Chapter 11
Statistics – an Indispensable Tool in Dynamical Modeling

by Hans von Storch

Abstract

The role of statistical analysis in the process of establishing and utilising ocean and other environmental models is discussed. A general state space model approach is adopted. In "quasi-realistic models," statistical thinking is encoded in the parameterisations and is required for extracting experimental evidence and for validation. Data assimilation techniques are used to systematically combine observational evidence and quasi-realistic models. While quasi-realistic models serve as complex substitute reality, is dynamical knowledge represented through simplified models. These "cognitive" idealised models have to be fitted to observational data when adapted to real situations.

11.1 Environmental Research

As outlined in a lecture of this school (von Storch 2000), two fundamentally different types of mathematical models are used in environmental research. One sort is "quasi-realistic," and is supposed to be a substitute reality, within which the otherwise impossible experiments can be conducted. A few decades earlier, such models were often mechanistic, but most of these apparata have now been replaced by mathematical models (Sündermann and Vollmers 1972). They are also used to extra- and interpolate in a dynamically consistent manner the sparse observations, so that spatially and temporally high resolution analyses of the system's state are constructed. A representative of this type are 3-D models of the North Sea with a resolution of a few tens of kilometers, simulating explicitly an array of processes such as advection, mixing, tides, bottom stress, wind stress, air-sea interaction and the like (e.g., Kauker 1999). The other type of model, named here "cognitive," is highly simplified and idealised. *Because of its reduced complexity,* such a model constitutes "knowledge." An example of this type of model is Frankignoul's model of the variability of the mixed layer depth (Frankignoul 1979). Both types of modeling require the use of statistical thinking, both in terms of design as well as analysis.

The present paper is a discussion about the different roles of statistics in designing and using these models. Throughout this discussion, we make use of the formalism of *state space models* (Sect. 11.2). In Sect. 11.3, the role of statistics in quasi-realistic modeling is considered: parameterisations, forecasting, simulation and numerical experimentation, and data analysis. Several examples, ranging from the specification of high frequency wind fluctuations used in wave modeling, hindcasting storm surge statistics over 40 years, and simulating the impact of increased atmospheric carbon dioxide concentrations on storm surge statistics are presented. In Sect. 11.3 the role of statistics for "cognitive" models is discussed; first the general concept of *Principal Interaction Patterns* is introduced, and an example of a *Principal Oscillation Pattern* Model of wave dynamics along the Pacific equator is presented.

11.2 State Space Models

Here, we introduce as a kind of overarching general view the concept of *state space models*.

We describe our system with a state variable, represented by an *m*-dimensional state vector ϕ_t. Often, this state vector can not be observed, sometimes because of lack of suitable sensors, but also because space-time continuous observations are not doable. The dynamics of this state variable are described by a system of difference equations with the dynamics ***F***.

$$\phi_{t+1} = F\left(\phi_t, \vec{\alpha}, t\right) + \varepsilon \tag{11.1}$$

The dynamics depend on a set of free parameters $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots)$; the fact that the dynamics are only approximately known as well as the fact that seemingly random effects act upon our system is taken care of by ε. Of course, the dynamics may generate internal noise as well. The dynamics ***F*** may be derived from theoretical arguments, such as the conservation of momentum or mass, or after an empirical fit.

Even if ϕ_t is not observable in its entirety, some empirical evidence will be available, for instance at some locations and at some times, or as indirect evidence from proxy information such as lake waves. If these observations are combined into an observation vector ω_t , we have an *observation equation* that relates the state variable to the observed variables

$$\omega_t = P(\phi_t) + \delta \tag{11.2}$$

Again, the observation equation is not exactly satisfied; there may be measurement uncertainties with respect to the value or to location and timing; also the link may be a bit fuzzy as in case of proxy data. The operator P is usually not invertible.

11.3 Statistics and Quasi-realistic Models

In this section we deal with quasi-realistic models, i.e., numerical models that incorporate all processes of first, second and sometimes third order. The level of complexity of such models is usually limited by the available computer power.

We discuss the problem of parameterisations, of analysing the output from quasi-realistic models and the problem of how to estimate the trajectory of an open system with the help of a quasi-realistic model and observational data.

11.3.1 Parameterisations

In the design of quasi-realistic models, such as GCM-type global or regional ocean models, the unavoidable truncation of the basic dynamical equations, necessitates the introduction of parameterisations of sub-grid scale processes (cf. von Storch et al. 1999). Such processes, such as the turbulent layers of the ocean at the surface and at the bottom, take place on scales too small to be resolved, but have a significant impact on the dynamics of the state variables such as the stream function. Therefore their *effect* on the resolved scales is considered to be partially determined by the configuration given at grid size scale (von Storch 1997). In *the state space formalism*, the fitting of parameterisations means to specify some of the unknown parameters α.

