ISSAOS 2001

Università degli Studi dell'Aquila

CHAOS IN GEOPHYSICAL FLOWS

International Summer School on Atmospheric and Oceanic Sciences

Edited by G. Boffetta, G. Lacorata, G. Visconti and A. Vulpiani

otto editore

ISSAOS 2001
Università degli Studi dell'Aquila

CHAOS IN GEOPHYSICAL FLOWS

International Summer School on Atmospheric and Oceanic Sciences

Edited by G. Boffetta, G. Lacorata, G. Visconti and A. Vulpiani

Otto Editore - Piazza V. Veneto 14 - 10123 Torino

www.otto.to.it

ISSAOS 2001
Chaos in Geophysical Flows

International Summer School on Atmospheric and Oceanic Sciences
September 10-14, 2001. L'Aquila, Italy

Edited by G. Boffetta, G. Lacorata, G. Visconti and A. Vulpiani

Prima edizione: ottobre 2003
OTTO editore – Torino
ISBN 88-87503-72-9

mail@otto.to.it
http://www.otto.to.it

CONTENTS

PREDICTABILITY: SHORT- AND LONG-TERM MEMORY OF THE ATMOSPHERE

Klaus Fraedrich
Universität Hamburg, Germany

Abstract

The predictability of a system is closely linked with its short- and long-term memory. Short-term memory is described by an exponentially decreasing auto-correlation with finite integral time-scale, which also characterises the sensitive dependence of, for example, weather and atmospheric flows on their initial conditions. Toy forecast experiments are introduced to demonstrate predictability analysis based on the statistics of forecast errors and ensemble forecasts in imperfect and perfect model environments: Persistence forecasts in red noise (systematic and non-systematic error growth; lag-averaged ensembles, their spread and error) and the Lorenz model for identical twin predictions (attractor topology identified by instantaneous uncertainty growth, ensemble dynamics, optimal growth). Ensemble forecasts of the real atmosphere supplement these toy experiments: The past weather analog ensemble statistics (ensemble spread and scaling) yield estimates of the correlation dimension and entropy when meeting the perfect ensemble and perfect model hypothesis. The low dimension obtained from local weather time series is supported by low degrees of freedom estimated from numerical weather prediction ensemble forecasts. - Long-term memory is described by an auto-correlation with a power-law decline and an infinitely large integral time-scale which, in part, characterises the variability of weather and climate extending over many time scales. This long-term memory is observed in almost all local near surface soil and air temperatures (but not in precipitation) both in observed and simulated time series when subjected to detrended fluctuation analysis. Similar long-term variability occurs in the conceptual model of Fickian diffusion which, if excited by random fluxes, relaxes towards the earth's greenhouse by Newtonian cooling. Long-term memory may also be induced through teleconnection patterns (North Atlantic Oscillation or NAO) and their variability. They are analysed in idealised atmospheric and coupled ocean-atmosphere model experiments: On shorter time scales, the atmospheric forcing (few days) by transient eddies drives the non-linear barotropic tendencies and leads to a NAO-life cycle of a couple of weeks. On longer time scales, the atmospheric NAO mode couples with the ocean and leads to its decadal variability. The outlook adresses predictability issues in Global Change.

I. INTRODUCTION: SOME BASICS

The time-evolution of weather or climate is conveniently characterised by its memory. The short-term memory of dynamical systems is associated with a finite integral time-scale due to an exponential decay of the auto-correlation between initial and future states, C(r) $\sim exp(-r/r^*)$ with the time lag r normalized by r^*. Tools of its exploration have been developed in predictability analysis and they are being exploited to improve weather forecasts; the appropriate key words are local Lyapunov exponents, optimal growth and breeding, ensemble forecasting, etc.. In contrast, the long-term memory is characterised by an infinite integral time-scale and linked to a power-law decrease of the auto-correlation, $C(r) \sim (r/r^*)^{-\alpha}$ with $0 < \alpha < 1$. The methods

of its analysis are still being developed and the understanding is far from being complete. Long and short term memory are the two major subjects of nonlinear dynamics with strong implications for weather and climate processes.

Weather is a nonlinear dynamical system whose limited predictability is due to the ever present initial analysis errors or uncertainties and its internal instabilities. Predictability is described by the evolution of the forecast-error and measured by the distance (or correlation) between the forecast and its corresponding verification map. Fig. 1.1 shows a schematic representation of the (squared) error growing with lead time after averaging over a set of individual forecasts of different initial conditions and weather patterns. Two domains of this diagram are of interest, which are related to the short- and long-term memory of the system: (i) The effect of initial uncertainties on the short-term memory is characterised by (instantaneous and effective) growth rates and by crossing-times of error levels, which are introduced as multiples of the initial error magnitude (error doubling, etc.) or as the error of a reference prediction (climate mean, etc.). The latter is commonly referred to as a predictability limit. (ii) Saturation is reached once the initial value is 'forgotten' and the forecasts become independent of the initial condition. Now long-term memory, if it exists, takes over and affects the time evolution of the dynamical system. This memory does not neccessarily provide sufficient predictive skill at ultra-long forecast lead times.

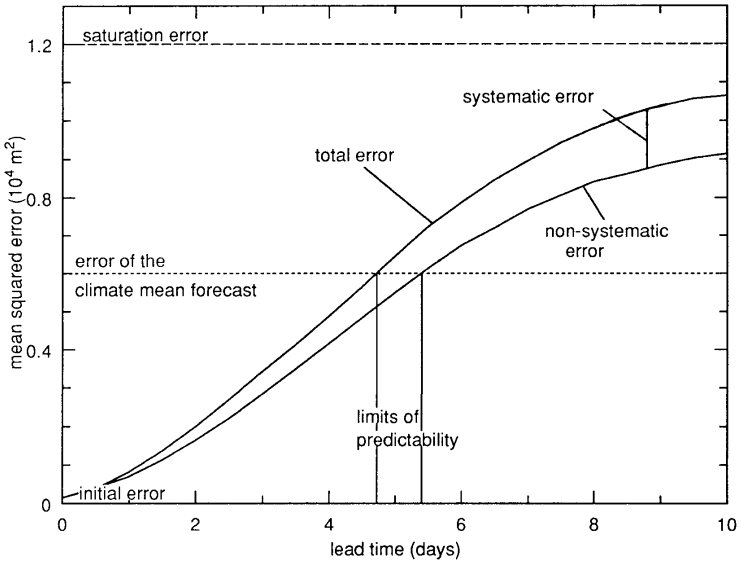


Fig. 1.1 - Error growth curve: Sample mean of individual weather forecasts (schematic).

The following lecture notes present predictability issues related to short- and long-term memory processes. Short-term memory (Section II) is explored using toy forecast experiments and the real atmosphere. Toy systems are introduced to demonstrate the predictability analysis in an imperfect and perfect model environment: Persistence forecasting (weather remains as defined by its initial state) is used here (Section II B) to provide imperfect predictions of a red noise atmosphere, which serves as null hy-

pothesis for significance tests. Experiments with individual forecasts, which demonstrate systematic and random error budgets, are supplemented by ensemble mean predictions generated by a sequence of individual forecasts. Such predictability experiments exhibit many features observed in practical weather forecasting. (ii) Identical-twin predictions with the Lorenz-Saltzman model serve as a paradigm of weather forecasting performed in a perfect model environment (Section II C): Analysing the forecast experiments by instantaneous asymptotic and optimal uncertainty growth unfolds a predictability related attractor topology, which appears to be typical for low-order chaotic systems. Effective optimal growth is introduced to demonstrate another ensemble forecasting technique by employing singular vector decomposition to the linear tangent version of the model.

Forecasts of the real atmosphere supplement these toy experiments: Past weather analog statistics (Section II D) and numerical weather prediction (NWP, Section II E) represent forecasts in a perfect and an imperfect model environment. In a perfect model-ensemble environment analog forecasts provide a prediction related interpretation of the (order-2) dimension and entropy. In the imperfect model environment, degrees of freedom have been estimated from the spread of NWP ensemble forecast maps to identify areas where the predictability needs to be improved.

Long-term memory (Section III) is documented for climate time series applying detrended fluctuation analysis (Section III A), which does not support the Brownian motion analog of climate fluctuations. Instead, a Fickian diffusion and Newtonian relaxation climate system driven by random fluxes is suggested. Atmospheric dynamics also induces long-term memory through teleconnections (North Atlantic Oscillation or NAO, Section III B). The analysis introduces idealised atmospheric and coupled ocean-atmosphere model experiments: On the shorter time scales, the atmospheric forcing (few days) by transient eddies drives the non-linear barotropic tendencies and leads to a NAO-life cycle lasting a couple of weeks. At longer time scales, the atmospheric NAO mode couples with the ocean leading to its decadal variability. Predictability issues in Global Change (Section IV) conclude this essay on predictability.

II. SHORT-TERM MEMORY: FROM PROBABILITY FORECASTS TO ENSEMBLE BREEDING

Predictability characterises the weather or climate system's sensitive dependence on initial and boundary conditions due to internal instabilities. Predictions of the first kind depend on initial conditions and are dominated by internally occurring instabilities at fixed boundary conditions; numerical weather prediction (NWP) is a prominent example. Predictions of the second kind relate the response of the system to changing boundary conditions; such predictions can be associated with structural stability and long-range forecasts. Third kind predictions may be related to the transient evolution of coupled systems, each with a different memory; climate shifts associated with man-made global change is one example.

Predictability analysis describes the error budget in terms of the time evolution of forecast errors measured by the distance between forecast and verification. These analyses can be interpreted in analogy to a diffusion process in state space: Single particle diffusion corresponds to an analysis of the verification trajectory only; its

distance from the origin represents the error growth of the persistence forecast. It is an important reference for forecast quality evaluation and will be discussed later. Two-particle diffusion provides the standard frame for an individual forecast evolving in relation to its verification. A larger set of particles represents ensemble forecasts with dispersing trajectories. Defining appropriate (initial conditions of) ensemble members and relating their stastistics to the verification trajectory is subject of the toy systems analysed.

Predictability experiments provide the data for diagnosing the error budget. External (practical) predictability experiments are linked with the practical task of weather or climate forecasting. The bias due to differences between model and real climate (systematic error) is one of the problems met in analysing predictability of models operating in an imperfect model environment. Internal (theoretical) predictability is related to small perturbations in initial or boundary conditions generated by model atmospheres; this leads to 'identical twin' experiments in a perfect model environment.

Exploration of the atmosphere's short-term memory and utilising its properties has improved the practical forecasting in detail and its performance in general. In particular, single forecasts have been complimented by ensemble predictions to forecast the forecast error and estimate the probability of future weather states. That is, ensemble forecasting plays a prominent role in weather forecasting and climate analysis, which is linked to chaos research and the sensitive dependence of weather or climate on initial conditions.

A. Who started it all?

A century ago the world's first probabilistic weather forecasts were issued in Western Australia and evaluated about eighty years later (in the Monthly Weather Review: Cooke 1906 [10] and Fraedrich and Leslie 1987 [19]). During the year 1905, daily weather forecasts for two districts in Western Australia were amended by quantitative weights of the forecasters' confidence in their predictions; these weights range from 1 to 5 or from not likely at all to almost absolute certainty. Evaluating their skill and reliability (Fig. 2.1), probabilities (0, 10, 60, 90 and 100%) are assigned to Cooke's weights. The Cooke's forecast trial showed that the high confidence weights were most frequently predicted so that the Brier scores (rms error of probability forecasts) attain high values. Since Cooke's time, automatic recording of long records of station data, sophisticated statistical and dynamical techniques and powerful computing devices have encouraged the issuance of weather forecasts. They provide the distribution of possible future weather states either directly as an ensemble forecast or in probabilistic terms.

Tab. 2.1 - Three highlights in forecasting predictability.

E. Cooke (1906 [10])	First probability forecasts (Australia), quantifying confidence in daily forecasts and thus anticipating later developments
E. Lorenz (1963 [51]) K. Popper (1965 [65])	Sensitive dependence on initial conditions (or chaos) 'Of Clouds and Clocks' and 'The Open Universe'
Weather Services (90s)	Routine ensemble-prediction: Forecasting of the forecast skill

Up to now considerable progress has been made in nonlinear dynamics (and the philosophical aspects of complexity, Tab. 2.1), commencing with the pioneering work of Lorenz [51] and Popper [65] in the sixties of the last century. Thus Cooke's first association of weather forecasts with forecast error estimates is common practice a century later, with predictability (or chaos) theory being applied to practical forecasting. The following subsections address basic predictability issues using both toy model experiments (Sections II B and II C) and the real atmosphere (Sections II D and II E); this includes ensemble predictions employing lag-averaged forecasting (LAF), optimal growth and breeding methods to identify and generate the initial ensemble members.

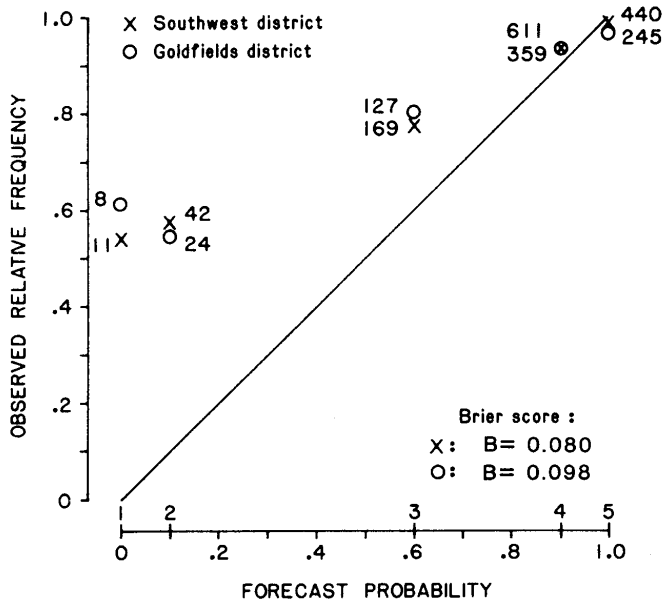


Fig. 2.1 - Reliability diagram (after Fraedrich and Leslie 1987 [19]) of the first probability weather forecasts for two districts near Perth in Western Australia during 1905. The first two of the five confidence weights have been quantified aposteriori in probabilistic terms (abcissa); the number of predictions is also indicated and the Brier scores (lower right).

B. Persistence in red noise

Persistence forecasts ('the weather will remain as it presently is') in a red noise environment are analysed analytically to simulate the imperfect model forecasting experiments with a toy model (Fraedrich and Ziehmann 1994 [22]). The results of individual and ensemble persistence forecasts in red noise are, in some aspects, similar to NWP experiments. Red noise is introduced as a substitute atmosphere using a first order auto-regressive Gaussian process, AR(1), which includes stochastic forcing z_i added

with each time step Δ. The process is discrete in time and continuous in the state variable X, represented by fluctuations, $X(t) =< X > +X'(t)$, about zero mean $< X >= 0$; time (sample) averaging is denoted by the brackets, $<>$, and the prime $X' = X$ describes the anomalies. The stochastic AR(1)-process is the discrete-time analogue of the continuous-time Langevin equation.

Tab. 2.2 - Red noise atmosphere (AR(1)-process).

autoregressive process	$X(t) = aX(t-1) + z_1 = a^r X(t-r) + \sum_i^{r-1} a^i z_{t-i}$
auto covariance	$C(r) =< X(t)X(t+r) >= \sigma^2 a^r$
spectrum integral	$S(\omega) = S_z^2/(1 + a^2 - 2acos(\omega))$
integral time scale	$\tau = \sum_{r=0}^{\infty} C(r)/\sigma^2 = 1/(1-a)$

Time integration of the auto-correlation, $C(r)/\sigma^2 = a^2$, defines the integral time scale, $\tau = \frac{1}{1-a}$ or $\frac{1}{\tau} = 1 - a \sim ln\frac{1}{a}$, as a measure of the life time of a perturbation. The white noise forcing z_i with zero mean $<z_i>$ = 0 is defined by the auto-covariance, $< z_i z_j >= q\delta(t_i - t_j)$ vanishing for $i \neq j$. The related white noise spectrum $S_z = \sigma_z^2 \Delta$ can be defined by the unit time step, Δ, so that $q = \sigma_z^2$. The variance of the response, $\sigma^2 =< (X - < X >)^2 >$, is related to the noise or random forcing intensity $\sigma_z^2 = \sigma^2(1 - a^2)$, which is $1 - a^2$ of the total variability. As the deterministic part of the fluctuations contributes the remaining a^2, the signal to noise ratio is $a^2/(1 - a^2)$. Red noise atmosphere (Tab. 2.2) $(0 < a < 1)$: This first order auto-regressive process is one of the simplest non-trivial processes simulating many observed aspects of the variability in the atmosphere. For $\omega > \frac{1}{\tau}$, the spectrum S drops by a ω^{-2} power-law. A large (small) red noise parameter, a, describes weather regimes with large (small) integral time scales, τ , which are associated with small (large) intensities of the white stochastic forcing spectrum S_z^2. For sufficiently low frequencies $\omega < \frac{1}{\tau}$, the response spectrum flattens to white noise. Note that the random walk $(a = 1)$ commences at $X(t) = 0$ and, after r time steps, reaches $X(t + r) = \sum_i z_i, i = 1, ..r$. It is non-stationary as its variance grows linearly with time $< X^2(t + r) >=< \sum_i z_i^2 >= r\sigma_z^2$, because $< X(t + r) >= r < z_i >= 0$.

Tab. 2.3 - Chance, climate, and persistence forecasts $F(r)$ for lead time r and their mean squared (ms) forecast errors E in the red noise atmosphere.

model	forecast	ms error	comments
chance	$F_R(r) = X_i$	$E_R(r) = 2\sigma^2$	distance of independent states
climate mean	$F_C(r) =< X >$	$E_C(r) = \sigma^2$	reference forecast
persistence	$F_P(r) = X(t-r)$	$E_P(r) = 2\sigma^2(1 - a^r)$	residing in the initial state
combination	$F^*(r) = \alpha F_P + \beta F_C$	$E^*(r) = \sigma^2(1 - a^r)$	AR(1)-process: $\alpha = a$; $\beta = 0$

Chance, climate, and persistence: These forecasts serve as reference predictions $F(r)$ for the lead time r (Tab. 2.3). They are commonly evaluated by the mean squared (ms) forecast error $E =< (F - X)^2 >$ sample averaged $<>$ over the forecast experiments; anomaly correlations are another measure of accuracy frequently used by national meteorological centres. Chance forecasts, F_R, select initial values at random; the ms-error corresponds to the squared distance between independent weather states, which is twice the variance of the system, $2\sigma^2$. Climate is a prediction, F_C, by the climate mean; the ms-error corresponds to the system's variance and defines a threshold for a predictability limit T. Persistence, F_P, is a fundamentally important reference forecast, because only forecasts better than persistence have skill in the forecast of the

time derivative. Persistence predicts the future weather states, $X(t)$, by the initially observed state $X(t_0)$; that is, a persistence forecast, $F_p(r)$, commences at the time $t_0 = t - r$ with the observation $F_P(r) = X(t_0) = X(t - r)$. It is evaluated (after the lead time r), by the verification $X(t)$. Forecast and verification are analysed as a pair of trajectories evolving in state space, whose squared Euclidean distance is the squared error $e^2 = (X - F)^2$ and the average is taken over all verification pairs $E = <e^2>$.

