Tellus (2002), 54A, 1–13
Printed in UK. All rights reserved

TELLUS
ISSN 0280–6495

Dynamic models for Icelandic meteorological data sets

By JOSEPH EGGER[1]* and TRAUSTI JÓNSSON[2] [1]*Meteorologisches Institut, Universitaet Muenchen, Theresienstrasse 37, 80333 Muenchen, Germany;* [2]*Vedurstofa Íslands, Bustadavegur 9, 150 Reykjavík, Iceland*

(Manuscript received 24 November 2000; in final form 28 May 2001)

ABSTRACT

Dynamic models are derived from pairs of time series of daily surface and 500 hPa height observations in the Icelandic region in order to interpret these data. The models describe the motion of data points in the phase plane spanned by a pair of variables. For each pair, a regression model of first order (R-model) is derived, which assumes linear motion in the phase plane. The coefficients of the corresponding Fokker–Planck equation (FP) are also derived from the data. This equation describes changes of the probability density distribution of the data pair and takes nonlinear aspects into account. The data are interpreted in the light of both models. The R-model yields a mean rotation and convergence which characterize the basic type of motion inherent in a data set. It is found in all cases that the timescale of contraction is shorter than the "dynamic" timescale linked to the rotation. The FP equation yields a mean motion which resembles quite often that in the R-model as well as a diffusive part. The R-model is not satisfactory in some cases which involve the meridional surface winds. For example, southerlies (northerlies) and temperature rotate clockwise (anticlockwise). Meridional and zonal surface winds rotate anticlockwise in a southerly flow, but clockwise in a northerly one. Therefore, new aspects of the Icelandic meteorology can be extracted form the data by use of the FP equation.

1. Introduction

Iceland is located close to the center of the Icelandic low, the semipermanent low-pressure cell in the North Atlantic, typically located between Iceland and Southern Greenland (Serreze et al., 1997). Individual cyclones are frequently found in this region, particularly in winter. Both cyclogenesis and cyclolysis are commonly observed. In short, cyclone activity dominates the climate of Iceland although high-pressure cells may be found as well. As is well known, the dynamics of the Icelandic region are of global importance. The Icelandic low plays a key role in the North Atlantic oscillation (e.g. Barnston and Livezey, 1987), a phenomenon of global scale. All this suggests that one should analyse data from this region as intensively as possible. A considerable number of meteorological observations have been collected over the years in Iceland, the longest record being that initiated by Thorlacius in 1845 in Stykkisholmur. These data form the basis of the comprehensive climatologies of Iceland by Eythorsson and Sigtryggson (1971) and Einarsson (1984). Einarsson (1991) discussed correlations between monthly mean temperatures at 32 stations. He found, for example, that variations of monthly temperatures from year to year are mostly in phase for the whole country except during summer. Hartmann (1974) analysed rawin sonde data of weather ships located south-west of Iceland. He found strong phase relationships in the vertical for waves with periods between three and seven days, i.e. for oscillations within the range of cyclone activity. Jónsson (1997) extended

* Corresponding author.
e-mail: J.Egger@lrz.uni-muenchen.de

the work of Einarsson (1991) by including pressure data and deriving corresponding regression equations. He found, *inter alia*, that monthly mean temperatures in Stykkisholmur are positively correlated with southerly winds at 500 hPa over Iceland as well as with the 500 hPa height. They are negatively correlated with the zonal wind component. Jónsson (1994) presented a precipitation climatology of Iceland. A comprehensive overview of research on Icelandic climate is found in Gardarsson (1999).

Despite all these efforts we are still far from an exhaustive evaluation of all the Icelandic data available. Here, we report on further, admittedly small steps towards this goal. Basically we extend the approach of Jónsson (1997) by relating pressure, geopotential height and temperature data of various stations, including sea surface temperatures. However, the analysis of the data will be based on an approach with novel features. First, a conventional regression model of first order (R-model) will be adapted to time series of selected pairs of stations. The eigenvalues of this model will be evaluated in order to capture the dynamics of the R-model. This technique is exploited extensively in the so-called POP method (von Storch and Zwiers, 1999; POP, Principal Oscillation Pattern). We extend this analysis by the derivation and interpretation of Fokker–Planck (FP) equations in the phase space of the station data. While the POP analysis assumes that the underlying process is linear, the FP equations do not rely on this assumption. It will be demonstrated that additional features of the time series can be derived this way. Even this part of our approach is not completely novel. For example, FP equations have been derived from turbulence and traffic flow data (Friedrich and Peinke, 1997; Friedrich, 1999, personal communication) and from ice core data (Ditlevsen, 1999). Demaree and Nicolis (1990) fitted an FP equation to Sahelian precipitation data assuming a drift term of third order. However, we are not aware that this technique has been used to analyse time series of standard meteorological data such as temperature or pressure. As pointed out by Egger (2001) it is more straightforward to derive a Master equation form the data instead of an FP equation. The Master equation can also be used to predict probability density distributions. However, the interpretation of the terms of the Master equation is more difficult. Thus the FP equation is to be preferred if one is specifically interested in data interpretation, as is the case here.