To allow for formalisation, let us write the state variable ϕ as a sum of the large-scale resolved component $\overline{\phi}$ and an unresolved part ϕ'

$$\phi = \overline{\phi} + \phi' \tag{11.3}$$

Then, our basic differential equations

$$\frac{\partial\phi}{\partial t} = F(\phi) \tag{11.4}$$

with the "dynamics" $\boldsymbol{F}$ is replaced by

$$\frac{\partial\overline{\phi}}{\partial t} = F_{\Delta x}(\phi) \tag{11.5}$$

with a modified operator $F_{\Delta x}$ resulting from the full operator $\boldsymbol{F}$ after introducing a truncated spatial resolution Δx. In general, this operator may be written as

$$F_{\Delta x}(\phi) = F(\overline{\phi}) + F'(\phi') \tag{11.6}$$

with an operator $\boldsymbol{F'}$ describing the net effect of the sub grid scale variations represented by ϕ'. With this set-up, the system (5) is no longer closed and can therefore no longer be integrated. For overcoming this problem, conventional approaches assume that the "nuisance" term $F'(\phi')$ is either irrelevant, i.e.,

$$F'(\phi') = 0 \tag{11.7}$$

or may be parameterised by

$$F'(\phi') = Q(\overline{\phi}) \tag{11.8}$$

with a dynamically motivated function Q. For the specification of this function, usually a number of parameters are to be determined. This is mostly done by statistically fitting Eq. 11.8 to data from observational campaigns.

While both specifications (11.7 – 8) return an integrable Eq. 11.5, they both have to assume that the local scale acts as a deterministic slave of the resolved scales. However, in reality there is variability at local scales *unrelated to the resolved scales*. Thus, Eq. 11.5 should take into account that $F'(\phi')$ can not completely be specified as a function of $\overline{\phi}$, but that formulation (11.5) should be replaced by

$$F'(\phi') \sim S(\vec{\alpha}) \tag{11.9}$$

with a random process S with parameters $\vec{\alpha}$ which are conditioned upon the resolved state $\overline{\phi}$:

$$F'(\phi') \sim S\left(R(\overline{\phi})\right) \tag{11.10}$$

When the mean value $\mu(\overline{\phi})$ is the only parameter in the vector $\vec{\alpha}$, which depends on $\overline{\phi}$, then the distribution S may be written as

$$S\left(R(\overline{\phi})\right) = \mu(\overline{\phi}) + S' \tag{11.11}$$

with a conditional mean value $\mu(\overline{\phi})$ and a random components with zero mean value ($E(S') = 0$) and uncertainty unrelated to the resolved scales. Specification (11.8) equals specification (11.11) if $S' = 0$ and $\mu(\overline{\phi}) = Q(\overline{\phi})$.

The role of statistics here is to first suggest a suitable distribution S, and later to condition the free parameters $\vec{\alpha}$ in a manner such that observed or otherwise physically meaningful relationships are approximately satisfied. These parameters often include parameters such as means, variances and lag correlations or spectra.

A comparison of a randomised parameterisation with a conventional parameterisation has been provided by Bauer and Weisse (2000), who ran the ocean wave model WAM for an extended period of 6 months with analysed winds. These winds are available only every 6 hours. Usually the variability within these 6 hours, related to various high-frequency meteorological events such as passage of fronts and the associated gustiness is disregarded, and the wind is either kept constant or, as in the present case, linearly interpolated.

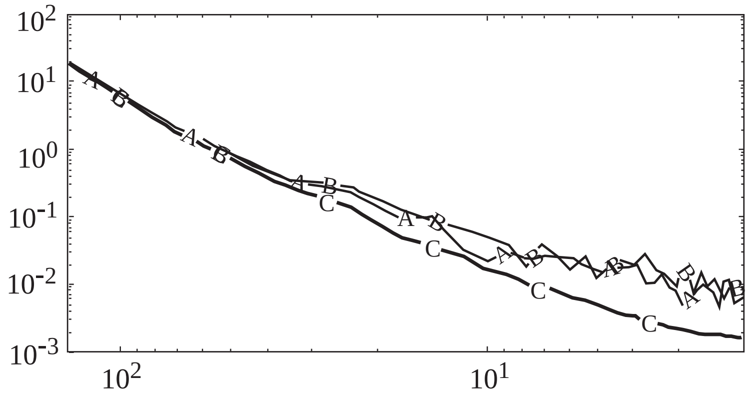

Fig. 11.1. Spectra of wind speeds at Ekofisk; time resolution 20 min (i.e., the abscissa of 9 corresponds to 9 x 20 min = 3 hours), units: (m/s) 2 x 20 min. A is the spectrum of the original series, B is for the 6 hourly wind speed with randomised interpolation and C with linear interpolation.
From Weisse and Bauer (pers. comm.)