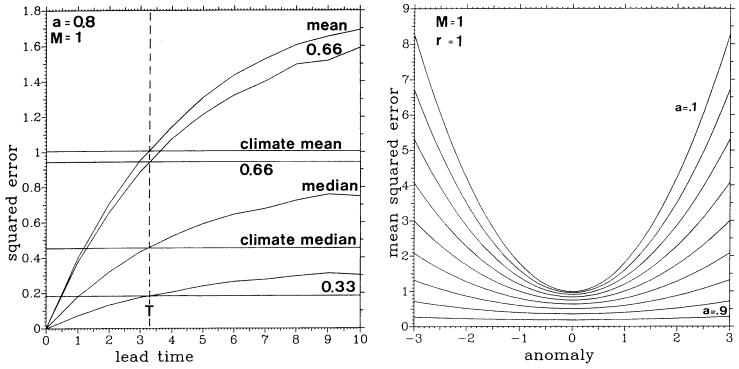
Tab. 2.4 - Imperfect model experiments: Forecasts F(r) and verification X.

error statistics	definitions	persistence in red noise
error	$e = X - F$	$e(t,r) = X(t) - X(t-r)$
mean squared error	$E(r) = <e^2>$	$E(r) = 2\sigma^2(1 - a^r)$
conditional	$E(r\|X_0)$	$<X_0^2> (1 - a^r)^2 + \sigma^2(1 - a^{2r})$
systematic	$SE = <(<F> - <X>^2>$	$\sigma^2(1 - a^r)^2$
non-systematic	$RE = E - SE$	$\sigma^2(1 - a^{2r})$
error growth law	$E_r = f(E, r)$	$E(\infty)[1 - E/E(\infty)]/\tau$
predictability limit	$E(r = T) = \sigma^2$	$T = ln(2)/ln(1/a)$
saturation error	$E(r = \infty)$	$2\sigma^2$

Predictability experiments (random and systematic errors): Mean squared errors, which grow with increasing lead time r, show two domains that deserve particular analysis (Tab. 2.4). The short term memory affects initial errors $E(r = 0)$ and error growth-rates $E_r = \frac{\partial E}{\partial r}$. A limit of predictability is defined as the lead time T at which the prediction error exceeds that of a reference forecast conveniently introduced by the error of the climate mean forecast, σ^2. That is, predictions at lead time $r > T$ needed to exceed the predictability limit $T = ln2/lna$ (setting $E_P(r = T) = \sigma^2$). As T is proportional to the integral time scale τ of weather regimes, $ln\frac{1}{a} \sim \frac{1}{\tau}$, the effective forecast range is limited by the life-span of its most energetic phenomenon. Saturation is reached at large lead times when forecast and verification become independent. That is, their correlation vanishes and the mean squared error $E(r \to \infty) = 2\sigma^2$ defines the mean difference between two randomly chosen weather states (chance forecast).

In an imperfect model environment the mean squared error can be separated in systematic, SE, and non-systematic or random components, RE. Given the initial anomaly X_0, the persistence forecast, $F(r) = X_0$ for lead time r, is associated with the verification time series commencing from this anomaly at the r-th step backward $X(t) = a^r X(t - r) + \sum_{i=0}^{r-1} a^i z_{r-i}$. Averaging the ms-error over a sample of the forecast-verification pairs, $<>$, conditional on a fixed initial anomaly, X_0, yields the conditional forecast error, $E(r|X_0) = X_0^2(1 - a^r)^2 + \sigma^2(1 - a^2 r)$ (see Fig. 2.2a-c). Averaging over all squared conditional anomalies X_0 leads to the unconditional error budget and to the distinction between the forecast error's systematic and non-systematic or random components, $SE = \sigma^2(1 - a^r)^2$ and $RE = \sigma^2(1 - a^{2r})$, which combine to $E = SE + RE$ (Fig. 2.2c). The systematic error is smaller than the random error and both approach unity for infinitely large lead times. At the limit of predictability, r = T, the systematic (non-systematic) error attains $\frac{1}{4}$ $(\frac{3}{4})$ of the climate variance. The initial error growth rate vanishes for systematic errors but is finite for the random part. (v) Persistence forecasts averaged over the same initial anomalies X_0 show the systematic error increasing with the distance of the initial condition from the climate mean, for which forecasts are expected to be better $(SE = 0)$.

a) b)



c)

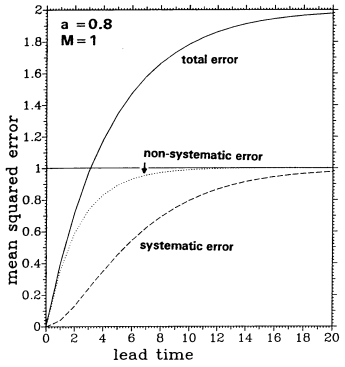


Fig. 2.2 - Error of persistence forecasts of a red noise atmosphere: (a) Time evolution of the squared error distribution: Mean, median, upper and lower terciles (0.66, 0.33) of persistence and climate mean predictions, (b) error at lead time $r = 1$ depending on initial anomaly-conditions, and (c) systematic, non-systematic and total error.

Error growth: This measure enters the laws of error growth, $E_r = f(E, r)$. Most error laws are similar to a Verhulst-type equation, where growth is confined by a quadratic saturation error feedback. Persistence in red noise confines the error growth rate by a linear term: $E_r = 2\sigma^2 a^r ln(1/a) \sim E(\infty)(1 - E/E(\infty))/\tau$ substituting $a = (1 - E/\sigma^2)^{1/r}$. Initial growth, $E_r(r = 0) \sim 1/\tau$, is large in processes with very short memory and the growth decreases with increasing error size E(r). That is, the often misinterpreted generalization that large errors grow slower, holds only for predictability experiments in the same forecast environment (a = const.) where the error growth decreases when approaching saturation. If regime change $(0 < a < 1)$ occurs, their climatology has to be included in the error growth analysis.

Lower and upper bounds of predictability: Persistence plus half-trend forecasts of

the state X(t) have been successfully used in empirical seasonal forecasting: $F_T(r) = X(t-r)+\frac{1}{2}[X(t-r)-X(t-2r)]$. The performance of this scheme shows also interesting similarities with NWP models. Smaller scales (that is, processes with shorter memory or smaller a) tend to have larger initial errors, $E(r=0)=\frac{1}{2}\sigma^2(1+a)$; smaller initial errors E(r=0) are associated with smaller initial error growth rates, $E_r(r=0)=2E(r=0)(1-a)(1+a)^{-1}ln(\frac{1}{a})$. Both results lead to useful predictability estimates, which are similar to results obtained from practical weather forecasts: (i) The link between decreasing initial errors E(r=0) and the increasing limit of predictability T may be extrapolated to yield a lower bound of the predictability limit. It characterises the potential for improving analysis schemes without simultaneous improvements in model performance. (ii) An upper bound may be reached when, in addition, the systematic forecast error can be reduced to zero.

Ensemble forecasts: Ensemble mean forecasts are introduced with the aim (i) to improve the forecast and to (ii) estimate the forecast error, so that an estimate of the future probability distribution of the atmospheric states is possible. The analysis consists of both the ensemble statistics and the error statistics of the ensemble forecasts.

Perfect model/ensemble hypothesis: The perfect model/ensemble hypothesis provides the background for introducing an ensemble of individual forecasts, $F_i(r)$ for $i = 1, ..., M$, and its ensemble average, $[F_i] = M^{-1}(\sum_{i=1}^{M} F_i)$, to predict a field variable X. A perfect ensemble (r = 0) consists of ensemble members F_i, which are chosen such that the mean distances between all members, $[{d_{ij}}^2] = M^{-1}\sum_{i=1}^{M}(\sum_{j=1}^{M-1}(F_i - F_j)^2/(M-1))$, represent the analysis error $[e_i^2] = [{d_{ij}}^2]$. In a perfect model forecast $(r>0)$ the ensemble remains a perfect one: The mean squared distance between all members, $[d_{ij}(r)^2]$, corresponds to the mean squared error of all individual forecasts, $[e_i(r)^2] = [(F_i(r)-X)^2] = [d_{ij}(r)^2]$. Thus the following results can be deduced (see Tab. 2.5, 2.6): (i) The mean-squared error of the ensemble average forecast is about half of the mean of the squared errors of the individual forecasts: $E_M = \frac{1}{2}[e_i^2](M+1)/M$. (ii) There is a linear relation between the ensemble spread S_M and the error of the ensemble mean, $E_M = S_M(M+1)/(M-1)$, because $[{d_{ij}}^2] = 2S_M/(M-1)$.

Tab. 2.5 - Perfect model/ensemble statistics.

ensemble statistics	$[(F_i-X)^2] = ([F_i]-X)^2 + [(F_i-[F_i])^2]$		perfect model/ensemble
ensemble mean	$F_M = [F_i]$		
spread (variance)	$S_M = [(F_i-[F_i])^2]$	set $X=0$	$S_M = [F_i^2] - [F_i]^2$
pairwise distance	$[d_{ij}^2] = [(F_i-[F_j])^2]$	set $X=F_j$	$[d_{ij}^2] = 2S_M M/(M-1)$
error of ensemble mean	$E_M = [(F_M-X)^2]$	set $X \neq 0$	$E_M = [e_i^2] - S_M$

Tab. 2.6 - Lag-average forecasts (LAF) of persistence in red noise.

ensemble mean	$F_M = [F(r+i)] = M^{-1}\sum_{i=0}^{M-1} F(r+i)$
ensemble spread	$S_M = \sigma^2\{1-(1+a)M^{-1}(1-a)^{-1}+2a(1-a^M)M^{-2}(1-a^{-2})\}$
mean squared error	$E_M = \sigma^2\{1+(1+a-2a^r(1-a^M))M^{-1}(1-a)^{-1}-$ $2a(1-a^M)M^{-2}(1-a)^{-2}\}$
systematic error	$SE_M = \sigma^2\{a^r-(1-a^M))M^{-1}(1-a)^{-1}\}$

Imperfect model/ensemble (lagged-average forecasts): Practical forecasts are almost always made in an imperfect model/ensemble environment. Here it is simulated by

an ensemble mean of M lagged persistence forecasts of the red noise atmosphere, $[F_i(r)] = [F(r+i)] = M^{-1} \sum_{i=0}^{M-1} F(r+i)$, where $F(r+1) = X(t-(r+i))$. The error and ensemble statistics can be deduced analytically. The predictability statistics include mean and spread (or distances) of the ensemble members and the conditional (systematic, non-systematic) errors made by ensemble mean forecasts. The results are summarised (Fig. 2.3): The ensemble mean and individual forecast errors, $E_M(r)$ and E(r), show that the lagged-average ensemble of persistence forecasts are, in general, worse than the latest individual forecast (before the predictability limit T is reached). The ensemble spread S_M is independent of the lead time r but changes with ensemble size M. Furthermore, there is no direct error-spread correlation as suggested by the perfect model/ensemble case. Instead, the following is noted. Depending on lead time r and red noise memory τ (or a), there is an optimal ensemble size (M = 8), whose members provide the largest error-spread (E_M, S_M)-correlation (0.31) and whose mean generates the minimum systematic error ($r = 1, \tau = 5$ or $a = 0.8$).

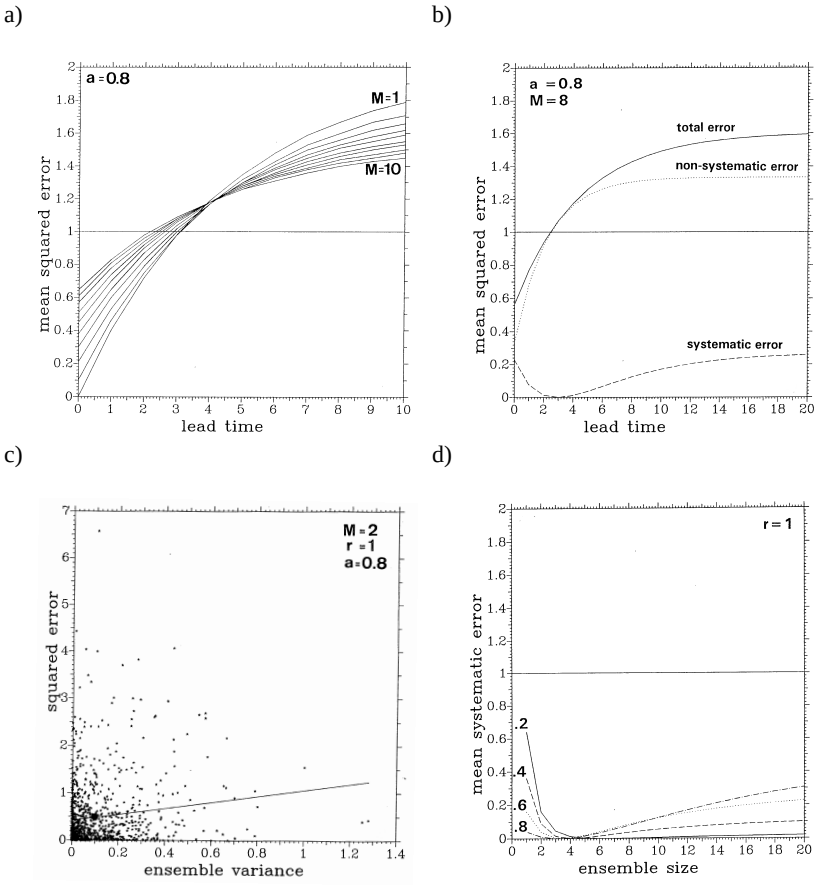


Fig. 2.3 **-** Lag-averaged persistence ensemble mean forecasts in a red noise atmosphere (a = 0.8 autocorrelation): (a) Error changing with lead time and ensemble size, (b) systematic, non-systematic and total error for M = 8 ensemble members, (c) scatter diagram of error versus ensemble spread, (d) systematic error changing with ensemble size M (lead time r = 1).

C. Lorenz attractor revisited

Benard-Rayleigh convection is conveniently described by the incompressible Boussinesq flow whose thermodynamic variables (pressure, density and temperature) are small fluctuations p^*, ρ^*, θ from a reference state (subscript '0') which is hydrostatic, adiabatically stratified and at rest. The density acts only in the buoyancy term and satisfies an equation of state, $\rho^*/\rho_0 = -\alpha\theta$ ($\gg p^*/p_0$), with the thermal expansion $\alpha = 1/T_0$. A y-independent motion with $()_y = 0$ evolves in the (x,z)-plane with the horizontal and vertical velocities (u, w). Friction $\nu\nabla^2(u, w)$ and thermal diffusion $\kappa\nabla^2\theta$ are also included. Other constants are gravity g, kinematic viscosity ν conductivity κ. Tab. 2.7 gives the momentum, thermal energy, continuity, and the equation of state.

Tab. 2.7 - Boussinesq approximation (2-dim).

x-momentum	du/dt	$= -(p^*/\rho_0)_x + \nu\nabla^2 u$
z-momentum	dw/dt	$= -(p^*/\rho_0)_z - g(\rho^*/\rho_0) + \nu\nabla^2 w$
thermal energy	$d\theta/dt$	$= \kappa\nabla^2\theta$
continuity	$u_x + w_z$	$= 0$
state	ρ	$= \rho_0(1 - \alpha\theta)$

From Benard-Rayleigh instability to the Lorenz-Saltzman model [69]: The motion can be described by the horizontal component of the vorticity, $w_x - u_z = \nabla^2\psi = \psi_{xx} + \psi_{zz}$, using the streamfunction with $(u, w) = (-\psi_z, \psi_x)$ and continuity $u_x + w_z = 0$. Combined with the thermal energy one obtains a dynamical system, which describes the convection process by vorticity and thermal energy changes $()_t$ and advection $J(\psi, a) = \psi_x a_z - \psi_z a_x$,

streamfunction $\quad (\nabla^2\psi)_t = -J(\psi, \nabla^2\psi) + \nu\nabla^4\psi + g\alpha\theta_x$
temperature $\quad \theta_t = -J(\psi, \theta) + \psi_x(\Delta T/H) + \kappa\nabla^2\theta$

Linearisation $(')$ is about a state of rest. The Fourier Ansatz, $(\psi, \theta)' = Re\sum_{k,m}(\psi, \theta)_{k,m} exp(ikx) sin(m\pi z)$, satisfies the upper and lower boundary conditions at $z = 0, H$: the temperature is constant $\theta' = 0$ and the vertical velocity and tangential stress vanish, so that $J(\psi', \nabla^2\psi') = J(\psi', \theta') = 0$. With $w' = \psi'_x = 0$ one sets $\psi' = 0$ and $u'_z = -\psi'_{zz} = 0$ gives $\nabla^2\psi' = 0$, so that $\psi' = \psi'_{zz} = \psi'_{zzzz} = 0$ etc.

streamfunction $\quad \psi_t = -\psi\nu(m^2 + k^2) - \theta\alpha g i k/(m^2 + k^2)$
temperature $\quad \theta_t = +\psi(\Delta T/H) i k - \theta\kappa(m^2 + k^2)$

Finally, the characteristic equation (dispersion) of the eigenvalue problem is obtained introducing the Fourier transform with the Ansatz exp(St). Introducing the Rayleigh number, $Ra = \alpha g H^3 \Delta T/\nu\kappa$, gives

dispersion $\quad (S/\nu + m^2 + k^2)(S/\kappa + m^2 + k^2)(m^2 + k^2) - k^2 Ra = 0$

The critical marginal instability (Ra_c) occurs at the smallest possible vertical wavenumber $m = 1$. Setting $d(Ra)/dk = 0$ for $S = 0$, introducing the aspect ratio $a = 1/\sqrt{2}$ or $k = am(\pi/H)$, one obtains the critical Rayleigh number

$Ra_c = \pi^4(1+a^2)^3/a^2 = \frac{27\pi^4}{4} \approx 675.5$, which is related to the most unstable mode (ψ_1, θ_1):

streamfunction $\psi = \psi_1 sin(\pi ax/H) sin(\pi_z/H)$
temperature $\theta = \theta_1 cos(xa\pi/H) sin(z\pi/H)$

Nonlinearity by advection of vorticity and thermal energy excites additional modes. As there is no vorticity advection due to the most unstable spatial mode, $J(\psi, \nabla^2\psi) = 0$, it is only the thermal energy advection, which excites an additional mode, $\theta_2 sin(2\pi z/H)$, because $J(\psi,\theta) = \frac{1}{2}\theta_1\psi_1 a(\pi/H)^2 sin(2\pi z/H)$. Thus the temperature advection $J(\psi,\theta) = (\psi_1\theta_2)a(\pi/H)^2[sin(\pi ax/H)sin(\pi z/H) + cos(\pi ax/H sin(3\pi z/H)]$, supports further modes besides the most unstable one, which is associated with the (ψ_1, θ_1)-amplitudes. The next smallest vertical wavenumber is $m^* = 2$, which describes horizontally homogeneous anomalies of the vertical temperature profile and adds opposing signs to the vertical distribution of temperature variability. Truncating the spatial eigenfunction expansion at this wavenumber

streamfunction $\psi = \psi_1 sin(\pi ax/H) sin(\pi_z/H)$
temperature $\theta = \theta_1 cos(xa\pi/H) sin(z\pi/H) + \theta_2 sin(2\pi z/H)$

and introducing it into the y-independent Boussinesq system gives the Lorenz model, once the orthogonality of the eigenfunctions has been utilised to derive the time evolution of the amplitudes $(\psi_1, \theta_1, \theta_2)_t$. Originally, this system was taken ad hoc from Saltzman's (1962) [69] numerical integrations of a higher order spectral model of the symmetric Benard-Rayleigh problem. Dimensionless time, wavenumbers, temperature, and streamfunction, $\tau = t\kappa(\pi/H)^2(1+a^2), k^2 + m^2 = (1+a^2)(\pi/H)^2$, $X = \psi_1 a/[\kappa(1+a^2)\sqrt{2}], Y = \theta_1\pi r/(\Delta T\sqrt{2})$, and $Z = \theta_2\pi r/\Delta T$ provide the final form of the Lorenz-System, with the Prandtl and relative Rayleigh number, $\sigma = \nu/\kappa$ and $r = Ra/Ra_c$ and another parameter $b = 4/(1+a^2)$ entering the last equation.