In principle, all these techniques can be applied to any type of data set, provided the length N of the available time series is such that results are statistically reliable. Correspondingly, the POP analysis has been applied to fields based on a rather large number of stations (e.g. Schnur et al., 1993). The number of data needed to determine the coefficients of FP equations is, however, much larger than that required by in the POP analysis. It is for that reason that we restrict our investigation to data pairs so that the resulting FP equation is just two-dimensional. Daily observations are, of course, more abundant than monthly means. This suggests that we concentrate on Icelandic time series of daily pressure and temperature values. Both data from pairs of stations separated horizontally as well as pairs at different levels will be analysed.

2. Data

The following data sets have been used.

2.1. Set RD

Time series of daily surface pressure values p_s (0.1 hPa; 12 h) and air temperature T (0.1 °C) are available for the stations Reykjavik at the south-eastern tip of Iceland and Dalatangi situated at a promontory at the east coast (Fig. 1). The data cover the interval 1 January 1949 to 7 April 2000. The first variable x of the pair is the pressure difference Dalatangi–Reykjavik while $y = T(\text{Reykjavik}) + T(\text{Dalatangi})$. Although Dalatangi is situated farther north (65°8′N) than Reykjavik (64°8′N), the variable x nevertheless represents the meridional component of the geostrophic surface wind to a good approximation. Dalatangi must be seen as a somewhat extreme station, strongly reflecting the oceanic conditions due to its exposed position on a peninsula. The length of the record is $N = 18\,813$.

2.2. Set MZRT

This set is part of the data set used by Jónsson (1997). The data are derived from grid point values of the NMC-reanalysis data covering the

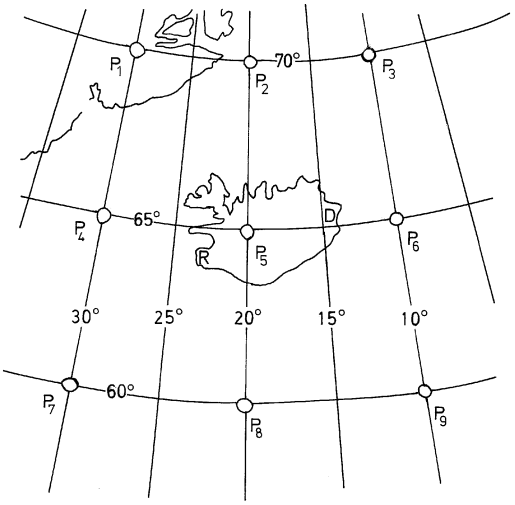

Fig. 1. Location of the grid points on which the data set *MZRT* is based. The positions of the stations Reykjavik (R) and Dalatangi (D) are also given.

interval 1 January 1958 to 30 June 1998 with $N = 14\,791$. Fig. 1 shows the position of the grid points where daily values of the 500 hPa and 1000 hPa heights *(P)* are available. The relative topography RT 500/1000 hPa is defined at the central point. Out of these data we form the parameters

$$M = (P_3 + P_6 + P_9) - (P_1 + P_4 + P_7) \tag{2.1}$$

(*B* in Jónsson, 1993) and

$$Z = (P_7 + P_8 + P_9) - (P_1 + P_2 + P_3) \tag{2.2}$$

(*A* in Jónsson, 1997) where M characterizes the meridional component of the geostrophic wind and Z the zonal component. Indices s and 5 will denote surface and 500 hPa values, respectively, of M and Z. All the parameter pairs to be subjected to the analysis procedure are listed in Table 1.

3. Analysis methods

Let X_n, Y_n denote the time series of the variables x, y where n is the time index and X_n, Y_n are the observations of the first and second station. Mean values are removed, as is the annual cycle. All data to be analysed are normalized with their standard deviation so that we investigate time series of dimensionless variables of unit standard deviation. We use also a dimensionless time scaled by the interval $\mathrm{D}t = 1$ day.