In their study, Bauer and Weisse considered the linear interpolation as the "conditional expectation" parameterisation (11.8) of short term atmospheric fluctuations. A randomised version was constructed from time series of wind observations every 20 minutes. It returns for any instantaneous pair of wind speeds 6 hours apart a consistent series of 20 minutes wind speeds. This constitutes their "randomised parameterisation" (11.11). The success of the random number generator is demonstrated in Fig. 11.1, displaying three auto spectra. One spectrum (A) is from the original series observed every 20 min. In the other two spectra, first a wind speed was sampled every 6 hours, and then either linearly interpolated (C) or random numbers were inserted (B). The spectra are very similar for time scales longer than 6 hours (corresponding to an abscissa of 18); the linear interpolation causes a severe underestimation of the high-frequency variance, which is entirely recovered by the randomised parameterisation.

The effect of the two different formulations of high-frequency wind variations on the statistics of ocean waves was analysed in terms of the distribution of wave heights every 20 minutes within 3 days. For each interval of three days, the 6th largest wave height is determined, i.e., the 97% from the sample of 216 wave height values within three days. The time series of these 97% quantiles for 60 consecutive 3-day intervals for the two WAM simulations is displayed in Fig. 11.2. Obviously, the two time series are very similar, so that the presence of short term fluctuations does not induce significant low-frequency variations; thus, the randomised parameterisation is not required for the simulation of the overall statistics. However, a closer inspection reveals that the distributions are shifted to taller waves (not only in terms of 97% quantiles but also 50% quantiles); that is, the presence of high-frequency variations may increase the height of extreme values by a few decimeters.

It seems that while a "randomised parameterisation" (11.9) is theoretically attractive, in most applications the additional introduction of variability is inconse-

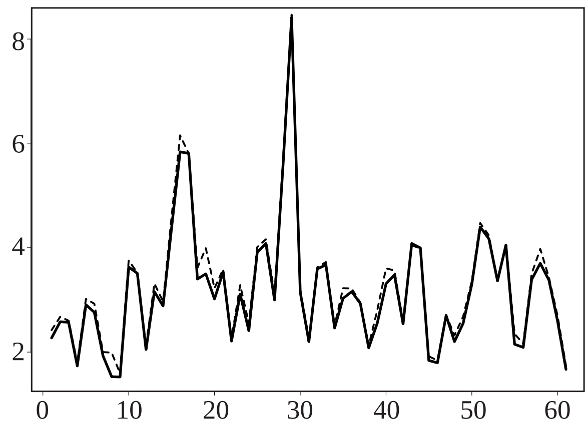

Fig. 11.2. Time series of 97 % quantiles of significant wave height at 30°N, 30°W calculated for consecutive 3-day intervals. The solid line is from the run with linearly interpolated winds, and the dashed curve from the simulation with a randomised specification. Units: m/s. Abscissa: number of 3-day long intervals.
After Bauer and Weisse (2000)

quential, because the dissipative mechanisms of the models efficiently remove variability on the smallest scales. However, experiments with an Energy Balance Model (von Storch et al. 1999), an idealised ocean model (Timmermann 1999) and high-frequency wind forcing of a wave model (Bauer and Weisse 2000) demonstrate the potential of this approach.

11.3.2 Analyzing Integrations of Quasi-Realistic Models

When quasi-realistic models are integrated, the purpose may be forecasting, simulation or executing a numerical experiment.

11.3.2.1 Forecasting

Forecasts, like the prediction of stream flows in rivers, are made to provide users with timely information to allow for adaptation to a changing environment. This information is usually a point-value, such as a river level at a given time, or an interval, such as a temperature range, or a probability, such as the probability for rain on a given day. The information provided by a single forecast is easy to grasp, but the information provided of the forecast system needs a *statistical analysis* of the predictive skill, concerning the frequency of hits or the expected error (Murphy and Epstein 1989). Sometimes efforts are made to add to the prediction of state variables a prediction of the skill of the forecast itself.

In terms of the *state space model*, forecasting means to first determine "initial states" by suitably solving the observational Eq. 11.2 and then integrating the state space Eq. 11.1 forward in time. From the large-scale forecast, local forecasts are derived from "model output statistics" (Klein and Lahn 1974), which amount to invoking another observation model. In most cases, the uncertainty terms ε and δ are disregarded, but not always.