Tab. 2.8 - The Lorenz (1963) [51] system.

fields	tendency = damping + forcing	interaction
streamfunction	$X_\tau = -\sigma X + \sigma Y$	
temperature	$Y_\tau = -Y + rX$	-XZ
'stratification'	$Z_\tau = -bZ$	+XY

This Lorenz system (Lorenz 1963 [51], Sparrow 1982 [72]) (Tab. 2.8) is a set of nonlinear autonomous differential equations, $x_t = f(x)$, where $x = (X, Y, Z) \in R^m$ and f: $R^m \to R^m$. The three fixed points are (0,0,0) and $(\pm(b(r-1))^{1/2}, \pm(b(r-1))^{1/2}, r-1)$. The common parameter constellation ($\sigma = 10$, $r = 28$ and $b = 8/3$) provides chaotic flow whose predictability is analysed in the following. Fig. 2.4 shows the temperature and streamline fields linked to the location of the trajectory in the (X,Y,Z) state space.

Predictability experiments (attractor topology, uncertainty growth, and ensemble forecasts): Given f and an initial condition x_0, the trajectory x(t) is uniquely determined. Suppose that there is an uncertainty ϵ_0 in the initial condition, it is its growth which needs to be quantified. The dynamics of an infinitesimally small uncertainty is governed by the linearisation of the flow, $\epsilon_t = J(x)\epsilon$, where $J(x)$ is the Jacobian of f at

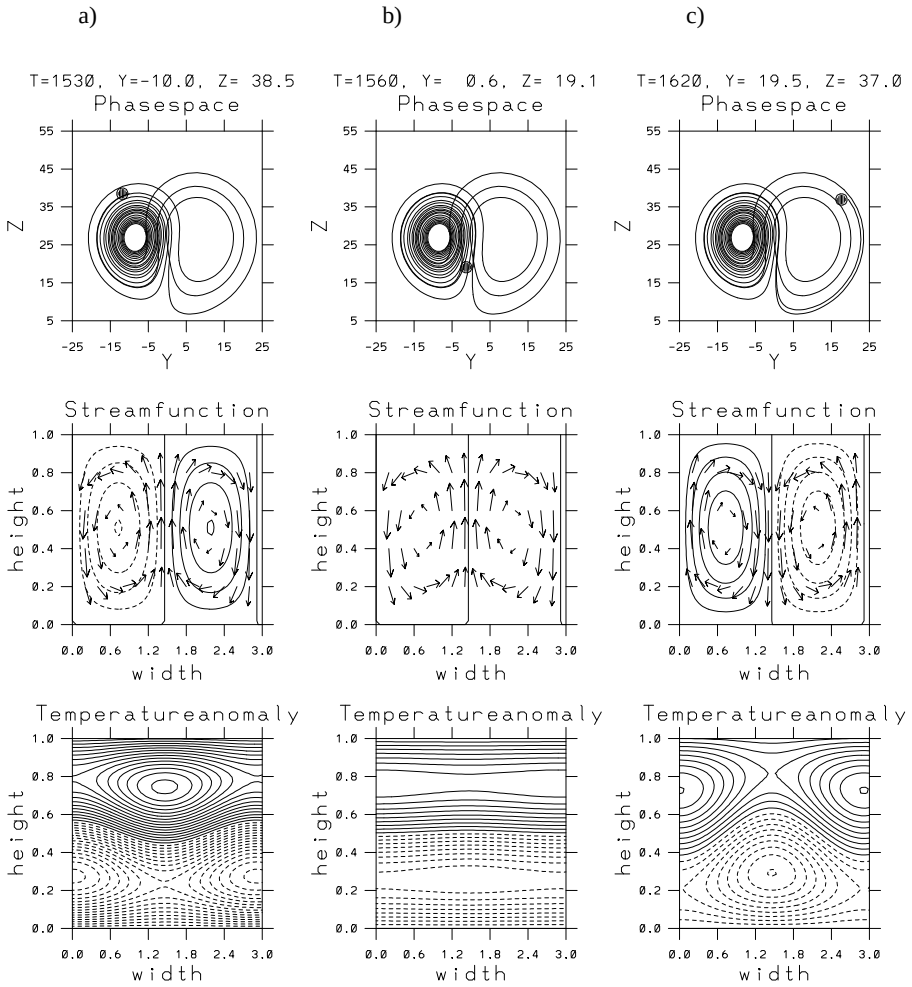


Fig. 2.4 - The Lorenz system trajectory in the (Y,Z)-plane: The three positions (full circles in top panel a to c) are associated with streamfunction fields with scaled velocity-vectors (middle panel) and temperature anomalies (bottom panel).

x. In this perfect model scenario, uncertainty is defined by the distance between the (inexact) forecast and the true state as a function of time. That is, the magnitude of uncertainty at time t is simply $\epsilon(t) = |\epsilon(t)|$. Two growth rates can be defined associated with the uncertainty dynamics (for applications and additional details see Ziehmann et al. 1995 [92], Smith et al. 1999 [73]).

Instantaneous uncertainty growth: Instantaneous growth is $g = \epsilon^{-1}\frac{d\epsilon}{dt} = \epsilon^{-1}(\epsilon(t)^T\epsilon(t))^{1/2} = \frac{1}{2}\epsilon^{-2}(\epsilon_t^T\epsilon + \epsilon^T\epsilon_t)$. Whether ϵ is increasing or decreasing (r less or greater than zero) depends on (i) the eigenvalues of J, (ii) on the projection of the uncertainty onto the eigenvectors, $\epsilon^T S$, and (iii) its decomposition into the eigenvectors $S^{-1}\epsilon$. The eigenvalues of J alone do not supply sufficient conditions to determine

the sign of g since the eigen-basis is, in general, not orthogonal. Combination with the linear uncertainty dynamics, $\epsilon_t = J(x)\epsilon$, gives $g = \epsilon^T[\frac{1}{2}(J^T + J)]\epsilon/\epsilon^T\epsilon$, where $\frac{1}{2}(J^T + J)$ is the symmetric (or deformation) part of the Jacobian with real eigenvalues and orthogonal eigenvectors (Tab. 2.9). The interpretation of the growth rate g is given by the Rayleigh quotient of ϵ_t. That is, the largest (smallest) eigenvalue of the deformation provides an upper (lower) bound on g for any ϵ at x. If all eigenvalues are positive (negative), g will be positive (negative) independent of the uncertainty orientation. Thus, if definite, the sign(g) is determined and only its magnitude depends on the orientation. For isotropically distributed uncertainties the ensemble average of instantaneous growth rates is given by the arithmatic mean, $<g>= \sum_m \lambda_i$, where the λ_i are the eigenvalues of the symmetric (deformation) part of the Jacobian.

Tab. 2.9 - Lorenz system linearised about (X_0, Y_0, Z_0).

matrix	Jacobian J		
streamfunction	$X_\tau = -\sigma X$	$+\sigma Y$	0
temperature	$Y_\tau = (r - Z_0)X$	$-Y$	$-X_0 Z$
stratification	$Z_\tau = Y_0 X$	$+X_0 Y$	$-bZ$
matrix	Symmetric $\frac{1}{2}(J^T + J)$		
streamfuncion	$X_\tau = -\sigma X$	$+\frac{1}{2}(\sigma + r - Z_0)Y$	$+\frac{1}{2}Y_0 Z$
temperature	$Y_\tau = \frac{1}{2}(\sigma + r - Z_0)X$	$-Y_0$	0
stratification	$Z_\tau = \frac{1}{2}Y_0 X_0$	0	$-bZ$

The exact bounds on uncertainty growth are obtained placing restrictions on every trajectory within regions of state space: (i) In regions G_1, where all eigenvalues of the Jacobian J have negative real parts, $Re(\lambda_i) < 0$, it is not sufficient to rule out positive growth rates for non-normal J. Yet these regions are observed (numerically) to be dominated by decreasing uncertainties. The eigenvalues of J are the roots of the characteristic polynomial $P(\lambda) = a_0\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$ with $a_0 = 1$, $a_1 = (1+\sigma+b), a_2 = \sigma(1+b-r+Z)+b+X^2, a_3 = \sigma(b+X^2-rb+bZ+XY)$. The Routh-Hurwitz criterion, which provides the subspace

$$G_1 = [x|T_0, ...T_m > 0],$$

states that all $Re(\lambda i) < 0$, if $T_0 = a_0 > 0$, $T_1 > 0$, $T_2 = a_1 a_2 - a_3 > 0$, and T_3 or $a_3 > 0$. (ii) In regions G_2, where each eigenvalue of the symmetric deformation map $\frac{1}{2}(J^T + J)$ is negative, $\lambda_i < 0$, no instantaneous growth rate is positive and all infinitesimal uncertainties will decrease with time for at least as long as the trajectory remains in that region. In general $G_2 \in G_1 \in R^m$. That is, all λ_i are negative definite only if all principal sub-matrices of $-(J^T + J)$ have positive determinants V_i. This provides the subspace

$$G_2 = [x|V_1, ...V_m > 0]$$

with $V_1 = 2J_{1,1} = 2\sigma > 0, V_2 = 4J_{1,1}J_{2,2} - (J_{1,2} + J_{2,1})^2 = 4\sigma - (\sigma + r - z)^2 > 0$ and $V_3 = det{-}(J + J^T) = 2b[4\sigma - (\sigma + r - z)^2] - 2y^2 > 0$. The fraction of time the system spends in G_1 or G_2 is related to the fraction of the attractor in this region. This reflects predictability (Fig. 2.5).

Effective growth: The linear resolvent operator (tangent propagator) of f: $M = M(x_0, \Delta t)$, maps ϵ_0 and x_0 to $\epsilon(\delta t)$ and $x(\Delta t)$ after Δt. Thus, the magnitude of

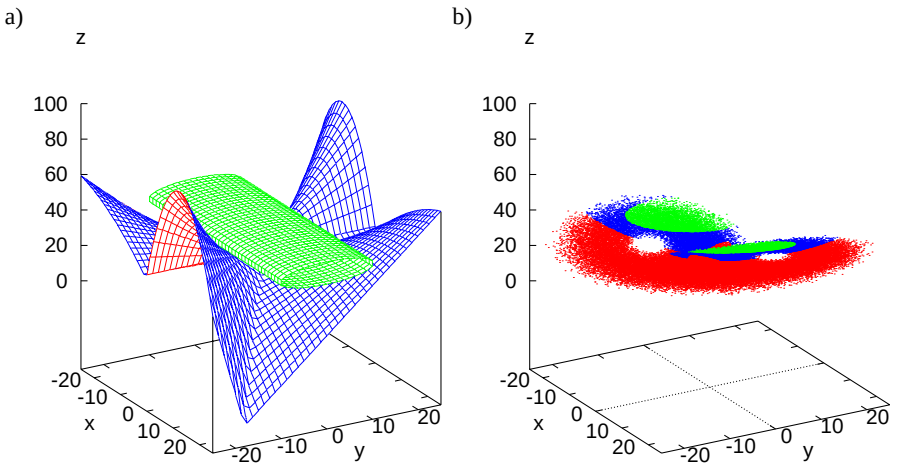


Fig. 2.5 - The Lorenz system with regions of uncertainty growth: (a) Surfaces of G_1 with blue and red side facing the stable und unstable region, and G_2 green. (b) Location of the attractor with blue, red and green dots lying within G_1 and G_2, outside G_1 and within G_2 (after Ziehmann et al. 1995 [92]).

uncertainty is $\epsilon(\Delta t) = |\epsilon(\Delta t)| = [\epsilon(\Delta t)^T \epsilon(\Delta t)]^{1/2} = [\epsilon_0 M_0^T M_0 \epsilon_0]^{1/2}$; its growth rate yields $r_{\Delta t} = \Delta t^{-1} log[(\epsilon_0 M_0^T M_0 \epsilon_0)/\epsilon_0^T \epsilon_0]^{1/2}$ (Tab. 2.10). The quadratic form associated with the uncertainty bounds the eigenvalues σ_i^2 of $M^T M$: Singular value decomposition (SVD) of M is $M = U\Sigma V^T$ with orthogonal matrices V(U) containing the right (left) eigenvectors as columns and Σ is the diagonal matrix of the singular values $\sigma_i > \sigma_j$ for $i < j$. Under M each right singular vector is rotated into the left singular vector and stretched (shrunk) by the factor σ_i greater (less) than one: $Mv_i = \sigma_i u_i$. The eigenvalues σ_i can be interpreted as finite-time Lyapunov exponents and the maximum growth rate $g_{\Delta t}$ is associated with the uncertainty aligned along the first right singular vector v_1. For $\Delta t \to \infty$, the growth rate $g_{\Delta t}$ approaches the largest Lyapunov exponent and the Lyapunov spectrum corresponds to the logarithms of the singular values of M. Accordingly, the first Lyapunov vector at each point on the attractor is $u_1(\Delta t)$. At a fixed location x_0 the singular value decomposition of $lim_{\Delta t \to \infty} M$ is evaluated over a bounded trajectory and thus this direction cannot be defined for (almost all) x_0. That is, the singular vectors can be determined from finite time local dynamics but the Lyapunov vector orientation cannot as they depend on an integral over the infinite past.

Tab. 2.10 - Uncertainty dynamics.

instantaneous growth $\epsilon_t = J(x)\epsilon$	$g = \epsilon^{-1} d\epsilon/dt = \epsilon^{-1}(\epsilon(t)^T \epsilon(t))^{1/2} = \frac{1}{2}\epsilon^{-2}(\epsilon_t^T \epsilon + \epsilon^T \epsilon_t)$ $= \epsilon^T[\frac{1}{2}(J^T + J)]\epsilon/\epsilon^T \epsilon$
effective growth $\epsilon(t_0 + \Delta t) = M_0(x_0, \Delta t)\epsilon_0$	$g_{\Delta t} = \Delta t^{-1} log\epsilon$ $= \Delta t^{-1} log[(\epsilon_0 M_0^T M_0 \epsilon_0)/\epsilon_0^T \epsilon_0]^{1/2}$

Ensemble forecasts: A perfect ensemble has members, which are analogs collected from a long time integration. They are nearest neighbors on the attractor with smaller dimension than the state space, where 'near' is defined by the statistical form of the observational uncertainty. Each member has equal weight (quantization noise). Thus

forming perfect ensembles requires more knowledge than specifying measurement uncertainties. Perfect ensemble and perfect model forecasts (with 512 members in a 6-bit box) evolving in time show the following (Fig. 2.6): (i) Early stages of predictability loss realise the ensemble spreading out. The spread develops quickly and becomes comparable with the diameter of the attractor. The information content of this distribution decays at a much slower rate than suggested by the standard deviation: Near t = 3.5 the root-mean-square error is near zero, although the system will have either large positive or large negative values. That is, this error measure has a fundamental shortcoming for non-linear systems. (ii) A return of skill is found for both large positive and negative X-values, which is due to the ensemble's residence in G_1 or G_2. This structure appears to occur in many low-order models. (iii) Saturation or the ultimate limit of predictability is reached at the time required for the ensemble to become statistically indistinguishable from climatology, which is the projection of the natural measure onto the variable of interest.

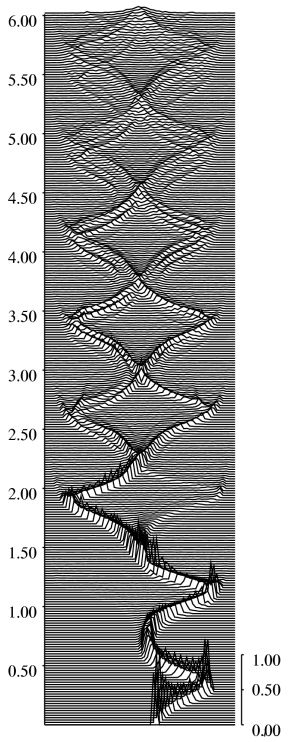


Fig. 2.6 - The Lorenz system with the evolution of variable x in a perfect ensemble: Probability density for the X-variable changing with time (plotted in the vertical, after Smith et al. 1999 [73]). The initially ($t = 0$) sharp distribution decays, but then shows a true return of skill (near $t \approx 0.4$) at the time the ensemble crosses the G_1 surface. The ensemble has members at each extreme of the attractor ($t \approx 2.75,\ 3.5$ etc.). The asymptotic distribution is not reached at the top of the panel ($t = 6.0$).