It is convenient to discuss the results in the phase plane of the variables x, y where the observation at time $n\mathrm{D}t$ is represented by a point (X_n, Y_n). This point jumps from one observation to the next, so that the total data set is represented by a complicated "trajectory". We introduce a probability density function $f(x, y, t)$ (pdf), which describes the probability of making an observation in a specified region G of the phase space at time t. Thus

$$L = \int_G f(x, y, t)\, \mathrm{d}x\, \mathrm{d}y$$

is this probability. In general, it is difficult if not impossible to estimate f from atmospheric data as a function of time. One would need ensembles of data sets for this purpose. However the time mean distribution $\bar{f}$ (where the bar denotes a time mean) can be approximated by the two-dimensional histogram of relative frequencies in the (x, y)-plane (see, for example, Fig. 2).

Table 1. *Parameter combinations analysed. Also given are the corresponding coefficients of the R-model*

Dataset	x	y	α_1	α_2	β_1	β_2
RD	p(Reyk) − p(Dala)	T(Reyk) + T(Dala)	−0.70	−0.34	0.04	0.13
*MZRT*1	M_s	RT	−0.47	−0.32	0.07	0.02
*MZRT*2	Z_s	RT	−0.32	−0.32	0.15	0.07
*MZRT*3	M_s	Z_s	−0.46	−0.32	0.15	0.06
*MZRT*4	M_5	RT	−0.40	−0.26	0.14	−0.14
*MZRT*5	M_5	Z_5	−0.34	−0.24	0.09	0.00
*MZRT*6	Z_5	RT	−0.24	−0.31	0.01	0.06
*MZRT*7	M_5	M_s	−0.56	−0.29	0.29	−0.19
*MZRT*8	Z_5	Z_s	−0.39	−0.31	0.20	−0.00

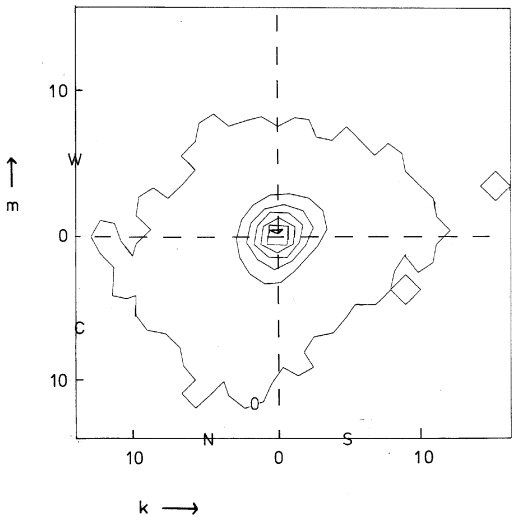

Fig. 2. Probability density distributions in the (k, m)-grid of the phase plane of the variables x, y defined for the data set RD (x is the pressure difference Dalatangi–Reykjavik; y is the sum of temperatures at both stations). The grid size $Dx = 0.5$ corresponds to half the standard deviation here and in all following figures. Contour interval 0.01. The symbols S(N) and W(C) denote southerly ($x > 0$) and northerly ($x < 0$) surface winds and positive (negative) temperature deviations, respectively. Dashed: coordinate axes; $x = kDx$; $y = mDx$.

3.1. *Regression model*

The model equations to be used are

$$X_{n+1} = X_n + \alpha_1 X_n + \beta_1 Y_n + Z_{1n}, \tag{3.1}$$

$$Y_{n+1} = Y_n + \beta_2 X_n + \alpha_2 Y_n + Z_{2n}, \tag{3.2}$$

where α_i, β_i are constants and z_i are white noise processes with standard deviations σ_i. It is straightforward to determine the model parameters from the observations. For example, α_1 and β_1 follow from

$$\overline{X_n X_{n+1}} = (1 + \alpha_1)\overline{X_n^2} + \beta_1 \overline{X_n Y_n}, \tag{3.3}$$

$$\overline{Y_n X_{n+1}} = (1 + \alpha_1)\overline{X_n Y_n} + \beta_1 \overline{Y_n^2}, \tag{3.4}$$

where the bar stands now for an average over all observation times. The standard deviations are obtained from

$$\sigma_1^2 = [1 - (1 + \alpha_1)^2]\overline{X_n^2} - 2(1 + \alpha_1)\beta_1 \overline{X_n Y_n} - \beta_1^2 \overline{Y_n^2}, \text{ etc.} \tag{3.5}$$