11.3.2.2 Simulations

Simulations are made to generate a quasi-realistic extended trajectory of the considered system. Often, it is simply impossible to observe with sufficient spatial and temporal resolution the development of the system. For instance, a high resolution current field of the North Sea can not determined from observations, let alone for an extended time. Then, it is advisable to run a quasi-realistic model instead. However, the result of this model is highly complex, not as complex as reality, but in practical terms much too complex to grasp the wide range of phenomena and their dynamics. Thus, the evaluation of a simulation requires the skillful *statistical analysis* of the output of such a model. Without such an analysis, the researcher can only consider a limited number of events, analyse these events in terms of the processes involved, and compare them with observations.

An important aspect with such simulations is the validation of the models. There are cases where models have shown sensitivities, mainly because the models replicate not the real system but somewhat different, overly sensitive systems.

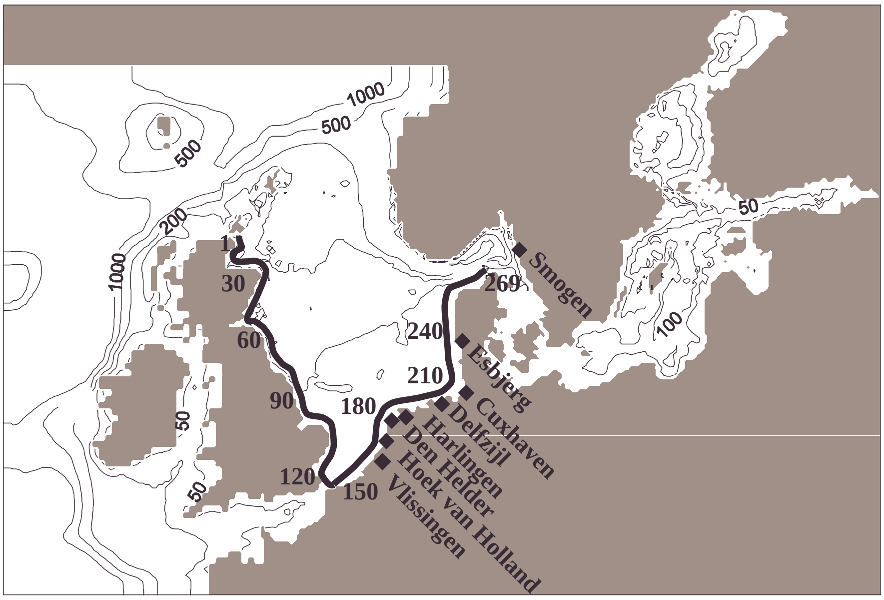

Fig. 11.3. Integration area in Langenberg's study. The black line indicates the location of the considered 270 near-coastal water level variations
From Langenberg et al. (1999)

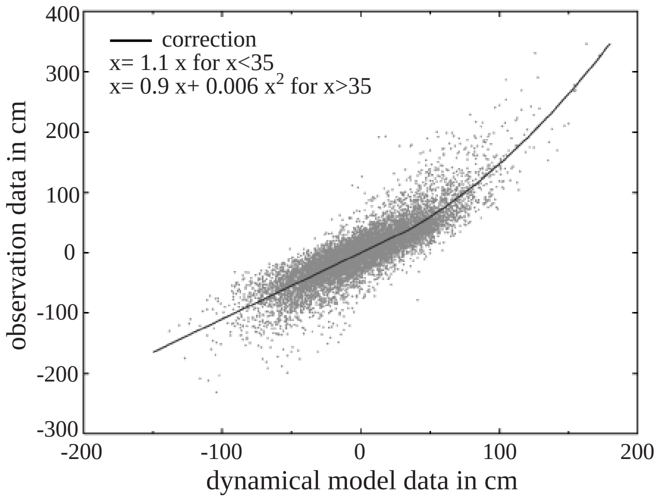

Fig. 11.4. Observations model, relating near coastal water level (horizontal axis) to shore-line water level (vertical axis), for Langenberg's 2-D North Sea model

In terms of the *state-space model*, simulations are done by integrating Eq. 11.1, mostly by disregarding the ε-term. Confirmation is done by invoking an observational model (11.2) and comparing the estimated ω-values with observations.

Examples of such simulations are numerous; one such example is from Langenberg et al. (1999), who ran a model of the hydrodynamics of the North Sea over 40 years. They analysed changes in the time-mean and in the intra-annual statistics of coastal sea level. Their model F was a 2-D barotropic model of the North Sea (Fig. 11.3). This model simulates water levels in the interior of the North Sea and in the near-coastal area. Observations, on the other hand, are available on shore locations only. With help of a tide gauge, an observational model P was designed (Fig. 11.4). Water levels at the shore-line are for moderate water level variations about 1.1 times the off-shore water levels; the amplification increases for water levels larger than 50 cm above normal.