D. Analog forecasting: Local dimension and entropy

From the predictability of surrogate atmospheres in an imperfect and a perfect model environment we turn to the forecasting of the real weather. Two approaches are discussed: Forecasting of the future weather by past weather analogs is an internal prediction experiment. 'Atmospheric predictability as revealed by naturally occurring analogues' dates back to Lorenz (1969) [52]. Although analog forecasting suffers from data scarcity, because the atmosphere's recurrence time is vast compared with its life time (van den Dool 1994 [83]), it has been frequently revisited, for example, by adapting a forecast error minimising metric for the analog search (Fraedrich and Rückert 1998 [27]) to improve individual and ensemble mean forecasts of dynamical systems and of Hurricane tracks (Sievers et al. 2000 [71], Fraedrich et al. 2003 [32]).

Consider a time series displayed in a time-delay coordinate phase space of sufficiently large embedding dimension m. The spread of initial ensemble members (they satisfy the perfect ensemble hypothesis) is measured by their mean mutual distance, and its scaling with the number of members leads to the D_2-dimension. The K_2-entropy is the change of ensemble spread (perfect model hypothesis) with increasing embedding dimension by adding future time-delays to the phase space. This describes the main axes of an initially small sphere expanding into an ellipsoid. This expansion is formed by diverging trajectories in the time-delay coordinate phase space and, therefore, characterises predictability. A reliability test with surrogate data may also be applied (Fraedrich 1988 [15]). That is, scaling and time change of the analog ensemble spread (nearest neighbours in phase space) correspond to the correlation dimension D_2 and entropy K_2, estimated by the Grassberger-Procaccia (1984) algorithm [39] (Fig. 2.7, Fraedrich 1987 [14]).

Compared with internal analog forecasting, numerical weather prediction (NWP) models, which play an eminent role in practical weather forecasting, provide external forecast. Their advancement has been remarkable since the first concepts (Richardson 1922 [67]) emerged and first applications (Charney, Fjortoft, and von Neumann 1950 [9]) were made. Here, a recent estimate of local degrees of freedom of NWP bred ensemble forecasts (Patil et al. 2001) is critically appraised.

Analog ensemble forecasts (local dimension and predictability): Given a time series x(t) of observed local weather variables measured at time steps τ. Pieces of this time series, which commence at t_i and last for $(m-1)\tau$ time steps, are used to define local weather states (or points) embedded in a m-dimensional phase-space spanned by time-delay coordinates, $x_m(t_i)$. Other weather states $x_m(t_j)$ in the same phase space are called weather 'analogs' or nearest neighbors to $x_m(t_i)$, if both are independent pairs of points, $|t_i - t_j| > \tau$, in this space and their Euclidean distance, $d_{ij} = |x_m(t_i) - x_m(t_j)|$, is small (Tab. 2.11). Note that a delay or sampling time-lag, corresponding to the decorrelation time-scale, may guarantee linear independence of the phase-space coordinates.

Scaling analysis of the statistics on the numbers of analog-pairs (which depend on their distance phase space and its embedding dimension) leads to estimates of the dimension of local weather dynamics and its entropy (as a measure of predictability). First the cumulative number distribution (or correlation integral) $C_m(l)$ of all N points in phase space is obtained and then their scaling behaviour is analysed:

(i) Analog ensembles: Counting the relative number of $j = 1, ..., N-1$ analogs

Tab. 2.11 - Analog pairs in the delay-coordinate phase space.

analogs	delay coordinate phase space
analog pair	$x_m(t_i) = [x(t_i), x(t_i - \tau), ..., x(t_i - (m-1)\tau)]$
	$x_m(t_j) = [x(t_j), x(t_j - \tau), ..., x(t_j - (m-1)\tau)]$
pairwise Euclidean distance	$d_{ij} = \|x_m(t_i) - x_m(t_j)\|, \|t_i - t_j\| \gg \tau$

$x_m(t_j)$, whose distance is smaller than a prescribed threshold ($d_{ij} < l$) from the local base point $x_m(t_i)$, provides a cumulative number distribution, $\sum_{j=1}^{N-1} \theta(l - d_{ij})/(N - 1)$, which increases to one with increasing threshold l. Here the Heaviside function is used with $\theta(a) = 0$ or 1, if $a > 0$ or $a < 0$. Repeating this procedure for all $i = 1, ..., N$ base points, $x_m(t_i)$, and subsequent averaging gives the correlation integral (Grassberger and Procaccia 1983, 1984 [38][39], for an extension to small and noisy data see [18], and [4]).

$$C_m(l) = N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N-1} \theta(l - d_{ij})/(N - 1).$$

That is, averaging over all analog-pairs in all possible m-dimensional spheres of size l provides an analog ensemble climatology.

(ii) Perfect ensemble hypothesis and distance scaling of the correlation integral (at fixed embedding) leads to the mean correlation dimension D: Consider the number of pairs of points, which are homogeneously distributed on a line (surface or volume) embedded in an m-dimensional space. This number grows according to a linear (quadratic or cubic) power-law:

$$C_m(l) \sim l^D$$

with D = 1, 2, or 3 at fixed m = 1, 2, or 3. Analoguously, local weather states evolving in a delay-coordinate phase space are characterised by the dimension of that object which is covered by its dynamics. Scaling with $l \to 0$ and sufficiently high (and fixed) embedding m occurs only, if perfect ensemble members are counted in the m-dimensional sphere of radius l.

(iii) Perfect model hypothesis and time scaling of the correlation integral leads to the estimates of the local mean predictability (entropy or information loss): Extending the pieces of weather trajectories from (m-1) to m , or increasing the embedding dimension by one, reduces the number of analog pairs trapped in a sphere confined by the l-threshold. Their chance being trapped in the m-dimensional l-sphere decreases proportional to

$$C_m(l) \sim exp(-m\tau L_k).$$

Thus predictability represents the system's dynamics in delay coordinate phase space. It can be estimated by the sum of the positive Lyapunov (or characteristic) exponents L_k, which contribute to the expansion of an infinitesimally small sphere of initially close trajectories into an ellipse with K expanding main axes: $H = \sum_{k=1}^{K} L_k$ with $L_k > 0$.

(iv) Interpretation: The results obtained from the analysis of local variables should not suggest that the weather (say within a Rossby radius about the station) may be

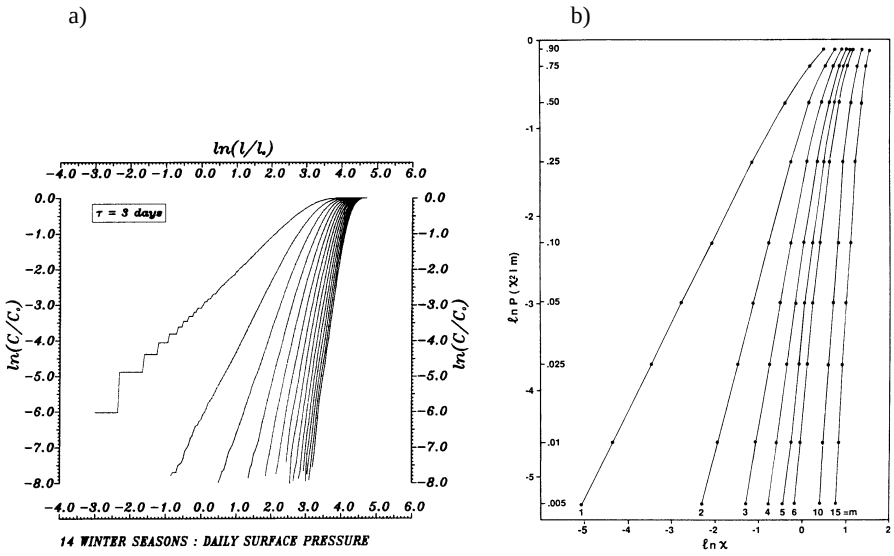


Fig. 2.7 - Correlation integrals (cumulative l-distance distribution), $C_m(l)$, changing with increasing embedding dimensions $1, ...(m-1)$: (a) Local surface pressure and (b) surrogate AR(1) determined analytically.

modelled on single station reading but that one projection of the weather attractor happens to reveal a relatively small dimension associated with a (climate) mean local predictability. Further distant stations are expected to characterise different regions and may lead to different estimates of local mean predictability and dimensionality (Fraedrich 1986 [13]).

E. Numerical prediction: Local degrees of freedom and forecast spread

Consider independent Gaussian random variables, y, with zero mean $\mu = 0$ and variance σ^2. Degrees of freedom, dof, characterise the linear independence of the members i = 1, ..., M of an ensemble y_i of these random variables. The ensemble statistic is conveniently described by the first and second moments: the ensemble's mean, $[y_i]$, and the ensemble's variance, var_M.

$$[y_i] = \frac{1}{M}\sum_{i=1}^{M} y_i \quad \text{and} \quad var_M = \frac{\chi^2}{M} = \frac{1}{M}\sum_{i=1}^{M}(y_i - [y_i])^2$$

Degrees of freedom: Sampling different realisations of ensembles shows that, due to the inherent degrees of freedom, their means and variances, $[y_i]$ and var_M, fluctuate and lead to distributions, which can be described theoretically: The ensemble means are Gaussian distributed and the variances are standardised chi-squared distributed, χ^2/M. Denoting Ex and Var as the first and second moments of sampled ensemble statistics (ensemble means and variances), one obtains the ensemble degrees of freedom, dof = M, by moment fitting. That is, realising that a chi-squared variable

(unweighted sum of normalised Gaussian variables) has the sample mean and variance, $Ex(\chi^2) = \sigma^2 M$ and $Var(\chi^2) = 2\sigma^4 M$, one obtains the following (Tab. 2.12):

(i) The combination of the sample averaged ensemble variances, $Ex(var_M) = \sigma^2$, which is related to the chi-squared distribution, and the sample variance of the ensemble means, $Var([y_i]) = \sigma^2/M$, which is related to the Gaussian distribution, gives $dof = M = Ex(var_M)/Var([y_i])$.

(ii) The second moment fit combines the (squared) mean and variance of the sample variances using the chi-squared distribution, $M = 2Ex^2(var_M)/Var(var_M)$, which, for σ-normalised variables, reduces to $dof = 2/Var(var_M)$.

Tab. 2.12 - First and second moments of samples of ensemble mean and variances and their relation to the ensemble degrees of freedom M.

ensemble	first moments	second moments
means	$Ex([y_i]) = 0$	$Var([y_i]) = \sigma^2/M$
variances	$Ex(var_M) = \sigma^2$	$Var(var_M) = 2\sigma^4/M$
degrees of freedom	$dof = Ex(var_M)/Var([y_i])$	$dof = 2Ex^2(var_M)/Var(var_M)$

Interpretation: Some of the statistical measures associated with the dof-estimate allow a direct geometrical or meteorological interpretation. (i) The sample mean of the ensemble variances characterises the ensemble spread $Ex(var_M)$ describing the average Euclidean distance of all ensemble members from their respective means $[y_i]$. In the perfect model/ensemble environment (see above sections), this provides a predicted forecast error. (ii) The sample variance of ensemble means, $Var([y_i])$, measures the fluctuation of the ensemble means due to the inherent degrees of freedom of its M members; likewise does the sample-variance of the ensemble-variances. (iii) The ensemble degrees of freedom, dof, increase with increasing ensemble spread $Ex(var_M)$ and decreasing sample variability of the ensemble means or variances, $Var([y_i])$ or $Var(var_M)$. (iv) The ensemble covariance or correlation matrix describes the sample averaged linear relation between all ensemble members. The eigenvector decomposition turns the second moment fitting from a dof-variance into a dof-eigenvalue notion, realising, $Ex(var_M) = \sum_{i=1}^{M} \lambda_i$ and $Var(var_M) = Var(\sum_{i=1}^{M} y_i^2/M) = 2M^{-2} \sum_{i=1}^{M} \lambda_i^2 = 2/dof$ (Fraedrich et al. 1995 [23], see also Wang and Sheng 1999 [89]). In this sense, dof-estimates provide a limit for the number of independent ensemble eigenvectors required to represent the ensemble's variability. (v) Thus an estimate of the confidence limits can be derived:

$$dof = 2Ex^2(var_M)/Var(var_M) = (\sum_{i=1}^{M} \lambda_i)^2 / \sum_{i=1}^{M} \lambda_i^2$$

$$\pm \Delta dof = \pm \nabla(dof) \cdot (\lambda_i) = \mp 4 dof (2/N^*)^{1/2}$$

The error-propagation leads to the error of the dof-estimate, where ∇ is the gradient with respect to the λ_i. And the North et al. (1982) [60] rule of thumb gives the confidence limit placed on the estimated spectrum of the eigenvalues, $\Delta\lambda_i = \pm 2\lambda_i (2/N^*)^{1/2}$, which depends on the number N^* of independent realisations in the sampling. For example, $N^* = 1000$ independent realizations lead to lower and upper confidence limits, $dof \pm \Delta dof = (1 \mp 4(2/N^*)^{1/2}) dof$, of about 18%. If all variables are mutually uncorrelated dof=M; if they are perfectly correlated (or anti- correlated), the lower bound of dof = 1. (vi) Note that normalisation matters for non-Gaussian

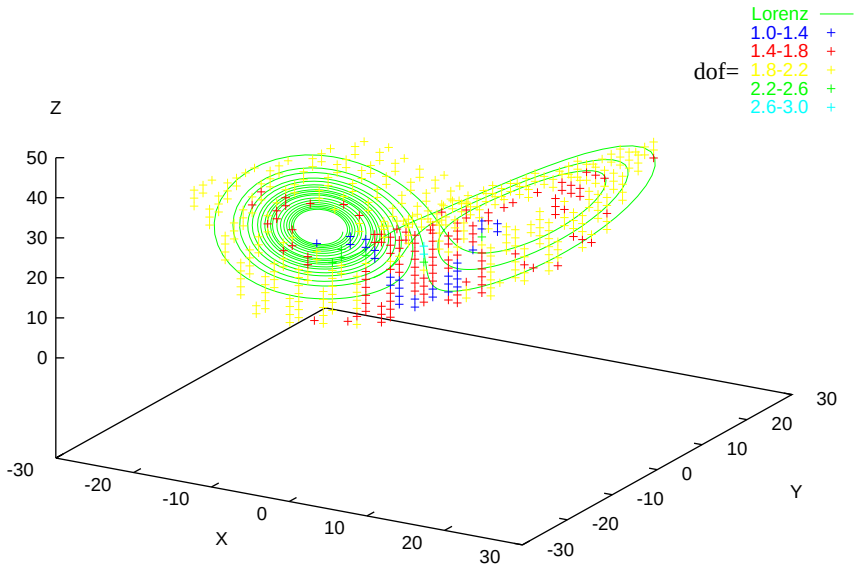
distributed data. For example, ellipses in the x, y-plane are subjected to dof-analysis after being modified by affine transformations: Rotation about the z-axis and stretching the major versus minor axes by the stretching factor a/b. The non-normalised or covariance dofs are independent of the orientation but decrease from dof = 2 to 1 with increasing stretching factor a/b, because the ellipse changes in shape from a circle, a/b = 1, to a line. The normalised or correlation dofs, $1 < dof < 2$, depend both on the rotation angle or orientation of the major (or minor) axis, and their ratio a/b. This is plausible considering the effect which normalisation (by the standard deviation) has on the embedding coordinates. That is, dof-fractality provides an indicator of shape and orientation of the object in the embedding space characterising its deviation from isotropy (Fraedrich 1997 [26]). (vii) Applications: In meteorology and climate research, time-space data sets are displayed in the (T,S)-data matrix, x_{ij}. Space or time sampling yield the following estimates for the degrees of freedom: Spatial ensemble statistics (spatial means [], and spatial variances, Var_S) fluctuating with time sampling is due to spatial degrees of freedom. Likewise in time: The time degrees of freedom characterise the time statistics (time means $<>$, and time variances, Var_T), which vary with spatial sampling. The degrees of freedom are estimated by first and second moment fitting; note that $(dof_T, dof_S) \leq (T, S)$:

$$dof_T = [var_T]/Var_S(<x_{ij}>) \qquad dof_S = <var_S> /Var_T([x_{ij}])$$
$$dof_T = 2[var_T]^2/Var_S(var_T) \qquad dof_S = 2\ <var_S>^2\ /Var_T(var_S)$$

Spatial degrees of freedom (Lorenz attractor): Applied to the Lorenz attractor, dof-estimates gives the following results: (i) The global dof-estimate, dof = 2, characterises the chi-squared distribution of the Euclidean distances between all states on the attractor from their arithmetic mean. Rotation about the Z-axis changes the values, $2 < dof < 3$, because they are not invariant under affine transformations. (ii) Local dof-estimates describe the distance distribution between a base point and its nearest neighbors (see Fig. 2.8a). The influence of sample size and the effect of varying noise-level can also be illustrated. (iii) The dof-estimates of higher dimensional systems tend to exceed the Lyapunov- or the correlation-dimension, D $< dof_s < M$. (iv) Global dof-estimates provide a minimum embedding dimension for measured variables, because they define the number of independent eigenvectors of the state space correlation matrix, onto which the time series needs to be projected.

