either when validating model results against observations, or when comparing regime properties in different periods, one is faced with the problem of estimating the significance of the products of complex statistical tools. Specifically, one needs to quantify the uncertainty in regime properties arising from the internal, chaotic dynamics of the atmosphere, and use such information to assess the predictability of such properties as a response to external variations in forcing terms.

Given the limitations of the observed record of upper-air fields, it is very difficult to address these issues from re-analysis data. However, ensembles of GCM simulations can provide reliable estimates of the effects of internal vs. external variability in regime properties.

In order to identify weather/climate regimes within the ENSEMBLES project framework we propose to use the cluster analysis technique (to date, despite its limitations, this is the most investigated methodology). In particular we propose the (modified) K-means cluster analysis method applied in Straus et al. 2007, which can be summarized in the following four steps:

1. Identification of clusters in the reduced phase space defined by the empirical orthogonal functions (EOFs), with the associated principal components (PCs) providing the new

coordinates. The leading EOFs (to explain about 75% of the space-time variance) are retained from the covariance matrix.

2. For a given number k of clusters, the optimum partition of data into k clusters is found by an algorithm that takes an initial cluster assignment (based on the distance from pseudorandom seed points), and iteratively changes it by assigning each element to the cluster with the closest centroid, until a ‘‘stable’’ classification is achieved. (A cluster centroid is defined by the average of the PC coordinates of all states that lie in that cluster.) This process should be repeated (at least) 100 times (using different seeds), and for each partition the ratio of variance among cluster centroids (weighted by the population) to the average intra-cluster variance is recorded.

3. At that point the goal is to assess the strength of the clustering compared to that expected from an appropriate reference distribution, such as a multidimensional Gaussian distribution. In assessing whether the null hypothesis of multi-normality can be rejected, it is therefore necessary to perform Monte-Carlo simulations using a **large number** of these synthetic data sets, which are described here. A large pool (500 to 1000) of synthetic data sets is created separately to be comparable to each of the “real” datasets (e.g. Reanalysis, XX Century simulations, Scenario simulations). In each case, each synthetic data set has precisely the same length as the observed or simulated data set against which it is compared, and it is generated from a series of *n* Markov processes, whose variance and firs-order auto-correlation are obtained from the first *n* PCs of the appropriated observed or simulated data set.

4. However, since the null hypothesis of a multi-normal distribution may be rejected not only because of multi-modality but also because of skewness (Christiansen 2007), there is the need of a further Montecarlo test which involves the comparison of the observed (or simulated) data sets with a large number (500 to 1000) of skewed time series of pseudo-random data obtained using a non-symmetric Markov model (see appendix (b) in Straus et al., 2007).

Of course the choice of the cluster analysis tool does not preclude the experimentation and the use of other innovative tools under investigation as for example the projection pursuit method (Bo Christiansen personal communication).

M4.4.5 (b) Methodologies for identifying linkages between the stratosphere and troposphere

In the extra-tropical winter zonal mean circulation anomalies often propagate from the stratosphere down into the troposphere on a timescale of 10-60 days. In the troposphere these perturbations typically manifest themselves as changes in the strength of the North Atlantic Oscillation/Arctic Oscillation. The mechanism involves the interactions between planetary waves and the mean flow.

Current knowledge does not allow to propose a single methodology for the analyses of linkages between stratosphere and troposphere. Therefore we suggest a basket of tests that need to be applied in order to better understand the problem and to identify the various aspects of the stratosphere-troposphere linkage.