Figure 11.5 displays the result of this simulation. Since observational models are available only for a few locations where tide gauges exist off-shore, unprocessed sea level trends are shown. The heavy line represents the change in the winter mean, which is of the order of 2 mm/year along the eastern coast and negligible along the western coast. Thus, within 1955 – 93, the mean sea level has risen due to changing meteorological conditions by about 8 cm. The intramonthly 90% quantiles, representative for storm activity, has remained practically stationary.

11.3.2.3 Numerical Experimentation

Quasi-realistic models, which have been tested in simulations of the historical record, may also be used to test the system's sensitivity to changes of boundary conditions, forcing functions and internal processes. One simulation is done with

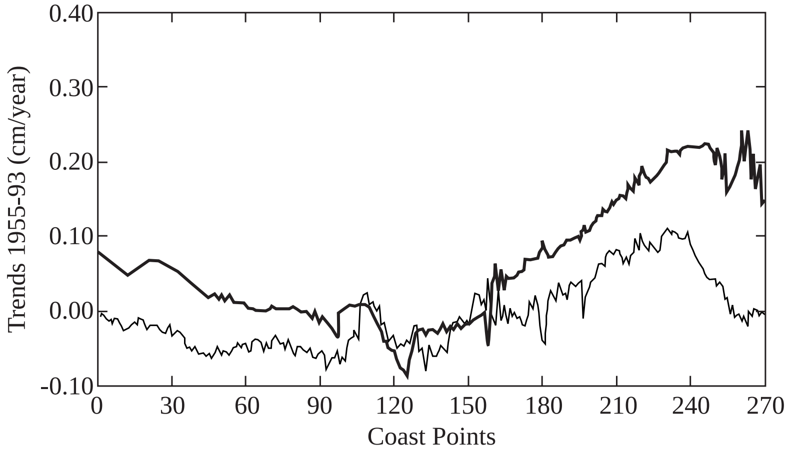

Fig. 11.5. Trends of winter means (heavy line) and winter 90% quantiles (light line) of winter high tide levels for 270 near-coastal locations along the North Sea coast (as indicated in Fig. 11.3) simulated in a 1955 – 93 model integration. Units: cm/year. From Langenberg et al. (1999)

"control" conditions, and others with a limited number of factors modified in a perfectly controlled manner. In this way, the effect of the "treatment" may be examined. Like in medical research, a *statistical analysis* of the outcome of such an experiment is often required because of the inherently noisy character of environmental systems. Usually, decision rules are required to reject null hypotheses, "treatment has no effect," with a given risk.

These "treatments" may be dramatic as the effect of opening the Central American Isthmus in Panama (Maier-Reimer et al. 1990), but may also feature different parameterisations of the cirrus clouds (Lohmann and Roeckner 1995), or

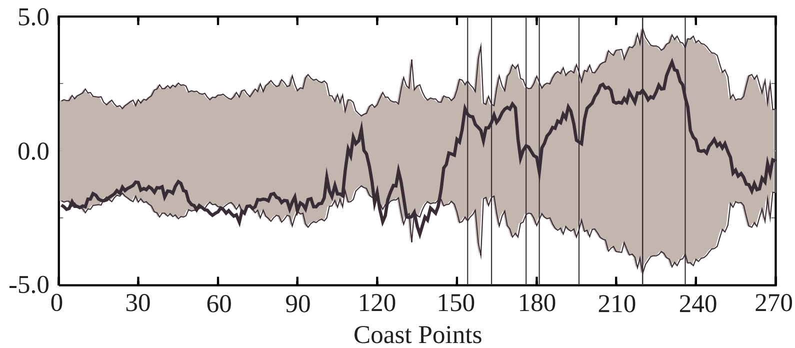

Fig. 11.6. Change of winter 90% quantiles of winter high tide levels for 270 near-coastal locations along the North Sea coast (as indicated in Fig. 11.3) from today to the time of expected doubling of atmospheric CO_2 concentration, in cm. Shaded area: 95% confidence interval for present natural variability; heavy line: expected anthropogenic change.
From Langenberg et al. (1999)

specific aspects such as the effect of an accelerated weather stream on the wave statistics (Bauer et al. 2000).