Forecast ensemble degrees of freedom (NWP forecasts): A recent analysis (Patil et al. 2001 [61]) suggests local weather dimensionality limited by $dof_T \sim 5$. This conclusion can be reached when proceding as follows: (i) Pick an ensemble of 5 local wind-weather forecasts (12-hourly, 500mbar zonal and meridional wind components) and sample the ensemble statistics over 50 realisations (or 25 gridpoints for the zonal and meridional wind) within a $1100 \times 1100km$ region. The NWP ensemble forecasts are obtained by a breeding procedure (Tab. 2.13) originally developed to estimate maximum Lyapunov exponents (following Wolf et al. 1985 [91]). (ii) Estimate the ensemble degrees of freedom (dof) by fitting the second moment of the sample to the theoretical chi-squared distribution utilising the eigenvalue formulation (Fig. 2.8b). A global climatology may be composed repeating this procedure for all regions and average over 150 days. (iii) Interprete these dofs, which give the number of ensemble (breeding or Lyapunov) eigenvectors required to describe the ensemble's variability, as a 'local low dimensionality of atmospheric dynamics' and conclude that this may provide hints of where to improve data assimilation.

a)



b)

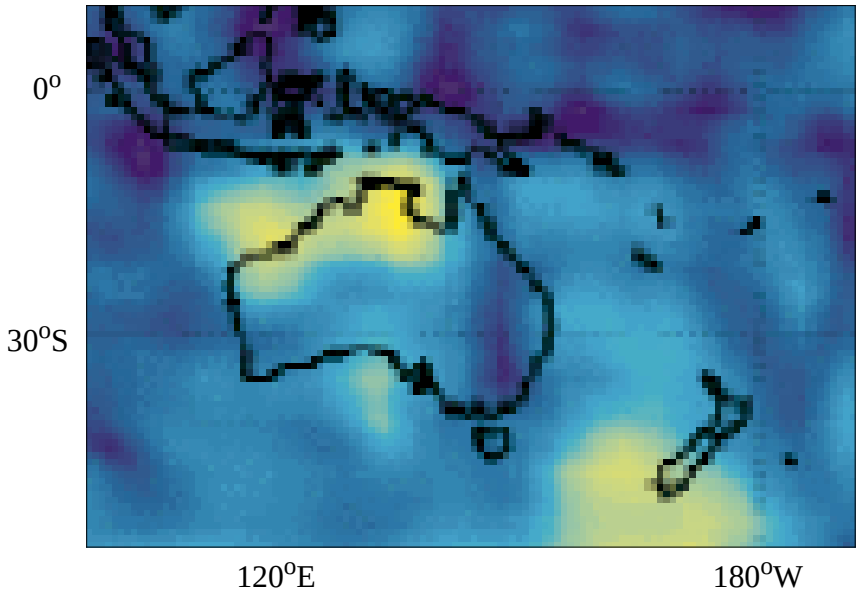


Fig. 2.8 - Degrees of freedom: (a) Spatial degrees of freedom (dof_S) at locations on the Lorenz (1963) [51] attractor; (b) local ensemble degrees of freedom estimated from ensemble forecast statistics (sampled over 25 grid points) for a 36hour NWP forecast in the Austral-Asian region (from Patil et al. 2001, shading from light to dark: $1 \leq dof_T \leq 5$).

Tab. 2.13 - Ensemble breeding method: A practical approach to generate ensemble members.

Breeding (in Δt-steps)	Control	Member	Error and Perturbation
(1a) Analysis A, perturbation P	A(0)	A(0)+P(0)	$E(0) = <P(0)>$
(1b) 1st forecasts $F(1)$, $F_1(1)$	$F(1)$	$F_1(1)$	$E(1) = <F_1(1) - F(1)>$
(2a) Analysis A(1), rescaled P(1)	A(1)	A(1)+P(1)	$P(1) = (F_1 - F) \cdot E(0)/E(1)$
(2b) 2nd forecasts $F(2)$, $F_1(2)$	$F(2)$	$F_1(2)$	$E(2) = <F_1(2) - F(2)>$
: Generate new members			
(n) Saturation (after days)	$F(n)$	$F_1(n)$	$E(n) = <F_1(n) - F(n)>$
short-term error growth			$dlnE/dt \sim 1.5d$

Some comments are in order: (i) Initially (t = 0) the numerically generated forecast ensembles can hardly be distinguished, because they start from very close initial conditions; that is, $dof(t = 0) \sim 1$. With increasing forecast lead time ($t = \Delta t$) the spread and, therefore, the degrees of freedom grow, $1 \leq dof(t = \Delta t) \leq 5$. (ii) Local predictability can be estimated by the ensemble spread, $Ex(var_M)$, and by the degrees of freedom, $dof = 2Ex^2(var_M)/Var(var_M) = (\sum_{i=1}^{M} \lambda_i)^2 / \sum_{i=1}^{M} \lambda_i^2$. Both measures are similar: The spread describes the mean Euclidean distance between ensemble members and their mean. It is a measure of the divergence in terms of spread per forecast lead time. The dof-estimate characterises the representative number of ensemble (Lyapunov) eigenvectors, which participate in the dispersion of the model forecast. (iii) Note that the sampling of 25 close grid points of a local wind-field map results in only $N^* \sim 2$ to 4 independent realisations when assuming a 500km Rossby radius. That is, the confidence interval is almost as large as the dof-estimate, and this requires improvements.

III. LONG TERM MEMORY - LOW FREQUENCY VARIABILITY

In the mid-70s the Brownian motion entered climate research as a paradigm for the Earth's climate fluctuations (Hasselmann 1976 [42]). Employing the Langevin approach to a greenhouse climate relaxing towards its equilibrium, for example, one obtains its high frequency red and a low frequency white noise temperature response on white noise stochastic forcing. As the autocorrelation of this system decreases exponentially with finite integral time scale, there is no indication of long-term memory. Still, such concepts have stimulated an intensive red noise search in observed data and simulations of comprehensive general circulation models (GCM). At the same time, observations and modeling of flicker noise or other power-law scaling regimes emerged (Voss and Clark 1976 [84], van Vliet et al. 1980 [85]) with new concepts, which are also close to the climate system's energy balance; that is, replacing the Newtonian or radiative-convective relaxation by the diffusive heat flux divergence and the forcing by additive noisy fluxes (and not by flux divergences). This system develops a power-law decorrelation (or power spectrum) leading to an infinite integral time scale, which indicates long-term memory. Since then power-law power spectra different from Brownian motion have been identified in observed records and model simulations of the climate system. For example, the near surface temperature (Manabe and Stouffer 1996 [56], Pelletier 1999 [62]) shows low-frequency behaviour which does

not, up to very long periods, asymptote towards a white plateau. While most of these studies (for a review see Pelletier and Turcotte 1999 [63]) are guided by self-affine scaling laws governing the non-linear dynamics, the associated long-term memory or correlation aspect has been emphasised only recently (Koscielny-Bunde et al. 1998 [50], Talkner and Weber 2000 [78]) analysing observed temperatures and GCM simulations (Müller et al. 2002 [59], Fraedrich and Blender 2003 [34]). In most of these studies, the variance spectrum analysis has been replaced by methods based on the analysis of a random walk (time series of accumulated anomalies) to determine the scaling laws and long-term memory: Fluctuation and detrended fluctuation analysis (FA and DFA, see e.g. Peng et al. 1994 [64]). These analyses suggest that the near surface temperature fluctuations are governed by a universal scaling behaviour showing long-term memory correlations up to twenty years. This feature has been attributed to the slow ocean and cryosphere dynamics also affecting the atmosphere as the fastest of the climate components. But, a quantitatively similar long-term memory has been identified in the dynamical properties (global angular momentum) of a GCM (PUMA) (Fraedrich et al. 1998 [28]) driven by Newtonian cooling and damped by Rayleigh friction. Both the observational support and the understanding of stationary climate processes with long-term memory is far from being complete, not to mention the existence of generally accepted conceptual models. Here we present information on the low-frequency behaviour of the temperature as the most important variable of the surface heat balance, which is analysed using observations and simulations. First, the method is briefly described and then applied to long time series of the near surface air and soil temperature observed and simulated. Finally, a simple conceptual model comprising Fickian diffusion and Newtonian cooling is introduced as a suitable climate concept which, in contrast to the conventional Brownian motion analog or red noise, reveals a long-term memory regime. The subsequent sections introduce dynamical mechanisms of low-frequency atmospheric variability generated by the North Atlantic Oscillation life-cycles initiated by spatial resonance. Finally, decadal variability is described by a coupled ocean-atmosphere mode, which establishes itself by a positive feedback loop modulated by the annual cycle.

A. Long-term memory everywhere?

Long-term memory and scaling regimes (correlation, fluctuation and power spectrum) are deduced from time series x_j, which are interpreted by random walk positions or the Y(j)-profile running through time j, where $Y(j) = y_1 + y_2 + ... + y_j$. The random step (or anomaly), $y_j = x_j - <x_j>$, is obtained after subtracting the ensemble mean annual cycle $<x_j>$. Thus the random walk profile starts from the zero position and ends there. The fluctuation analysis applied to the profile is based on the structure function approach (Monin and Yaglom 1975 [58]) to obtain scaling properties:

$$F(r) = [(F_1^2 + F_2^2 + ... + F_N^2)/N]^{1/2}$$

The steps or squared distances, $F_m^2(r) = (Y_{(m)r+1} - Y_{(m-1)r+1})^2$, change with segment length r and are averaged over all segments of the profile, that is from m = 1 to N. Stationarity and normalisation by the total variance σ^2 yields the following relation between auto-correlation and structure function, $C(r) = 1 - \frac{1}{2}F(r)$. That is, a power-law scaling of the auto-correlation corresponds to a power-law scaling of the

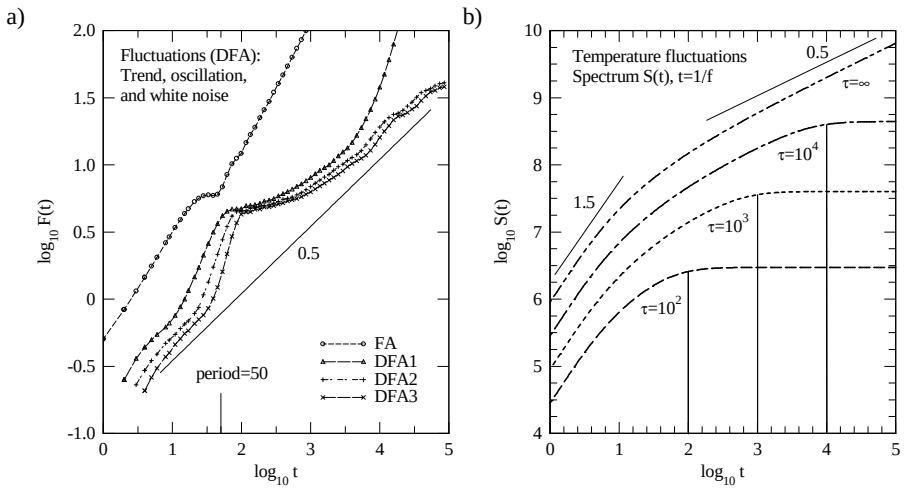


Fig. 3.1 - Power law scaling of simple systems: (a) Fluctuation (FA) and detrended fluctuation analysis (DFA1 to 3) of the artificial time series given by a 50 day wave superimposed by a linear trend and white noise. (b) Normalized variance spectra of mixing-length averaged temperature anomaly fluctuations in a simple energy balance climate model, which is characterized by one-dimensional Fickian diffusion excited by white noise fluxes and damped by Newtonian cooling: The diffusivity and mixing length are fixed at $D = 10^6 m^2 s^{-1}$ and $L = 1000 km$, the Newtonian relaxation time (in days) varies with $\tau = 10^2,\ 10^3,\ 10^4$ and ∞.

structure function, $C(r) \sim r^{-\gamma}$ and $F(r) \sim r^{\alpha}$, with $\alpha = 1 - \frac{1}{2}\gamma$. Fourier transformation of the correlation function leads to the variance spectrum, which is also governed by a power-law, $S(\omega) \sim \omega^{-\beta}$ for $\omega \to \infty$ with $\beta = 1 - \gamma = 2\alpha - 1$.

Short- and long-term memory can now be introduced by time series analysis. They are formally distinguished by their integral timescale $\tau = \int_0^\infty C(r)dr$ being finite or infinitely large. Thus stationary processes with long-term memory are limited in the range of exponents to exclude divergence of moments:

$$\begin{array}{lll} \text{correlation} & C(r) \sim r^{-\gamma} & \text{for } 0 < \gamma < 1 \\ \text{fluctuation} & F(r) \sim r^{\alpha} & \text{for } 1 > \alpha > \frac{1}{2} \\ \text{spectrum} & S(\omega) \sim \omega^{-\beta} & \text{for } 1 > \beta > 0 \end{array}$$

The scaling range is limited by white noise ($\gamma = 1, \alpha = \frac{1}{2}$, and $\beta = 0$), which is stationary without memory (zero or finite integral time-scale) and by Flicker- or 1/f-noise ($\gamma = 0, \alpha = 1$, and $\beta = 1$), which is non-stationary with long-term memory (infinite integral time scale). A pure random walk, for example, attains $\gamma = -1, \alpha = \frac{3}{2}$, and $\beta = 2$. Fractional stochastic processes (Tab. 3.1) (Hosking 1981 [44]), for example, show long-term memory. The statistical properties associated with it are power-law power spectra in the low frequency range (with $\beta = 2d$), the power-law decorrelation with increasing time-lag, $C(r) \sim k^{2d-1}(-d)!/(d-1)!$, and an infinitely long integral time scale τ, if $0 < d < \frac{1}{2}$. Not unlike the auto-regressive or AR-models being fitted to data with short-term memory, the fractional AR-models are used to determine the scaling behaviour of time series with long-term memory. Next, the detrended fluctuation analysis is introduced.

Tab. 3.1 - Fractional auto-regressive process (Hosking 1981 [44]).

moving average	$x(t) = \sum_{k=0}^{\infty} a(k) z_{t-k}$	
weights ($d < \frac{1}{2}$)	$a(k) = (k+d-1)!/k!(d-1)!$	$k^{d-1}/(d-1)!$ for $\omega \to \infty$
auto-correlation	$C(r) = (-1)^k(-2d)!/(k-d)!(-k-d)!$	$k^{2d-1}(-d)!/(d-1)!$ for $k \to \infty$
spectrum	$S(\omega) = (2sin\frac{1}{2}\omega)^{-2d}$	ω^{-2d} for $\omega \to \infty$
integral time-scale	$\tau = \sum_{r=0}^{\infty} C(r) = \infty$ for $0 < d < \frac{1}{2}$	

Detrended fluctuation analysis (DFA): The power-law scaling of the fluctuations is not affected, if the squared distances $F_m^2(r) = (Y_{(m)r+1} - Y_{(m-1)r+1})^2$ of all time-segments are replaced by variances $F_m^2(r) = Var(Y'_{(m)r+j})$ within the segments. Now, detrended fluctuation analysis goes a step further. It removes the linear, quadratic, cubic, etc. trends from the segments fitting linear or higher order polynomials to all segments of the profile before the segmential variances are determined (see Peng et al. 1994 [64], Bunde et al. 1998 [50], Heneghan and McDarby 2000 [43]). Note that, conceptually, the linear detrending of the time-lag intervals returns the accumulated anomalies to the anomalies. Thus, a proper linear detrending of the anomalies commences for quadratic, cubic, etc. detrending of the profile. This can be seen when analysing a surrogate data set, which serves as an introductory example for the DFA methodology. The fluctuation analysis based on the structure function (FA) and the detrended fluctuation analysis employing linear, quadratic and cubic polynomials (DFA, DFA2, and DFA3) are applied to an artificial time series which consists of a linear trend plus a periodic 50day-wave plus a first order auto-regressive or red noise process (Fig. 3.1a). All methods show the periodicity. Only the detrended fluctuation analysis, employing trend-removing polynomials of second and higher order, capture the basic low frequency fluctuation feature of the time series, which is white noise. That is, for long time lags, FA and DFA do not scale with the white noise power-law, $\alpha = \frac{1}{2}$, as DFA2 and DFA3 do. Therefore, DFA2 will be used for further analysis.

Air and soil temperatures (observed and simulated): Temperature time series of local air and soil columns are subjected to the DFA analysis to identify regimes of long-term memory. This is the observed near surface air and soil temperature at the station Potsdam (Germany, 1893-2001, Fig. 3.2a) and simulated air temperature in continental Europe (coupled atmosphere-ocean GCM ECHAM-HOPE, 1000 year control run, Fig. 3.2b): (i) The observed near surface air (and soil) temperatures show a power-law scaling regime, which extends over two decades from less than one month to about several decades and longer. The DFA-scaling exponent $\alpha \sim 0.7$ corresponds to a power-law variance spectrum $\beta \sim 0.4$ (plotted with increasing log of the period). This characterises (a) the long-term memory with long lasting persistence of temperature anomalies and (b) the invariance of the same power-law scaling for a long time span. The occasional break down at long periods may indicate a transition to a new scaling regime ($\alpha = 0.9$ or $\beta = 0.8$), a very long quasi-periodic oscillation, or it may simply occur, because the results are less reliable for time scales close to the record length. In the soil, the long-term memory regime extends to about 2m. At the lowest layer of 12m, a red spectrum dominates the fluctuations in the time range between years and several decades. There is a transition region between 2 and 12m soil depth, where a flicker noise regime characterises the temperature variability in the time span between months and many decades. (ii) Simulated air temperatures are required to supplement the observations, because long records of tropospheric temperature ob-

servations are not available. The long-term memory regime, $\beta \sim 0.4$, extends well into the troposphere turning white between the mid-troposphere and the tropopause. The simulated long-term memory regime in the soil is comparable to the observed, up to the levels computed. (iii) Whether the long-term memory regime, which governs the temporal temperature fluctuations between the mid-troposphere and the near surface soil layers, is a universal phenomenon, has not been sufficiently resolved; coupled GCM simulations and observations indicate different behaviour over the oceans and in the tropics. In summarising: A power-law scaling different from that generated by the Brownian motion analog of climate variability, governs the fluctuations on time scales between a month and several decades or even century. That is, the scaling, which characterises long-term memory, remains unchanged over a large time span. This suggests a suitable cooperation of fundamental processes contributing to the climate variability, which are different from those generating red noise or Brownian motion.

In the Brownian motion case, the fundamental climate processes may be attributed to Newtonian cooling and additive stochastic forcing. Newtonian cooling relaxes the earth system to its greenhouse equilibrium by radiative-convective heat exchange; the relaxation time scale τ increases with the system's heat capacity (for simple climate models see, for example, (Fraedrich 2001 [30])). Additive white noise represents uncorrelated fluctuations of the atmosphere, which forces the climate fluctuations out of equilibrium. Thus the climate response shows two regimes: The red spectrum, $\beta = 2$, defines a transitional frequency regime, $\omega > \tau^{-1}$, where the damped system responds on the short frequency part of the white noise forcing before, for long periods, $\omega < \tau^{-1}$, the low frequency part of the white forcing dominates the system's response approaching the white noise plateau.

Fickian diffusion is another fundamental concept describing the variability of the climate system, which characterises its lateral and/or vertical Austausch. It exists besides Newtonian cooling which, as a parameterisation of the radiative-convective processes, relaxes the system towards its greenhouse equilibrium. Here Fickian diffusion dominates the heat balance and may generate long-term memory. Diffusive and random fluxes provide the equator to pole heat exchange by synoptic scale eddies. In addition, the turbulent mixing affects a large spatial domain (mixing length) within a relatively short time, measured by the Rossby-radius and the diffusion coefficient. Finally, Newtonian relaxation of the near surface properties to the deep soil restoration temperature, introduces a very long time scale, beyond which the system stabilises about white noise fluctuations. A toy model is presented (Fraedrich 2002 [33]).