The model equations (3.1) and (3.2) specify a "velocity"

$$\boldsymbol{v}_r = (\alpha_1 x + \beta_1 y, \beta_2 x + \alpha_2 y) \tag{3.6}$$

which describes the motion of the state vector (x, y) per time step in the absence of white noise forcing. In other words, eq. (3.6) describes a mean motion of vorticity

$$\zeta = \beta_2 - \beta_1 \tag{3.7}$$

and divergence

$$\delta = \alpha_1 + \alpha_2. \tag{3.8}$$

Of course, $\zeta(\delta)$ is a vorticity (divergence) in phase space and not in physical space. The variable x leads y in an anticlockwise flow with $\zeta > 0$, whereas it is the opposite for a clockwise flow. Here and in the remainder the terms clockwise and anticlockwise will be reserved for flows in the phase plane, while the terms cyclonic and anticyclonic characterize physical flows. The R-model assumes that vorticity and divergence of the mean flow do not depend on x, y.

Further information on the motion in the phase plane can be obtained by computing the eigenvalues $\lambda_{1,2}$ of eqs. (3.1) and (3.2) following from the assumption

$$(X_{n+1}, Y_{n+1}) = \lambda(X_n, Y_n). \tag{3.9}$$

Inserting eq. (3.9) in eqs. (3.1) and (3.2) one obtains

$$\lambda_{1,2} = 1 + (\alpha_1 + \alpha_2 \pm d^{1/2})/2 \tag{3.10}$$

damped oscillatory modes or purely damped modes depending on the sign of the discriminant

$$d = (\alpha_1 - \alpha_2)^2 - 4\beta_1\beta_2.$$

Thus

$$4\beta_1\beta_2 > (\alpha_1 - \alpha_2)^2 \tag{3.11}$$

is required for oscillatory modes to exist. One may also introduce an eigenfrequency ω such that $x_{n+1} - x_n = \omega x_n$ and $\omega = \lambda - 1$. The related eigenvectors are the Principal Oscillation Patterns (POPs). Note that the motion is spiralling around the origin only if oscillatory modes exits.

3.2. Fokker–Planck equation

The FP equation is based on the assumption that the data to be analyzed can be modeled by two Langevin equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = A_x + B_{xx}\frac{\mathrm{d}W_x}{\mathrm{d}t} + B_{xy}\frac{\mathrm{d}W_y}{\mathrm{d}t}$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = A_y + B_{yx}\frac{\mathrm{d}W_x}{\mathrm{d}t} + B_{yy}\frac{\mathrm{d}W_y}{\mathrm{d}t} \qquad (3.12)$$

where A_x, A_y, B_{xx}, B_{xy}, B_{yx} and B_{yy} are functions of x and y and where $\mathrm{d}W$ is a Wiener process (e.g. Gardiner, 1983). Obviously, eq. (3.12) are generalizations of eqs. (3.1) and (3.2) where the drift velocity $\boldsymbol{v}_\mathrm{d} = (A_x, A_y)$ is related to $\boldsymbol{v}_\mathrm{r}$ via

$$(\alpha_1 x + \beta_1 y) \sim A_x\,\mathrm{D}t,$$
$$(\beta_2 x + \alpha_2 y) \sim A_y\,\mathrm{D}t, \qquad (3.13)$$

and where Z_1 and Z_2 correspond with the white noise terms in eqs. (3.12). Note, however, that eq. (3.12) are stochastic differential equations while eqs. (3.1) and (3.2) are discrete in time.

The FP equation corresponding to eqs. (3.12) is

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}(A_x f) - \frac{\partial}{\partial y}(A_y f) + \frac{1}{2}\left(\frac{\partial^2}{\partial x^1} D_{xx} f + 2\frac{\partial^2}{\partial x\,\partial y} D_{xy} f + \frac{\partial^2}{\partial y^2} D_{yy} f\right) \qquad (3.14)$$

(Gardiner, 1983). In eq. (3.14)

$$D_{xx} = B_{xx}^2 + B_{xy}^2,$$
$$D_{xy} = B_{xy}B_{yx} + B_{xx}B_{yy}, \qquad (3.15)$$
$$D_{yy} = B_{yy}^2 + B_{yx}^2.$$