In the following we briefly sketch the outcome of an experiment with the North Sea model mentioned above, dealing with expected changes of coastal sea level statistics associated with the expected ongoing increase of atmospheric greenhouse gas concentrations (Langenberg et al. 1999). In this case, two atmospheric model runs had generated quasi-realistic 5-year sequences of weather, representative for present day conditions and for the year 2050, when atmospheric CO_2 levels are expected to have doubled. These forcing fields were used in two 5-year simulations with the above mentioned 2-D North Sea model, and changes in the intra-monthly 90% quantiles calculated. These changes were compared with the range of variations to be expected (at least in the framework of the model) in the present climate regime. For the coastal points displayed in Fig. 11.3, this range of variations is displayed as 95% confidence intervals in Fig. 11.6, and the simulated change of storm surge statistics as a solid line. Obviously, the simulated change varies within the 95% confidence band, indicating that the treatment "modified atmospheric composition" has little effect on the statistics of storm surges along the North Sea coast.[1]

11.3.3 Merging Dynamical Knowledge and Observational Evidence

A standard exercise is the determination of parameters $\vec{\alpha}$ in the state space Eq. 11.1. In that case, the observational evidence ω is usually a series of statistics, such as spatial variances, and the parameters are determined such that

$$E\left(\left\|\omega_{t+1} - P\left(F(\phi_t, \vec{\alpha}, t)\right)\right\|\right) = \min, \tag{11.12}$$

where $E(\cdot)$ represents the expectation operator. In practical cases, only a limited number of observations is available, and the usual statistical assumptions about stationarity and ergodicity have to be made before reasonably replacing the expectation operator by a finite sum. The minimisation can be achieved by variational methods. Among many others, an example related to the diffusion in the ocean is provided by Schröter and Wunsch (1986), for nutrient transports in rivers by Bülow et al. (1998).

Another problem is that of the operational, consecutive analysis of spatial distributions. This procedure has been developed in weather forecasting, where good knowledge about the (initial) state of the atmosphere is mandatory for a good weather forecast. Therefore, the routine measurements taken from radio sondes, weather stations etc., are blended with a recent forecast. In this way, a complete 3–D analysis of the state of the atmosphere is obtained, which is then used as the initial state for a forecast. The forecast for the next analysis time serves as the "first guess" to be blended with the new observations.

[1] Note that the effect of an overall increase of the ocean's volume, due to thermal expansion and changes of the mass of glaciers and ice sheets is not included in this analysis.

A large number of different techniques have been developed in the past; an overview, illustrated with examples, is provided by Robinson et al. (1998). The above sketched consecutive approach can be seen as a kind of *Kalman filter* (e.g., Honerkamp 1994, or Jones 1985) with a state model (11.1) and an observation model (11.2). With the state model a forecast

$$\phi_t{}^f = F\left(\phi_{t-1}, \vec{\alpha}, t-1\right) \tag{11.13}$$

is prepared, from which consistent observations ω_t^f are estimated with the help of the observation model:

$$\omega_t{}^f = P\left(\phi_t^f\right) \tag{11.14}$$

These estimated values are compared with the actual observations ω_t and a final "best estimate" ϕ_t is determined as linear combination

$$\phi_t = \phi_t{}^f + K\left(\omega_t - \omega_t{}^f\right) \tag{11.15}$$

with a suitable operator K, which depends on the covariance matrices of the noise terms ε and δ in Eq. 11.1 – 2.

11.4 Reduced "Cognitive" Models

As already discussed, are quasi-realistic models are too complex to allow for *understanding,* in much the same way that simply looking at environmental phenomena without a-priori theoretical reasoning is rarely a means of understanding the underlying mechanisms. Therefore, certain conceptual "models" of the phenomena are usually set up and data are checked to the extent they are consistent with the conceptual model. The conceptual model is an idealised description of the real situation, stripped down to a few relevant processes and their interaction. When supported by empirical evidence, the conceptual model embodies "understanding."

"Conceptual model" means that the functional from of dependencies is fixed, but that certain parameters are left to be fitted to the data. They may formally expressed in terms of *Principal Interaction Patterns*, as introduced by Hasselmann (1988). The concept has been implemented fully only in a few cases (e.g., Kwasniok 1996; Achatz and Schmitz 1997), but simpler cases such as *Principal Oscillation Patterns* (von Storch et al. 1995) and regressions techniques are abundant in the literature.

11.4.1 Principal Interaction Patterns[3]

To formalise the PIP concept, we assume that the full system is represented by a state space model (11.1) and a linear *observation equation* (11.2) for the observed variables ω,

$$\omega_t = P\phi_t + \text{noise} = \sum_{j=1}^{k} \vec{p}^{\,j}\phi_t + \text{noise}. \tag{11.16}$$

with the row vectors $\vec{p}^{\,j}$ forming the matrix P – being the Principal Interaction Patterns, spanning a low-dimensional state space. In the state space equation $\phi_{t+1} = F(\phi_t, \vec{\alpha}, t) + \varepsilon$ the operator F represents a class of models that may be nonlinear in the dynamical variables ϕ_t and depends on a set of free parameters $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots)$.

Different from the data analysis and model verification problem, the dimension *k* of the state variable ϕ is considered to be much smaller than the dimension of the observations *m*. Indeed, *k* is usually of the order of 20 or much less. Matrix *P* generally has many more columns (*m*) than rows (*k*). The system equations 11.1 therefore describe a dynamical system in a smaller phase space than the space that contains ω_t.