A toy model (one-dimensional diffusion forced by random fluxes): A simple linear reaction-diffusion equation can be used to describe a meridionally extending energy balance climate. The tropospheric temperature, $T = \theta + T_E$, fluctuates about a globally averaged restoration temperature T_E (of the deep soil, for example). These temperature fluctuations θ are induced by two processes: (i) Meridional eddy heat fluxes in the troposphere are parameterised by Fickian diffusion plus a random volume noise, $F = -D\theta_y - F'$, and (ii) the near surface soil relaxes towards the equilibrium deep soil temperature by Newtonian cooling, $-\theta/\tau$:

$$\theta_t = -\theta/\tau + (D\theta_y + F')_y$$

Random heat flux and its intensity, $< F'(y,t)F'(y',t') >= 2\pi F_0^2 \sigma(y-y')\sigma(t-t')$ with $dim(F_0^2) = K^2 m^3 s^{-1}$, may be attributed to the basically white noise fluctuations of the water-cycle. The Fickian diffusion represents the eddy heat fluxes, which

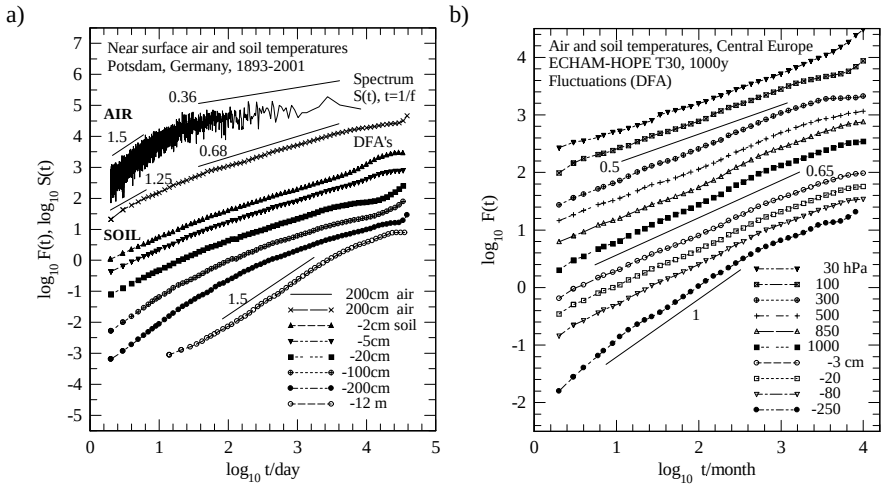


Fig. 3.2 - Temperature time series subjected to detrended fluctuation analysis (DFA2): (a) Observed near surface air and soil temperatures at Potsdam (Germany, 1893 to 2001); stations in other climates show similar results, (b) simulated air temperatures in central Europe (100 years, ECHAM-HOPE T30).

is the most prominent contributor to mix lower tropospheric temperature contrasts. The diffusion coefficient $D =< V\Delta y >$ may be scaled by meridional particle displacement, $\Delta_y = 2L$, and the teleconnection life-time, $t \sim 20days$ (section III B), as the mixing time scale. It provides the short-term limit of the climate system's long-term memory, which is of the magnitude of the life cycle of the North-Atlantic Oscillation (NAO). With the Rossby-radius as mixing-length, L=1000km, and the advective time-scale, V = L/t, one obtains $D \sim 2L^2/t \sim 10^6 m^2 s^{-1}$ (see also Green 1970 [41]). The subsequent analysis (see Voss and Clarke 1976 [84]) applies Fourier transform, so that the temperature anomaly $\theta(k,\omega) = ikF_0(k,\omega)/(Dk^2 + \tau^{-1} - i\omega)$ in (k,ω)-space and the response spectrum $S(k,\omega) = \theta\theta^*$ with the conjugate complex θ^* can be deduced

$$S(k,\omega) = F_0^2 k^2/[(Dk^2 + \tau^{-1})^2 + \omega^2]$$

Integrating over the 2L mixing-domain gives a mean mixing temperature $[\theta]$. It is controlled by the diffusion due to eddy life-cycles and describes the tropospheric and near surface soil temperature, $[\theta(t)] = (2L)^{-1} \int_{-L}^{L} \theta(y,t)dy$:

$$[\theta(t)] = (2L)^{-1} \int_{-L}^{L} (y,\omega)dy = (2\pi)^{-1} \int_{-\infty}^{\infty} sin(kL)\theta(k,\omega)(kL)^{-1}dk$$

Now the variance spectrum $S(\omega)$ of the time fluctuations of the temperature averaged over the mixing domain can be determined:

$$\begin{aligned} S(\omega) &= F_0^2(2\pi)^{-2} \textstyle\int_{-\infty}^{\infty} dk k^2 sin^2(kL)[(Dk^2 + \tau^{-1})^2 + \omega^2]^{-1}(kL)^{-2} \\ &= F_0^2(2\pi)^{-2} L^{-2} \cdot \pi D^{-1/2}(\tau^{-2} + \omega^2)^{-3/4}[sin(1/2\phi) \\ &\quad -exp(-\beta)[cos(1/2\phi)sin(\alpha) + sin(1/2\phi)cos(\alpha)]]/sin(\phi) \end{aligned}$$

where $\phi = arctan(\tau\omega)$, $\alpha = 2L\mu sin(\frac{1}{2}\phi)$, $\beta = 2LD^{-\frac{1}{2}}(\tau^{-2} + \omega^2)^{-1}/4cos(\frac{1}{2}\phi)$ with $dim(S) = K^2 s$. The spectrum is characterised by two fundamental time scales (or frequencies), which govern the deterministic dynamics of the system's mixing-length averaged temperature fluctuations: (i) The temperature relaxation time scale is associated with the Newtonian cooling and characterises the slow relaxation process towards the deep soil restoration temperature. (ii) The diffusive mixing, $\omega_0 = \frac{1}{2}D/L^2 \sim 1/10$days, is short due to the fast mixing processes in the lower atmosphere. The following three spectral regimes are identified (Fig. 3.1b): (a) High frequencies show a slope ($\omega_0 \ll \omega$): $S(\omega) \sim \omega^{-3/2}$. (b) Intermediate frequencies are characterised by long-term memory effects, which are limited by the time scale of the slow climate component ($\omega_\tau \ll \omega \ll \omega_0$): $S(\omega) \sim \omega^{-1/2}$. (c) At very low frequency, the white noise plateau is reached ($\omega_\tau \ll \omega \ll \omega_0$): $S(\omega) \sim \omega^0$. The white noise response level at zero frequency is $S(\omega_0) = F_0^2(2\pi)^{-2}L^{-2} - \frac{1}{4\pi}D^{-1/2}\tau^{3/2}[1 - exp(-\beta)]$. Neglecting Newtonian cooling ($\tau \to \infty$), that is $\mu \to \omega^{1/2}, \phi \to \frac{1}{2}\pi, \alpha = \beta \to a(2\omega)^{1/2}$, means that the system cannot be attracted by the stable equilibrium but the variance density increases $S(\omega) \to \infty$ for $\omega \to 0$:

$$\begin{aligned} S(\omega) &= F_0^2(2\pi)^{-2}\int_{-\infty}^{\infty} k^2 sin^2(kL)(D^2k^4 + \omega^2)^{-1}(kL)^{-2}dk \\ &= F_0^2(2\pi)^{-2}L^{-2}\pi D^{-1/2}(2\omega)^{-3/2}[1 - exp(-Z)[sin(Z) + cos(Z)]]Z^{-3} \end{aligned}$$

where $Z^2 = \omega/\omega_0$ is the frequency normalised by the natural one associated with the diffusion process, $\omega_0 = \frac{1}{2}D/L^2 \sim 1/10$days. This special case reveals two power-law scaling regimes: For low frequencies ($\omega < \omega_0$) the spectrum follows $S(\omega) \sim \omega^{1/2}$, and for higher frequencies ($\omega > \omega_0$) one obtains $S(\omega) \sim \omega^{-3/2}$. This system is similar to the one analysed by Voss and Clarke 1976 [84], which has been adopted as a paradigm for the observed power-law or self-affine scaling of the climate system's low-frequency fluctuations by Pelletier and Turcotte 1999 [63]. However, they use the vertical heat flux divergence associated with vertical overturning and a vertical mixing length, which affects the whole depth of the atmosphere.

Resonance: The long-term memory of the near surface temperatures is associated with the diffusive energy flux affecting the length 2L of a mixing domain. That is, the temperature fluctuations (averaged over the 2L-domain) are long-term correlated for frequencies, $\omega < \omega_0 = \frac{1}{2}D/L^2 \sim 1/10$days, below the natural frequency representing the diffusion process. This may be characterised as a resonance effect (Benzi et al. 1982 and 1989 [6][5]), which commences at periods long enough for diffusive mixing to spatially correlate temperatures across the mixing length 2L. Now the (domain averaged) temperature fluctuations become time correlated, not unlike a positive feedback, so that a long-term memory regime evolves with a process related power-law spectrum, which extends over a wide frequency band. The low frequency limit of this regime is set at the relaxation time-scale, ω_τ, of the stabilising influence of Newtonian cooling. The high frequency limit, ω_0, defines the diffusion time scale, up to which temperature fluctuations at the boundaries of the mixing domain remain independent of one another, so that a long-term memory cannot evolve in the mixing domain.

B. Teleconnections: Observations, idealised experiments, and mechanisms

The large scale weather or teleconnection pattern, which governs the European climate, reveals also long-term memory. This is the North-Atlantic Oscillation (NAO), which is often quantified by an index given by the surface pressure difference between the Iceland Low and the Azores High in winter. The index decribes the zonality of the cross-Atlantic flow on a monthly mean basis (Fig. 3.3a), although the NAO life cycle lasts typically only 10 to 20 days. The detrended fluctuation analysis (Fig. 3.3b) shows scaling regimes of NAO indices obtained from recent observational (Hurrell 1995 [46]) and historical climate records (Luterbacher et al. 1999 [55]), simulations with comprehensive coupled ocean-atmosphere models (Raible et al. 2001 [66]), and idealised experiments with a simplified atmospheric GCM (Franzke et al. 2000 [36]). With the power law slope $\alpha \sim 0.58$ the long-term memory is less obvious in the NAO pressure index than in the temperature data. This may be due to the strong influence of the weather related processes dominating the intense fluctuations of pressure or wind. For observational data the scaling obtained by DFA-analysis agrees well with an adapted fractional auto-regressive process (see above, Stevenson et al. 2000 [77]), but it differs from the simulated indices.

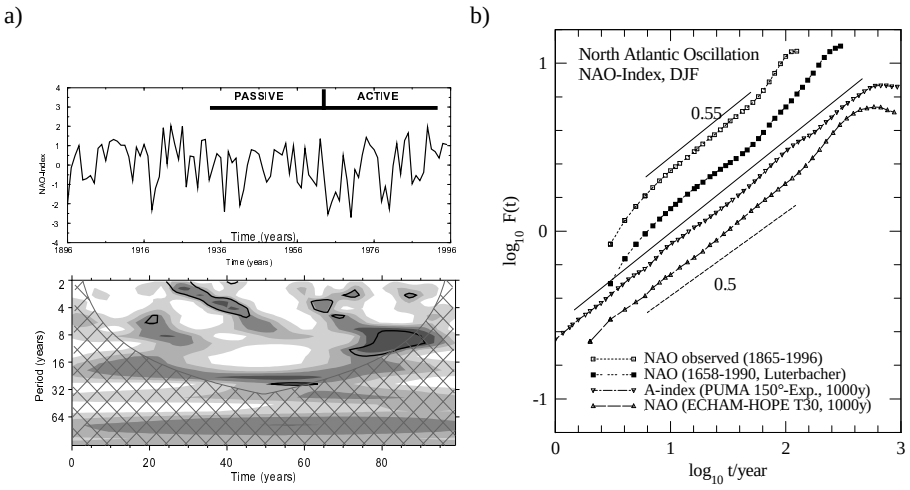


Fig. 3.3 - North Atlantic Oscillation (NAO): (a) The NAO-index as the difference between the normalized sea level pressure at Ponta Delgada, Azore, and Stykkisholmur-Reykjavik, Iceland, averaged over the winter season DJF (data from Hurrel, top panel) and its wavelet power spectrum (based on Morlet wavelet; cross-hatched region corresponds to the cone of influence where zero padding has reduced the variance; software by www.ResearchSystems.com, bottom panel); (b) detrended fluctuation analysis (DFA2) of the observed, recontructed and simulated NAO indices.

Climatological embedding: The North-Atlantic oscillation and the Pacific-North America pattern are embedded in an environment of jetstreams and stormtracks, which represent the atmosphere's first and second moments (Fig. 3.4a-c). Comparing the figures, one observes that the PNA pattern emerges with a wavetrain-like structure in regions, where the Subtropical Jet in 30° latitude attains maximum intensity poleward of the largest heating over the tropical Pacific. The NAO develops as a North-South dipole (Iceland Low and Azores High) in association with an eddy driven or

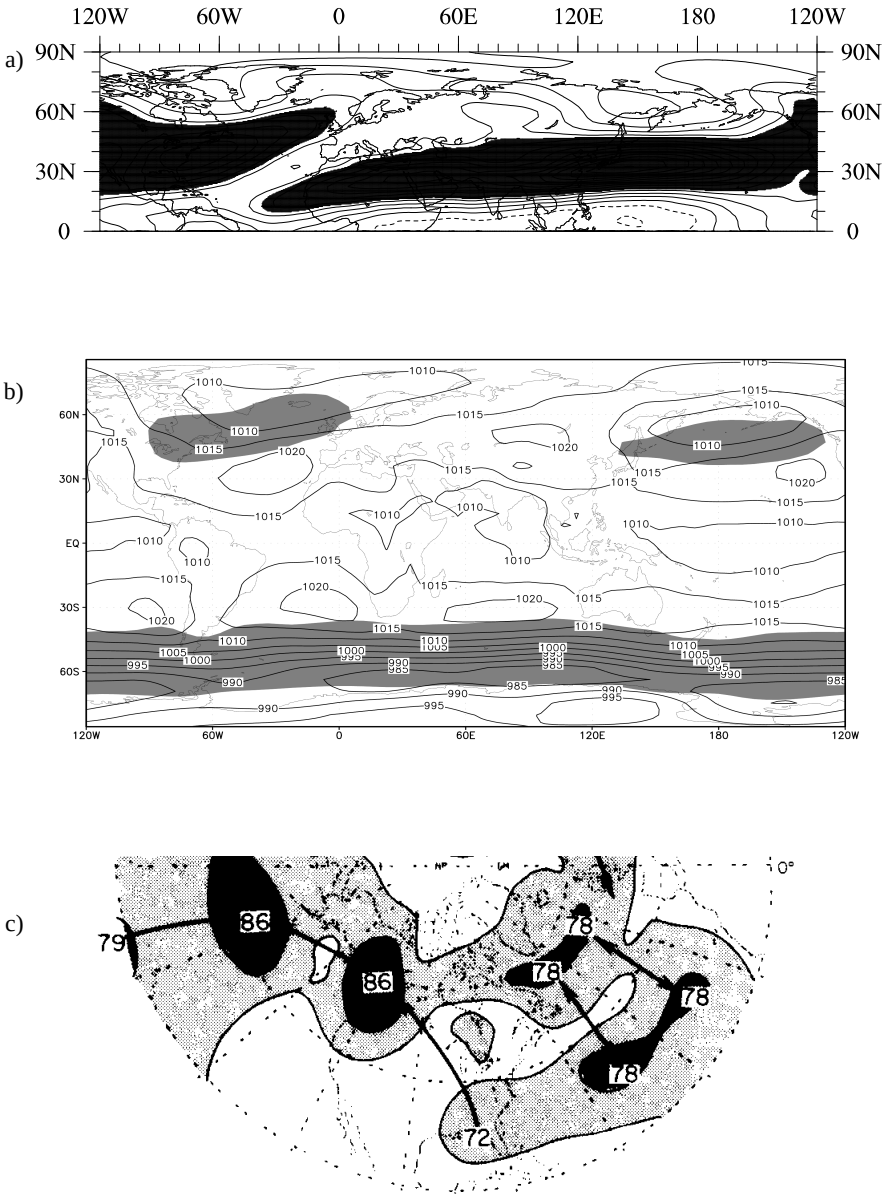


Fig. 3.4 - Jetstreams, stormtracks, and teleconnection patterns: (a) The mean 300hPa zonal wind in the northern hemisphere in winter (December to February, above $15ms^{-1}$ shaded), (b) the annual means of 3-6 day bandpass filtered standard deviation of the 500hPa geopotential (above 40gpm shaded) with the surface pressure (in hPa, thin lines), and (c) teleconnection patterns (after Gutzler and Wallace 1981): The Pacific-North America pattern (PNA, an east-west dipole embedded in an arc from the tropical Pacific to North America) and the North Atlantic Oscillation (NAO, a North-South dipole with centres near Iceland and Azores).

Polar Front Jet, which changes its meridional position in a region poleward of the weaker heating over the tropical Atlantic. The NAO variability covers a wide spectral band ranging from the relatively short life cycle of about ten days to a distinct low-frequency contribution with enhanced (active) and reduced (passive) phases (Fig. 3.3a and b, Appenzeller et al. 1998 [3]). These two aspects will be discussed in this and the following subsection using idealised atmospheric and coupled ocean-atmosphere GCM experiments.

Teleconnections observed (teleconnectivity, patterns, and indices): Teleconnection patterns describe centres of action which characterise the low-frequency variability of atmospheric fields between remote locations. Following Wallace and Gutzler 1981 [86] these patterns are deduced from the correlation matrix, whose elements are temporal correlation coefficients, r_{ij}, between monthly mean geopotential height fluctuations, ϕ_i', normalised by their local standard deviation at any selected grid point (subscript i) and every other gridpoint (subscript j): $r_{ij} = <\phi_i' \phi_j'>$. The centres of action inducing low-frequency variability are identified by the teleconnectivity, given by the (negative) correlation minimum of each column of the correlation matrix: $T_i = |r_{ij}$-minimum of all $j|$. Teleconnection patterns are defined by the columns of the correlation matrix which, spatially displayed, represent one-point correlation maps (Fig. 3.4c, 3.6). A teleconnection index is defined as the difference of normalised fluctuations between anti-correlated regions (averaged over several grid points). It is used to describe the space-time variability of a teleconnection pattern. A large set of teleconnection patterns has been identified in both observations and general circulation models (Fig. 3.4c): North Atlantic Oscillation (NAO), Pacific North America pattern (PNA), El Nino/Southern Oscillation (ENSO).