The drift velocity $\boldsymbol{v}_\mathrm{d}$ in eqs. (3.12) and (3.14) and the diffusion terms (3.15) can be estimated from data. However, that can be done at discrete points only, i.e. the FP equation (3.14) must be replaced by a finite difference equation. We introduce $J \times J$ grid boxes of grid size Dx and define corresponding grid points (k, m) such that $k = x\mathrm{D}x$, $m = y\mathrm{D}x$ are the locations of the grid points. An estimate of the drift velocities and D_{xx} is obtained by

$$(A_x)_{km} = \langle X_{n+1} - X_n \rangle / \mathrm{D}t,$$
$$(A_y)_{km} = \langle Y_{n+1} - Y_n \rangle / \mathrm{D}t, \qquad (3.16)$$
$$(D_{xx})_{km} = \langle (X_{n+1} - X_n)^2 \rangle / \mathrm{D}t - (A_x)_{km}^2\,\mathrm{D}t,$$

with corresponding expressions for D_{xy}, D_{yy}. The brackets in eq. (3.16) denote an average over all cases where the data point (X_n, Y_n) is found in grid box (k, m). In principle, a limiting process should be performed in order to obtain these coefficients (e.g. Siegert et al., 1998). Given, however, only data of finite resolution in time with unit time step, eq. (3.16) is the best approximation possible. It is clear from eq. (3.16) that the drift terms are simply estimates of the mean tendency in a grid box, while the diffusion coefficients are proportional to the variances of these tendencies. There is, however, the technical problem that the drift velocities (3.16) may be too large given a grid size Dx. A numerical integration of the EP equation (3.14) is possible only if the stability criterion $(A_{x,y})\,\mathrm{D}t < \mathrm{D}x$ is satisfied at all grid points. If not, the time step has to be reduced.

Note that both the R-model and the FP equation rely only on covariances with lags zero and one. Correspondingly we may capture reasonably well dynamical processes with timescales of a few days, i.e. cyclone dynamics. One would have to use autoregressive models of higher order to detect oscillations with longer periods with reasonable confidence. That would also require the introduction of FP equations of higher dimension than two. At the moment we prefer to deal with the simplest case. Correspondingly we select variables where we can expect to find signals with timescales of a few days. The results of Hartmann (1974) suggest that both meridional wind and temperature are good candidates. The variations of the zonal wind tend to be concentrated at motions with longer timescales and may be more difficult to analyse with our models.

Given eq. (3.16) we can easily find out if the velocities predicted by the R-model are satisfactory, at least in principle. We just have to compare $\boldsymbol{v}_\mathrm{r}$ with $\boldsymbol{v}_\mathrm{d}$. This intercomparison is facilitated by introducing streamfunction ψ and velocity potential χ, so that

$$\bar{f}u = -\frac{\partial \psi}{\partial y} + \frac{\partial \chi}{\partial x},$$
$$\bar{f}v = \frac{\partial \psi}{\partial x} + \frac{\partial \chi}{\partial y} \qquad (3.17)$$

is the related pdf transport. It follows that

$$\nabla^2 \psi = -\frac{\partial}{\partial y}(\bar{f}u) + \frac{\partial}{\partial x}(\bar{f}v),$$
$$\nabla^2 \chi = \frac{\partial}{\partial x}(\bar{f}u) + \frac{\partial}{\partial y}(\bar{f}v). \qquad (3.18)$$

With eqs. (3.17) and (3.18) we are able to investigate both the rotational and the divergent flow of states in the phase plane. Boundary conditions pose no problem when solving eqs. (3.18) for ψ and χ because ψ and χ are constant outside the domain where $\bar{f} \neq 0$. We are free to choose $\psi = \chi = 0$ there.

The rotational part of the transport does not contribute to changes of f in eq. (3.14), as follows from inserting eq. (3.17) in eq. (3.14). Moreover, eq. (3.14) yields in steady state with $D_{xx} \sim D_{yy} \sim D$, constant, $D_{xy} \sim 0$ the relation, $\nabla^2\chi = \frac{1}{2}\nabla^2 \mathrm{D}\bar{f}$, so that

$$\mathrm{D}\bar{f} \sim 2\chi. \tag{3.19}$$

Diffusion has to balance convergence.