The error-term ε in (11.1) is considered here a noise term. It is often disregarded in nonlinear dynamical analyses. However, disregarding the noise in low-order systems ($k < 20$) usually changes the dynamics of the system significantly, since the low-order system is a closed system without noise (cf. Timmermann 1999; von Storch et al. 1999). However, components of the climate system, such as the tropical troposphere or the thermohaline circulation in the ocean are never closed; they continuously respond to "noise" from other parts of the climate system, hence the noise term. It is doubtful if the fundamental assumption, namely that the low-order system is governed by the same dynamics as the full system, is satisfied when the noise is turned off.

When fitting the state space model from Eq. 11.1, 11.16 to a time series, the following must be specified: the class of models *F*, the patterns *P*, the free parameters α and the dimension of the reduced system *m*. The class of models *F*, must be selected *a priori* on the basis of physical reasoning. The number *m* might also be specified *a priori*. The parameters α and the patterns *P* are fitted simultaneously to a time series by minimising the mean square error $\varepsilon[P;\vec{\alpha}]$ of the approximation of the (discretised) time derivative of the observations ω by the state space model:

$$\varepsilon\big[P;\vec{\alpha}\big] = E\Big(\big\| \omega(t+1) - P\big(F\big[\phi(t), \vec{\alpha}, t\big] - \phi(t)\big) \big\|^2 \Big) \tag{11.17}$$

[3] Following the presentation in von Storch and Zwiers, 1999, Section 15.5.

11.4.2 Principal Oscillation Pattern Analysis[4]

A standard “cognitive model” is the identification of normal modes or “waves” in geophysical fluid dynamics. In one approach, the basic equations are simplified and linearised until they may be formulated as

$$\omega_{t+1} = A\omega_t + \varepsilon \tag{11.18}$$

where A is the “system’s matrix” and ε the error introduced by the manipulations. The eigenvectors of the matrix A are the normal modes. Typical examples for such modes are all sorts of “waves”, as for instance Kelvin waves.

Besides this entirely dynamical approach, there is a statistical variant of Eq. 11.18, namely when then matrix A is *not* derived from dynamical reasoning but fitted to data. In that case ε is usually considered to be *white noise*. Then Eq. 11.18 describes a discrete multivariate first-order autoregressive process.[5] The system matrix A may be estimated through

$$A = \sum_1 \sum^{-1} \tag{11.19}$$

where $\sum$ and denote $\sum_1$ the lag-0 and lag-1 covariance matrices of ω, which are easily calculated from the data. The eigenvectors of this estimated matrix A are the *Principal Oscillation Patterns* (POPs). Each state ω may then be expanded, or approximated by the POPs $\vec{e}^{\,k}$.

$$\omega = \sum \phi_k \vec{e}^{\,k} \tag{11.20}$$

The ϕ_k are the POP coefficients, and represent the dynamical state of the system:

$$\phi_{k,t+1} = \lambda_k \phi_{k,t} + \varepsilon_k \tag{11.21}$$

where λ_k is the eigenvalue of belonging to . is the noise projected into the subspace spanned by the POPs.

Note that Eq. 11.20 is an observation equation (11.2) and (11.21) a state equation (11.1).

The POP analysis is illustrated by an application to equatorial variability (von

4 Following von Storch (1993).

5 The relation between empirical and dynamical modes has been investigated by Schnur et al. (1993), who calculated from quasi-geostrophic theory the dynamical modes describing the extratropical atmospheric variability, and also used the POP approach on a long sequence of analysed geopotential height data. The spatial and temporal characteristics of the most significant POPs were very similar to the most unstable waves in the stability analysis, but the POPs also identified modes representative of the evolution of finite-amplitude waves. Thus, the POPs appear to be useful descriptors of the variability in cases where the dynamics were complex.

Fig. 11.7. Amplitude time series of two POP modes identified in the daily intraseasonal variability monitored by three Equatorial moorings at 165°E, 140°W and 110°W. The left curve refers to the 120 day mode, and the right one to the 65 day mode. The amplitudes are normalised to standard deviation one. The years are given as May to April intervals (thus "1984" represents the time from May 1984 until April 1985). From von Storch (1993)

Storch 1993). The goal was to investigate the modes of intraseasonal variability in 7 years of moored measurement in the upper equatorial Pacific ocean (Hayes et al. 1991). Oscillatory modes were searched for by using a POP analysis of daily averages of horizontal current at three equatorial locations, 165°E, 140°W and 110°W, for various depths.