Idealized GCM experiments: Atmospheric eigenmodes similar to the observed NAO and PNA structure occur in experiments with an idealised atmospheric GCM (PUMA). This idealised atmospheric circulation is represented by a global spectral model solving the primitive equations on sigma levels (pressure/surface pressure). The diabatic processes are represented by Newtonian cooling with the model being driven to a reference temperature, $T_R - T$, within a relaxation time scale $\tau_R \sim 20 days$. Dissipation is formulated by Rayleigh friction. The model is used with T21 spectral truncation and 5 equally spaced sigma levels (PUMA, Portable University Model of the Atmosphere, Fraedrich et al. 1998). The idealised experiments are designed with and without the mid-latitude stormtracks and weak or strong tropical heating to explore the atmospheric mechanisms underlying low frequency variability (James et al. 1994 [47], Frisius et al. 1998 [37], Franzke et al. 2000 [35] and 2001 [36], Walter et al. 2000 [87]).

Jetstreams and eddies: Two classes of mechanisms contribute to the observed state of the atmosphere, diabatic and dissipative processes, and adiabatic nonlinear dynamics. The physical processes link the time mean flow to the external forcing (radiation, latent heat release) and the dynamics modify this basic state by its time-mean effects, which are due to divergence or convergence of momentum- and temperature transports by transient eddies. The interaction of both physical and dynamical processes and the sensitivity of the basic state on both mechanisms is documented by four PUMA simulations. A zonally and hemispherically symmetric T_R-field with a 90 K equator-pole difference provides the diabatic forcing. With or without an additional diabatic heat source at the equator (1K/day), one can simulate the situations of stronger or weaker release of latent heat over the tropical Pacific and Atlantic. The effect of ed-

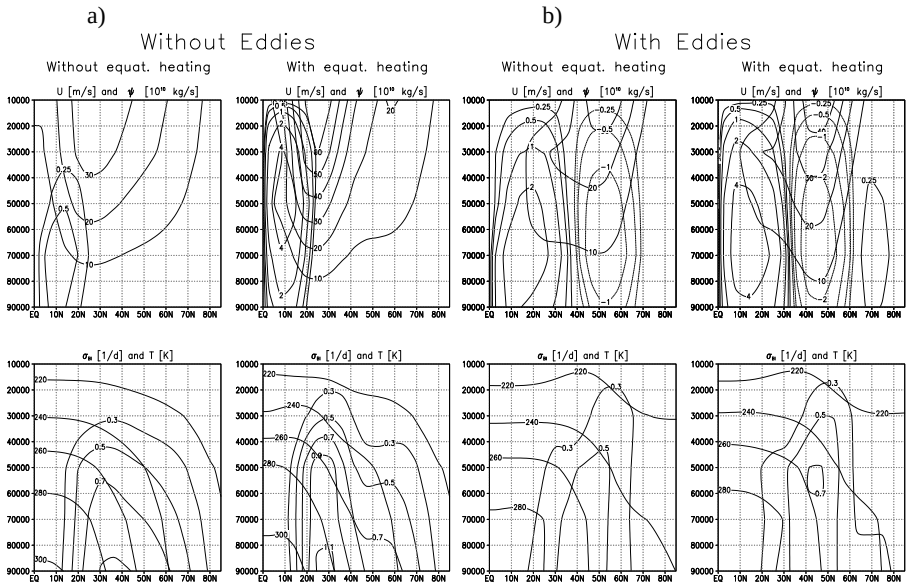


Fig. 3.5 - Meridional cross-sections of the zonal mean zonal wind (thin line) and the mass streamfunction (heavy line, top panels), the temperature (thin line) and the baroclinic growth rate (heavy line, bottom panels) of PUMA simulations with and without additional equatorial heating (left and right column; height in Pascal): (a) without transient eddies and (b) with transient eddies.

dies on these two basic states is analysed by integrations with and without initialising transient eddies. Vertical-meridional cross-sections of four zonally averaged parameters describe the time-mean flow responses: Zonal wind, mass-stream function, temperature and baroclinicity or baroclinic growth (Eady 1949 [11]). Without transient disturbances (Fig. 3.5a), an additional equatorial heat source enhances the Hadley cell and the subtropical jet. Baroclinicity is also enhanced and concentrates near the strengthened subtropical jet; a secondary maximum appears in the mid-latitudes (50° to 60°). With transient eddies (Fig. 3.5b) the circulation changes significantly: The eddy driven jet of the mid-latitudes dominates the zonal wind field in both cases of weak and strong tropical heating with the subtropical jet at the poleward branch of the Hadley cell being reduced to a secondary minimum. Thus the zone of largest baroclinicity is shifted polewards and weakens. That is, the transient eddies affect the meridional circulation which, by building the Ferrel cell, strengthens the Hadley cell.

Teleconnections and stormtrack phenomenology: An almost realistic single stormtrack is generated by superimposing a zonally asymmetric contribution to the symmetric global reference temperature field, which consists of an idealised dipole representing a cold continent located northwestward of a warm ocean. This T_R-dipole anomaly decreases with height and vanishes at the tropopause. The teleconnection pattern of the single stormtrack experiment and its fluctuation are characterised by a retrograde travelling Rossby wave number two (Frisius et al. 1998 [37]). Two-stormtrack experiments (Franzke et al. 2001 [36]) are performed introducing two T_R-dipoles at varying zonal distance with one dipole being shifted from 180° to 130°. To further

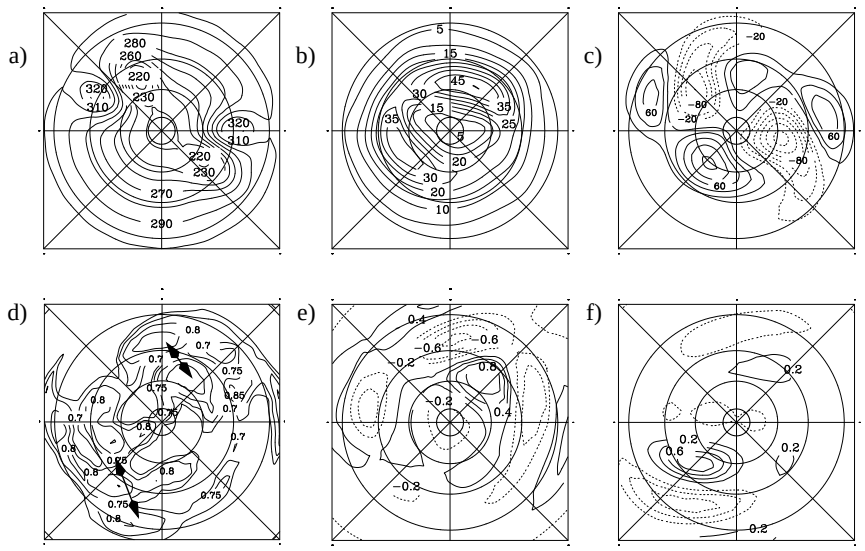


Fig. 3.6 - Idealized two stormtrack experiment of $150°$ separation: (a) Restoration temperature at 900hPa (contour interval is 10K), (b) stormtrack or standard deviation of band-pass filtered 500hPa geopotential (high-frequency transient eddies, contour interval is 5pgm), (c) long-term mean of 300hPa geopotential minus zonal mean (stationary eddies, contour interval 20 gpm), (d) teleconnectivity (contour interval is 0.05), (e) and (f) local correlations of monthly mean 300hPa geopotential (contour is 0.2, negative values are dashed, zero line is not displayed) for base point in P-area ($56.3°$E, $47.1°$N) and A-area ($230.6°$E, $47.1°$N).

compare idealised experiments with the observed stormtracks and teleconnections in the North Atlantic and North Pacific, the area downstream of the shifted TR-dipole is denoted as A-area, the other one as P-area; the respective A- and P-teleconnection indices and their patterns are deduced: (i) Separations of 200° to 160° between the two stormtrack generating T_R-dipoles create teleconnection patterns which, not unlike the single stormtrack case, describe a retrograde travelling Rossby wave number two (Franzke et al. 2000 [35]) with a strong correlation between the two teleconnection indices. The low frequency variability and exitation of the Rossby wave is due to the size and location of the stormtracks, which induce amplification (damping) of the anti-cyclonic (cyclonic) part of the wave (see below). (ii) Separations of 140° to 150° between the two stormtrack generating T_R-dipoles yield a more realistic climatology of the location of the stormtracks and the asymmetric time mean flow (stationary waves, Fig. 3.6a-c) where the A-area and P-area lie in the lower and upper half of each graph. Two teleconnections coexist whose teleconnectivities and local one-point correlations (A- and P-pattern) are, in contrast to the 180° experiment, unrelated and similar to the observed NAO and the PNA pattern (see Wallace and Gutzler 1981 [86], Fig. 3.6d-f). Both patterns undergo a low frequency life cycle of about 20 days (from -10 to +10 day lag, Fig. 3.7a-c for the A-pattern only), which is also comparable with the observed features. Life cycle composites show the following results: The positive A-index is associated with a retrograde travelling barotropic Rossby wave, which originates in the center of the A-storm track (-20 day lag) and propagates upstream. When the high pressure area of this wave passes the storm track, its amplitude am-

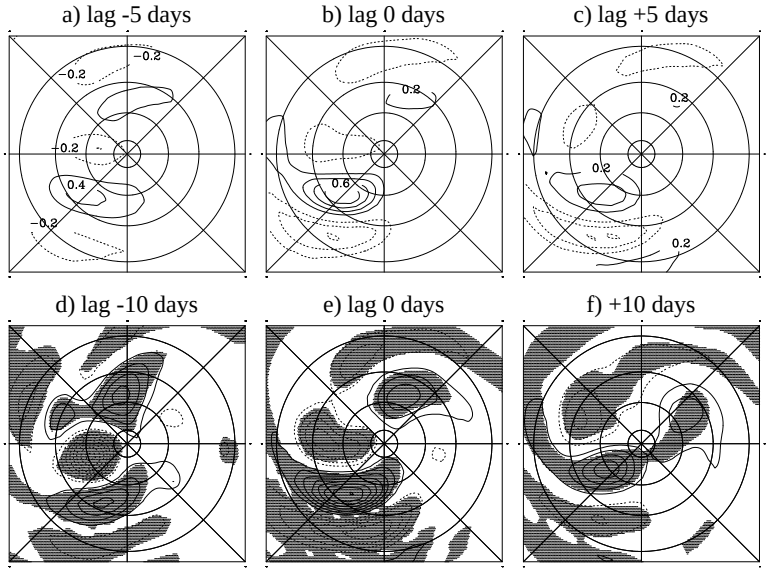


Fig. 3.7 - Idealized two stormtrack experiment of $150°$ separation (A-area): Lag correlation patterns of 300 hPa geopotential height are based on low pass filtered daily data of (a) -5 days, (b) 0 days and (c) +5 days. The base point is at $230.6°$E, $47.1°$N in the A-area, which is located in the lower part of all panels (contour interval is 0.2, the zero contour is not displayed). Composites of the vertical mean streamfunction anomaly for positive A-index events at d) day -10, e) day 0, f) day 10. The contour interval is $5 \cdot 10^5 m^2 s^{-1}$, shading indicates significance above the 99%-level.

plifies (until zero-lag) and decays after zero-lag (Fig. 3.7d-f). At zero-lag the high pressure area of the barotropic wave composite coincides with the positive area of the one point correlation map and with the high pressure anomaly of the stationary wave; the center of the low pressure area (further downstream) does not coincide with the negative area in the one point correlation map. The negative phase of the index is similar to the positive phase but of opposite signs (not shown).

Mechanisms (location, life-cycle, and spatial resonance): From the phenomenological analysis we turn to the physical mechanisms, which generate and support the teleconnection pattern. (i) As the positive area of the one-point correlation map coincides with the high pressure area of the stationary eddies, the asymmetric time mean flow appears to be responsible for the geographical location of the low frequency variability (see Branstator 1990 [8], Kang 1990 [49]). It is shown that the responses of forced Rossby waves depend strongly on the mean flow structure. This is supported by the barotropic streamfunction tendency with eight terms of linear and nonlinear interactions (see Franzke et al. 2001 [36]), which is applied to the composite life-cycle (shown in Fig. 3.7d-f). (ii) The A-pattern life cycle or low-frequency variability is exited (as early as -20 day lag) by the zonally asymmetric time mean flow (stationary wave, Fig. 3.6c), which interacts with the retrograde travelling barotropic Rossby wave number two. Note that its amplitude amplifies, when its high pressure area passes the storm track, which characterises spatial resonance (see iii). The decay

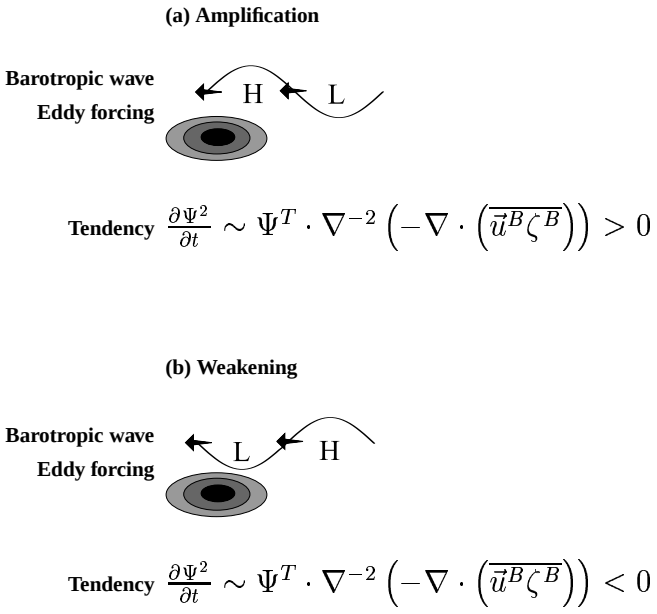


Fig. 3.8 - Mechanism of 'spatial resonance' leading to ridge amplification (H) and trough weakening (L): Both ridge (a) and trough (b) of the eastward traveling Rossby wave (arrows) generate positive (a) and negative (b) streamfunction tendencies, $\frac{\partial \Psi^2}{\partial t}$, when projected onto the positive high frequency eddy forcing, $\nabla^{-2} \left(-\nabla \cdot \left(\overline{\vec{u}^B \zeta^B} \right) \right)$, (shaded areas). The observed phase shift between high frequency eddy forcing (shaded areas) and thephase of the Ψ-wave is also indicated.

of the anomaly (after zero lag) commences with the end of the exitation phase after passage of the Rossby wave. The streamfunction tendency of this decay-part of the life cycle is forced by the interaction between the zonal mean time mean flow with the low frequency flow, and the low frequency contribution to the divergence term. (iii) The 'spatial resonance' mechanism characterises the storm track induced amplification (decay) of the anti-cyclonic (cyclonic) part of the wave. It can be formally derived by the low frequency barotropic streamfunction or vorticity tendency, whose dominating term is the divergence of the high frequency vorticity flux (Franzke et al. 2000 [35]) (Fig. 3.8a, b). In this sense the high frequency eddies create low frequency variability. The same mechanism occurs also in ocean circulations, which are driven by a spatially inhomogeneous stormtrack (Sura et al. 2000 [74]).

Summarising: This section is an introduction to long-term memory and the mechanisms, which may etablish low frequency variability in the climate system. The methods of analysis introduced are data analysis techniques, idealised circulation model experiments, and the development of conceptual models: (i) Long-term memory is documented by detrended fluctuation analysis applied to observed and simulated data. A toy diffusion model with random forcing and Newtonian relaxation is presented, which shows qualitatively similar behaviour. (ii) The dynamics of teleconnections and their life cycles is analysed; they are part of the atmosphere's low frequency vari-

ability. Spatial resonance appears as a suitable mechanism to describe the life-cycle of teleconnection patterns and to generate low frequency variability in idealised models of the atmospheric circulation (Franzke et al. 2000, 2001 [35][36]). A similar mechanism has been discovered for a shallow water ocean forced by the atmosphere (Sura et al. 2000 [74]). (iii) Finally (not discussed), noise induced regime transitions may be added as another mechanism producing low-frequency variability. A shallow water ocean forced by spatially inhomogeneous random windstress shows such behaviour (Sura et al. 2001 [75]). Although the subject of long-term memory and the associated processes is far from being completely described, the next section, which serves as an outlook, extends the predictability issue into the new field of Global Change.

IV. GLOBAL CHANGE PREDICTABILITY: SEVEN GLOBAL RISKS

Lectures on predictability of weather and climate would be incomplete without an outlook to some future issues. One issue is related Global Change and a scientifically based guidance for political initiatives, decisions, and actions. Not unlike the information provided for clients of weather forecasts (should she take her umbrella or leave it at home), Global Change politics expects and requires recommendations. In both cases, the diagnosed past and present join the prediction of future states and lead to recommendations respecting normative-ethical Leitmotifs (Tab. 4.1). Working tools for Global Change advice are developed and extensively discussed by the German Advisory Council of Global Change (WBGU) in their reports to the government (http://www.wbgu.de; see also Fraedrich 2000 [29]).

Tab. 4.1 - Global Change: Diagnosis, prognosis, and recommendations

Diagnosis	Symptoms: trends of global environmental changes Syndromes: positive feedbacks between symptoms Criticality: loading of natural systems, complex indicators
Prognosis	Risks, risk classes and associated strategies Risk dynamics
Recommendation	Normative-ethical Leitmotif Qualitative evaluation of diagnosis and prognosis Quantitative evaluation: guard rails and tolerance windows

The predictability issue may suffice here as a first introduction. It is associated with the risk concept, by which the diagnostics of damage and its circumstances turns into a prognosis provided by an estimate of the probabilities of its occurrence. Thus damage versus probability of occurrence span a two-dimensional phase space of risks, in which three regions are identified: Normal, transition and prohibited areas (Fig. 4.1). Adding confidence intervals to the damage and the probability estimates enlarges the two-dimensional phase space of Global Change risks, whose dimension can be further increased by including information on their spatial extent, their duration and type of evolution, that is the ubiquity, persistence and irreversibility of the global risks. In this effectively higher dimensional phase space risk classes are formed, which may contain risks of different origin. Seven global risk classes have been identified which, except for the unknown risk, are coined after the Greek mythology: *Cyclops*, *Cassandra*, *Medusa*, *Pandora*, *Damocles*, *Phythia* (Beese et al. 1998 [90]).