- Lagged correlations: The timescale and depth of the downward propagation can be determined by lagged correlations. E.g., the zonal mean wind at 10 hPa or near the surface can be correlated with the zonal mean wind at other altitudes and for different positive and negative lags.
- Coupled patterns: The connection between the stratosphere and the troposphere may also be studied with methods that identify coupled patterns such as Canonical Covariance Analysis, Maximum Covariance Analysis, and Principal Oscillation Patterns. These methods may find the pairs of patterns that dominate the coupling and are important for both studies of the mechanism and studies of the potential predictability.
- Ep-flux at the tropopause level: The EP-flux at the tropopause level can to some extent be seen as a measure of the wave-forcing from the troposphere on the stratosphere. Spikes in this value is expected to be followed by perturbations in the stratosphere and subsequently downward propagation.
- Statistical tests: The limited length of the time-series under considerations makes it necessary that the statistical significance of the results is properly assessed. A special concern when dealing with daily values is the serial correlations in the data that reduce the effective number of degrees of freedom. One way to handle this problem is to use Monte-Carlo procedures where the results from original data are compared to results from surrogate data. However, the surrogate data have to be generated so they have same serial correlations as the original data.

References

Alhamed A, Lakshmivarahan S, Stensrud DJ, 2002: Cluster analysis of multimodel ensemble data from SAMEX. *Mon Wea. Rev*. **130**, 226-256

Cassou C, Terray L, Hurrell JW, Deser C, 2004: North Atlantic winter climate regimes: Spatial asymmetry, stationarity with time, and oceanic forcing *J.Clim.*, **17**, 1055-1068.

Christiansen, B., 2007: Atmospheric circulation regimes: can cluster analysis provides the number? *J. Clim*, in press

D'Andrea, F. and Vautard R., 2001: Extratropical low-frequency variability as a low dimensional problem. Part I: a simplified model. *Q. J. R. Meteorol. Soc*, 127, 1357-1374.

Haines, K., and Hannachi A., 1995: Weather regimes in the Pacific from a GCM. *J. Atmos. Sci*., 52, 2444-2462

Hansen, A.R. and Sutera A., 1990: Weather regimes in a general circulation model. *J.Atmos.Sci.*, 47, 380-391.

Itoh H, Kimoto M, 1999: Weather regimes, low-frequency oscillations, and principal patterns of variability: A perspective of extratropical low-frequency variability, . *J. Atmos. Sci*. **56**, 2684-2705

Kondrashov D, Ide K, and Ghil M, 2004: Weather regimes and preferred transition paths in a three-level quasigeostrophic model. *J. Atmos. Sci*. **61**, 568-587

Kravtsov S, Robertson AW, and Ghil M, 2006: Multiple regimes and low-frequency oscillations in the Northern Hemisphere's zonal-mean flow *J. Atmos. Sci*. **63**, 840-860

Legras, B., and M. Ghil, 1985: Persistent anomalies, blocking and variations in atmospheric predictability. *J. Atmos. Sci*., 42, 433–471

Marshall, J., and Molteni F., 1993: Toward a dynamical understanding of planetary-scale flow regimes. *J.Atmos.Sci.*, **50**, 1792-1818.

Molteni, F., F. Kucharski and S. Corti, 2006: On the predictability of flow-regime properties on interannual to interdecadal timescales. In: *Predictability of Weather and Climate*; Eds T.N. Palmer and R. Hagerdon. Cambridge University Press.

Monahan, A.H., J.C. Fyfe and G.M. Flato, 2000: A regime view of Northern Hemisphere variability and change under global warming. *Geophys.Res.Lett*., **27**, 1139-1142.

Reinhold, B., and Pierrehumbert R.T., 1982: Dynamics of weather regimes: quasi-stationary waves and blocking. *Mon.Wea.Rev.*, **110**, 1105-1145.

Robertson AW, Mechoso CR, 2003: Circulation regimes and low-frequency oscillations in the South Pacific sector. *Mon Wea. Rev*. **131**, 1566-1576

Straus, DM, Corti S. and Molteni F., 2007: Circulation regimes: chaotic variability vs. SST-forced predictability. *J. Climate*, in press

Vautard, R. and Legras B., 1988: On the source of midlatitude low-frequency variability. Part II: Nonlinear equilibration of weather regimes. *J.Atmos.Sci.*, **45**, 2845-2867.