The data were first filtered by an EOF analysis to suppress small scale noise, and the EOF coefficients filtered in the time domain to eliminate the variability on periods larger than about half a year. The POP analysis then yielded two oscillatory modes (complex pairs of eigenvectors with the above properties). The normalised amplitude time series are displayed in Fig. 11.7.

One oscillatory mode has a period of $T = 2\pi / \omega = 65$ days and a damping time of $\tau = 73$ days. The amplitude time series reveals an annual cycle with a semi-annual component. The intraseasonal mode activity is strongest during solstice conditions and weakest during equinoctial conditions, and it is enhanced during warm ENSO conditions (1986/87 and 1990).

The other oscillatory mode, operating at a period of about 120 days and a damping time of about 105 days, is affected by the state of the Southern Oscillation as well with enhanced activity during warm episodes and reduced activity during the cold 1988 event.

The spatial amplitudes and phases of the two modes, in terms of zonal currents, are displayed in Fig. 11.8. Both modes represent eastward traveling signals.

The 120-day mode has its largest amplitude, with typical maximum values of about 16 cm/s, at 50 m depth at 65°E and 160 m depth at 140°W. In contrast, the 65 day mode has maximum zonal current anomalies at upper levels (50 m and above) in the eastern part of the basin, with a typical maximum of 12 cm/s at 140°W and 19 cm/s at 110°W. The zonal current 120-day signal propagates in about 60 days from 165°E to 110°W, so that the phase speed is about 1.8 m/s. The

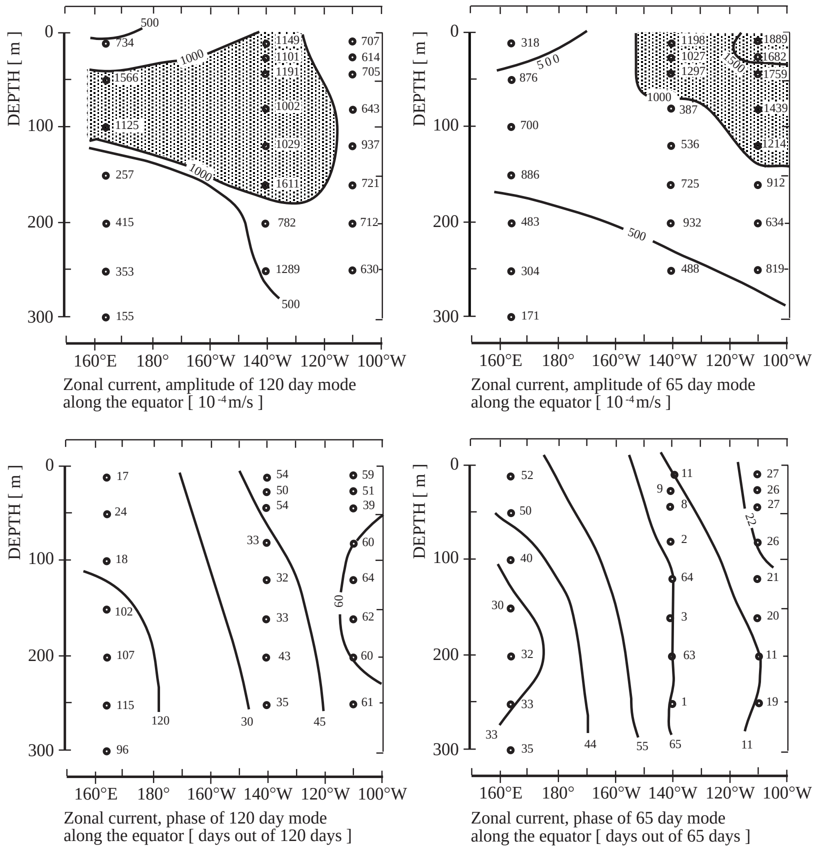

Fig. 11.8. Amplitude and phase distributions of the two eastward propagating oscillatory POPs of the zonal currents at three Equatorial moorings at 165°E, 140°W and 110°W. The coefficient time series are normalised to unity so that the amplitude pattern represents typical distributions in 10^{-4} m/s. The phases are given in days relative to the base period of 120 and 65 days. From von Storch (1993)

phase lines are vertically tilted, with the upper levels lagging the lower levels by about 15 days. The phase speed for the 65 day mode is estimated to be 2.1 m/s. At the two eastern positions, the phase lines are tilted, with the lower levels leading the upper levels by about 8 days. The two modes are not correlated; their time coefficients share a correlation of about –0.25. The two modes are, however, pattern-wise similar and are **not** orthogonal. Indeed, the POP analysis does not require that the modes be orthogonal.

Acknowledgments

Ralf Weisse provided me with the example on the sensitivity of the wave model to the different formulation of the high frequency wind forcing.