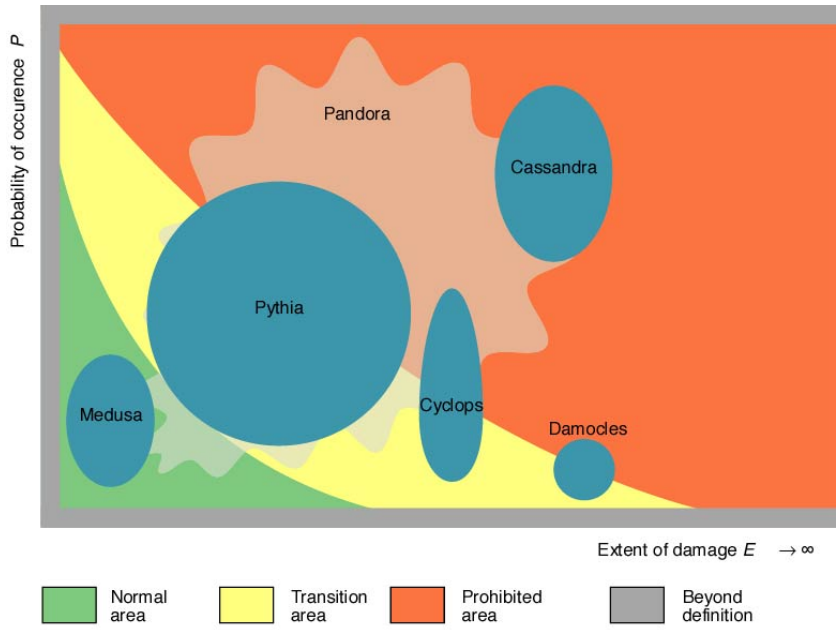


Fig. 4.1 - Risk Classes of Global Change: Their location in the normal, transition and prohibited regions of the (damage, probability)-space (after WBGU 1998 [90]).

Action of society may move a risk from one region to another, even by perception. A concerted action, however, is required, if risks are to be reduced and shifted to the normal region. Therefore, each risk class requires its own reduction strategy, which is related to its location in phase space. For example: Cyclops are mighty giants who, for all their strength but being one-eyed, can only perceive one side of reality. Thus, Cyclope class risks are characterised by a probability of occurrence, which is highly uncertain but, if damage occurs, it is very large. Many natural events like floods and earthquakes fall in this class, and the occurrence of AIDS. The high uncertainty of occurrence requires three strategies: Ascertaining the probability, prevent surprises, and manage emergency. - This outlook on predictability and forecasting in Global Change and this specific example may suffice as a snapshot of a new research area linking the social, economic, and natural science communities.

Acknowledgement: These notes comprise, more or less linearly combined, lectures and talks prepared for 2001 (based on published und still unpublished material): 'Nonlinear Geophysics - Special Session' American Geophysical Union (AGU, Boston, June 1-2), 'Second Workshop on Stochastic Climate Models' (Chorin, July 9-11), 'Climate Conference 2001' (Utrecht, August 20-24), 'Chaos in Geophysical Flows' (ISSAOS, L'Aquila, Italy, September 10-14), School of Earth Sciences (Pennsylvania State University, October 11). The support by Frank Sielmann, Richard Blender, Frank Lunkeit, Ute Luksch, and by the patient audiences is greatfully acknowledged.

[1] Abarbanel, H.D.I., R. Brown, J.H. Sidorowich, and L.S. Tsimring 1993: The analysis of observed chaotic data in physical systems. *Rev. Mod. Phys.*, 65, 1331-1392.
[2] Alexander, M.A. and C. Deser 1995: A mechanism for the recurrence of wintertimemidlatitude SST anomalies. *J. Phys. Oceanogr.*, 25, 122-137.
[3] Appenzeller, C., T.F. Stocker, and M. Anklin 1998: North Atlantic dynamics recorded in Greenland ice cores. *Science*, 282, 446-448.
[4] Barahona, M. and C.-S. Poon 1996: Detection of nonlinear dynamics in short, noisy time series. *Nature*, 381, 215-217.
[5] Benzi, R., G. Jona-Lasinio, and A. Sutera 1989: Stochastically perturbed Landau-Ginzburg equation. *J. Stat. Phys.*, 55, 505-522.
[6] Benzi, R., G. Parisi, A. Sutera, and A. Vulpiani 1982: Stochastic resonance in climatic change. *Tellus*, 34, 10-16.
[7] Bjerknes, J. 1964: Atlantic air-sea interaction. *Adv. Geophys.*, 10, 1-82.
[8] Branstator, G., 1990: Low-frequency patterns induced by stationary waves. *J. Atmos. Sci.*, 47, 629-648.
[9] Charney, J.G., R. Fjortoft, and J. von Neumann 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, 2, 237-254.
[10] Cooke, W.E. 1906: Forecasts and verifications in Western Australia. Mon. Wea. Rev., 34, 23-24.
[11] Eady, E. 1949: Long waves and cyclone waves. *Tellus*, 1, 33-52.
[12] Feldstein, S. B. 1998: The growth and decay of low-frequency anomalies in a GCM. *J. Atmos. Sci.*, 55, 415-428.
[13] Fraedrich, K., 1986: Estimating the dimensions of weather and climate attractors. *J. Atmos. Sci.*, 43, 419-432.
[14] Fraedrich, K., 1987: Estimating weather and climate predictability on attractors. *J. Atmos. Sci.*, 44, 722-728.
[15] Fraedrich, K. 1988: El Nino/Southern Oscillation predictability. *Mon. Wea. Rev.*, 116, 1001-1012.
[16] Fraedrich, K. and L.M. Leslie, 1991: Predictability studies of the Antarctic atmosphere using both analogue and chaos theory. *Austral. Meteorol. Mag.*, 39, 1-9.
[17] Fraedrich, K. and K. Müller, 1992: Climate anomalies in Europe associated with ENSO extremes. *Intern. J. Climatol.*, 12, 25-31.
[18] Fraedrich, K. and R. Wang 1993: Estimating the correlation dimension of an attractor from noisy and small datasets based on re-embedding. *Physica D*, 65, 373-39.
[19] Fraedrich, K. and L. M. Leslie 1987: Evaluation of techniques for the operational, single station, short term forecasting of rainfall at a mid-latitude station (Melbourne). *Mon. Weather Rev.*, 115, 1645-1654.
[20] Fraedrich, K., C. Bantzer, and U. Burkhardt 1993: Winter climate anomalies in Europe and their associated circulation. *Clim. Dyn., 8*, 161-175.
[21] Fraedrich, K. 1994: ENSO Impact on Europe? A Review. *Tellus*, 46A, 541-552.
[22] Fraedrich, K. and C. Ziehmann 1994: Predictability experiments of persistence forecasts in a red noise atmosphere. *Q. J. R. Meteorol. Soc.*, 120, 387-428.
[23] Fraedrich, K., Ziehmann, C., and F. Sielmann 1995: Estimates of spatial degrees of freedom. *J. Climate*, 8, 361-369.
[24] Fraedrich, K. and C. Ziehmann 1995: Praktische Vorhersagbarkeit: Persistenz inrotem Rauschen. Meteorol. Zeitschrift, NF4, 139-149.
[25] Fraedrich, K. 1996: Das Lorenz-Modell - ein Paradigma für Wetter und Vorhersagbarkeit. *Promet*, 25 (Heft 3), 62-79.
[26] Fraedrich, K. 1997: Atmospheric variability: Modelling, diagnostics, and forecasting. In Past and Present Variability of the Solar-Terrestrial System: Measurements, Data Analysis and Theoretical Models (G. Cini Castagnoli and A. Provenzale, eds.). International School of Physics 'Enrico Fermi' Course, 133, 431-483.
[27] Fraedrich, K. and B. Rueckert 1998: Metric adaption for analog forecasting. *Physica A*, 253, 379-393.
[28] Fraedrich, K., E. Kirk, and F. Lunkeit 1998: Portable University Model of the Atmosphere. Technical Report 16, Deutsches Klimarechenzentrum (available http://www.dkrz.de/forschung/reports.html), 37pp.

[29] Fraedrich, K., 2000: Advising Global Change Politics: A Model (Politikberatung zum globalen Wandel: Ein deutsches Modell), Festschrift Albrecht Kessler, Universität Freiburg (ISSN 1435-618X), 9-34.
[30] Fraedrich, K. 2001: Simple Climate Models. *In Progress in Probability*, 49, (eds. P. Imkeller and J.-S. von Storch), Birkhäuser Verlag, 65-100.
[31] Fraedrich, K. and C.-D. Schoenwiese 2002: Space-time Variability of the European Climate. In The Science of Disasters (eds. A. Bunde, J. Kropp, H.J. Schellnhuber), Springer Verlag, 104-139.
[32] Fraedrich, K., C.C. Raible, and F. Sielmann 2003: Analog ensemble forecasts of tropical cyclone tracks in the Australian region. *Wea. Forecast.*, 18, 3-11.
[33] Fraedrich, K. 2002: Fickian diffusion and Newtonian cooling: A concept for noise induced climate variability with long-term memory? *Statistics and Dynamics*, 2, 403-412.
[34] Fraedrich, K. and R. Blender 2003: Scaling of atmosphere and ocean temperature correlations in observations and climate models. *Phys. Rev. Lett.*, 90, 108501-(1-4).
[35] Franzke, C., K. Fraedrich, and F. Lunkeit 2000: Low frequency variability in a simplified atmospheric global circulation model: Storm track induced 'spatial resonance'. *Q. J. R. Meteorol. Soc.*, 126, 2691-2708.
[36] Franzke, C., K. Fraedrich, and F. Lunkeit 2001: Teleconnections and low frequency variability in idealized experiments with two storm tracks. *Q. J. R. Meteorol. Soc.*, 127, 1321-1340.
[37] Frisius, T., F. Lunkeit, K. Fraedrich, and I.N. James 1998: Storm-track organization and variability in a simplified atmospheric global circulation model. *Q. J. R. Meteorol. Soc.*, 124, 1019-1043.
[38] Grassberger, P. and I. Procaccia 1983: Measuring the strangeness of strange attractors. *Physica D*, 9, 189-208.
[39] Grassberger, P. and I. Procaccia 1984: Dimensions and entropies of strange attractors from a fluctuating dynamics approach. *Physica D*, 13, 34-54.
[40] Götz, G. 1995: Predictability of non-linear dynamical systems. *Időjárás (Quart. J. Hungarian Meteor. Service)*, 99, 1-32.
[41] Green, J.S.A. 1970: Transfer properties of the large-scale eddies and the general circulation of the atmosphere. *Quart. J. R. Met. Soc.*, 96, 157-185.
[42] Hasselmann, K. 1976: Stochastic climate models. *Tellus*, 28, 473-485.
[43] Heneghan, C. and G. McDarby 2000: Establishing the relation between detrended fluctuation analysis and power spectral density analysis for stochastic processes. *Phys. Rev. E*, 62, 6103-6110.
[44] Hosking, J.R.M. 1981: Fractional differencing. *Biometrica*, 68, 165-176.
[45] Hoskins, B.J. and D. Karoly 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, 38, 1179-1196.
[46] Hurrell, J.W. 1995: Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science*, 269, 676-679.
[47] James, P. M., K. Fraedrich, and I. N. James 1994: Wave zonal flow interaction and ultra low frequency variability in a simplified global circulation model. *Q. J. R. Meteorol. Soc.*, 120, 1045-1067.
[48] Kalnay, E. and R. Livezey, 1983: Weather predictability beyond a week: An introductory review. From: Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics. International School of Physics 'Enrico Fermi' Course 88, 311-346.
[49] Kang, In-Sik, 1990: Influence of zonal mean flow change on stationary wave fluctuations, *J. Atmos. Sci.*, 47, 141-147.
[50] Koscielny-Bunde, E., A. Bunde, S. Havlin, H.E. Roman, Y. Goldreich, and H.J. Schellnhuber 1998: Indication of a universal persistence law governing atmospheric variability. *Phys. Rev. Lett.*, 81, 729-732.
[51] Lorenz, E.N. 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20, 130-141.
[52] Lorenz, E.N. 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, 26, 636-646.
[53] Lorenz, E.N. 1991: Dimension of weather and climate attractors. *Nature*, 353, 241-244.
[54] Luksch, U. 1996: Simulation of North Atlantic low-frequency SST variability. *J.Climate*, 9, 2083-2092.
[55] Luterbacher, J., C. Schmutz, D. Gyalistras, E. Xopolaki, and H. Wanner 1999: Reconstruction of monthly NAO and EU indices back to A.D. 1675. *Geophys. Res. Lett.*, 26, 2745-2748.
[56] Manabe, S. and R.J. Stouffer 1996: Low-frequency variability of surface air temperature in a 1000 year integration of a coupled atmosphere-ocean-land surface model. *J. Climate*, 9, 376-393.
[57] Megreditchian, G. der 1990: Meteorological networks optimization from a statistical point of view. *Computational Statistics and Data Analysis*, 9, 57-75.
[58] Monin, A.S. and A.M. Yaglom 1975: *Statistical Fluid Mechanics*, Vol.2. MIT Press, 869pp.

[59] Müller W., R. Blender, and K. Fraedrich 2002: Low-frequency variability in idealised GCM experiments with circumpolar and localised stormtracks. *Nonlin. Proc. Geophys.*, 9, 37-49.
[60] North, G.R., T.L. Bell, R.F. Cahalan, and F.J. Moeng 1982: Sampling errors in the estimation of empirical orthonogal functions. *Mon. Wea. Rev.*, 110, 699-706.
[61] Patil, D.J., D.R. Hunt, E. Kalnay, J.A. Yorke, and E. Ott 2001: Local low dimensionality of atmospheric dynamics. *Phys. Rev. Let.*, 86, 5878-5881.
[62] Pelletier, J. 1999: Analysis and modelling of the natural variability of climate. *J. Climate*, 10, 1331-1342.
[63] Pelletier, J. and D. Turcotte 1999: Self-affine time series II: Applications and models. *Adv. in Geophys.*, 40, 91-166.
[64] Peng, C.-K., S. V. Buldyrev, A. L. Goldberger, S. Havlin, M. Simons, and H. E. Stanley 1994: *Phys. Rev. E*, 51, 1685.
[65] Popper, K.R. 1972: Of Clocks and Clouds. From: Objective Knowledge. Oxford, At the Clarendon Press, 206-255.
[66] Raible, C.C., U. Luksch, K. Fraedrich, and R. Voss 2001: North Atlantic decadal regimes in a coupled GCM simulation. *Clim. Dyn.*, 17, 321-330.
[67] Richardson, L.F. 1922: Weather Prediction by Numerical Process. Cambridge University Press, pp. 236.
[68] Roeckner, E., K. Arpe, and L. Bengtson 1996: The atmospheric general circulation model ECHAM-4: Model description and simulation of present day climate. Technical report 218, Max Planck Institut für Meteorologie, Hamburg, Germany.
[69] Saltzman, B. 1962: Finite amplitude convection as an initial value problem. *J. Atmos. Sci.*, 19, 329-341.
[70] Shirer, H.N., C.J. Fosmire, R. Wells, and L. Suciu 1995: Estimating the correlation dimension of atmospheric time series. *J. Atmos. Sci.*, 54, 211-229.
[71] Sievers O., K. Fraedrich, and C. Raible 2000: Self-adapting analogt ensemble predictions of tropical cyclone tracks. *Wea. Forecast.*, 15, 623-628.
[72] Sparrow, G. 1982: The Lorenz Equations. Springer-Verlag, USA.
[73] Smith, L.A., C. Ziehmann, and K. Fraedrich 1999: Uncertainty dynamics and predictability in chaotic systems. *Q. J. R. Meteorol. Soc.*, 125, 2855-2886.
[74] Sura, P., F. Lunkeit, and K. Fraedrich 2000: Decadal variability in a simplified wind-driven ocean model. *J. Phys. Oceanogr.*, 30, 1917-1930.
[75] Sura, P., K. Fraedrich, and F. Lunkeit 2001: Regime transitions in a stochastically forced double-gyre model. *J. Phys. Oceanogr.*, 30, 411-426.
[76] Sutton, R.T. and M.R. Allen 1997: Decadal predictability of the North Atlantic sea surface temperature and climate. *Nature*, 388, 563-567.
[77] Stephenson, D.B., V. Pavan, and R. Bojariu 2000: Is the North Atlantic a random walk? *Intern. J. Climate*, 20, 1-18.
[78] Talkner, P. and R.O. Weber 2000: Power spectrum and detrended fluctuation analysis: Application to daily temperatures. *Phys. Rev. E*, 62, 150-160.
[79] Tennekes, H. 1991: Karl Popper and the accountability of numerical forecasting. In New Developments in Predictability, ECMWF Workshop proceedings, Reading, UK.
[80] Toth, Z. 1995: Degrees of freedom in Northern Hemisphere circulation data. *Tellus*, 47A, 457-472.
[81] Torrence, C. and G.P. Compo 1998: A practical guide to wavelet analysis. *Bull. Amer. Met. Soc.*, 79, 61-68.
[82] Tsonis, A.A. and Elsner, J.B. 1989: Chaos, strange attractors and weather. *Bull. Amer. Met. Soc.*, 70, 14-23.
[83] van den Dool, H.M. 1994: Searching for analogs: How long must we wait? *Tellus*, 46A, 314-324.
[84] Voss, R.F. and J. Clarke 1976: Flicker (1/f) noise: Equilibrium temperature and resistance fluctuations. *Phys. Rev. B*, 13, 556-573.
[85] van Vliet, K.M., A. van der Ziel, and R.R. Schmidt 1980: Temperature-fluctuation noise on thin films supported by a substrate. *J. Appl. Phys.*, 51, 2947-2956.
[86] Wallace, J. M. and D.S. Gutzler 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, 109, 785-812.
[87] Walter, K., U. Luksch, and K. Fraedrich 2001: A response climatology of idealized SST anomaly experiments with and without storm track. *J. Climate*, 13, 467-484.
[88] Walter K. and H. F. Graf 2002: On the stability of the regional connection between the North Atlantic Oscillation and sea surface temperature. *J. Geoph. Res.*, 107, D17, 4338.
[89] Wang X. and S.S. Shen 1999: Estimation of spatial degrees of freedom of a climate field. *J. Climate*,

12, 1280-1291.

[90] WBGU: Beese F. O., K. Fraedrich, P. Klemmer, J. Kokott, L. Kruse-Graumann, C. Neumann, O. Renn, H.-J. Schellnhuber, E.-D. Schulze, M. Tilzer, P. Velsinger, and H. Zimmermann 1999: World in Transition: Strategies for Managing Global Environmental Risks. German Advisory Council of Global Change (WBGU), Annual Report 1998. Berlin, Heidelberg, New-York: Springer Verlag, 383pp.

[91] Wolf, A., J. Swift, H.L. Swinney, and J.A. Vasrano 1985: Determining Lyapunov exponents from time series. *Physica D*, 16, 285-317.

[92] Ziehmann, C., K. Fraedrich, and L.A. Smith 1995: Ein internes Vorhersagbarkeitsexperiment im Lorenzmodell. *Meteorol. Zeitschrift*, NF4, 16-21.