3.3. Statistical significance

An application of the more advanced FP equation makes sense only if the results differ significantly from those obtained with the R-model. For example, the difference of the velocity potential as derived from eq. (3.18) and that given by the R-model is displayed in Fig. 4c for the dataset RD. There is a relatively large domain of negative deviations and a small positive one. *A priori*, it is not clear if these patches of positive and negative deviations are significant. They may be random deviations caused by the finite length of the available time series. The related significance test is described briefly in the Appendix. It is based on Monte Carlo runs with the R-model. We will perform such significance tests for streamfunction and velocity potential.

4. Results

4.1. Surface data

The RD data set provides a pressure difference and a temperature T at the surface. The set *MZRT* allows us to calculate also u at the surface. Strictly speaking, the relative topography available in *MZRT* is not a surface variable. We include it, nevertheless, as a variable in this section on surface data when it comes to relating surface winds to this type of temperature.

4.1.1. Pressure and temperature at Reykjavik and Dalatangi; RD. As has been mentioned, the variables are $x = p(\text{Dalatangi}) - p(\text{Reykjavik})$ and $y = T(\text{Dalatangi}) + T(\text{Reykjavik})$. Given the locations of these stations, x is closely related to the meridional component of the geostrophic surface wind. Of course, the pressure at Dalatangi is highly correlated with that in Reykjavik (correlation coefficient 0.92). Deviations in Reykjavik are leading those in Dalatangi. It is the same with the temperatures.

The time mean pdf of x, y is displayed in Fig. 2. There is only one maximum and the distribution tends to lie on the line $x = y$. That part of the domain where $\bar{f} < 0.01$ is relatively large when compared to that more frequently visited with $\bar{f} > 0.01$. There is also an disconnected area of events with rather strong southwesterlies. Such events must be discarded, of course, in a data based derivation of the coefficients of the FP equation.

The auto- and crosscorrelations of both variables are shown in Table 2 for lags up to ten days. It is seen that the pressure difference deviations decay more rapidly than those of temperature. The correlation coefficient is 0.28. The crosscorrelation is antisymmetric with respect to lag $\tau = 0$. Southerlies (i.e. $x > 0$) bring higher temperatures the next day while there is no increase in the strength of the southerlies after a warm day.

The model parameters resulting from the R-analysis are given in Table 1. With $\alpha_1 \sim 2\alpha_2$, the memory of the pressure deviations is much shorter than that of the temperature anomalies, in

Table 2. *Autocorrelations AC and crosscorrelations CC of the data in set RD, where x = surface pressure difference Dalatangi–Reykjavik, y = sum of air temperatures at both stations. In CC_{ab} the variable b lags a*

Lag (days)	AC_{xx}	AC_{yy}	CC_{xy}	CC_{yx}
0	1·00	1·00	0·28	0·28
1	0·31	0·69	0·32	0·12
2	0·16	0·48	0·19	0·08
3	0·14	0·37	0·16	0·06
4	0·09	0·29	0·13	0·02
5	0·05	0·22	0·11	0·01
6	0·04	0·18	0·09	0·01
7	0·03	0·15	0·07	0·01
8	0·04	0·12	0·06	0·01
9	0·03	0·15	0·06	0·01
10	0·01	0·12	0·04	0·00

agreement with Table 2. Moreover, $|\alpha_i| \gg \beta_i$, so that decay dominates. Nevertheless, there is positive vorticity with $\zeta = 0.1$. The eigenvalues are real with $\lambda_1 = 0.29$, $\lambda_2 = 0.67$, so that $\lambda_i \sim 1 + \alpha_i$ because of the dominance of the damping.

The whiteness of the forcing as assumed in eqs. (3.1) and (3.2) has been checked by calculating Z_i in eqs. (3.1) and (3.2) as a residual and evaluating the corresponding autocorrelations. It is found that the forcing is indeed rather white. Autocorrelations of lag one are smaller than 0.02.

These results of the R-analysis must be compared to those obtained from the FP equation. The asymptotic pdf of eq. (3.14) resulting from an initial distribution where $f = 1$ at the central grid box comes rather close to that observed in Fig. 2 (not shown). The drift velocities (3.16) are shown in Fig. 3. The flow is strongly convergent as suggested by the R-model. Moreover, the x-component is attracted more strongly to the center than the y-component.

With $\zeta > 0$, the R-model predicts that ψ_r is negative. However, Fig. 4 demonstrates that ψ is positive for positive x, y and negative only in the rest of the domain. The temperature decreases the next day for strong northerlies, at least as far the rotational part is concerned. It tends to increase for southerlies except for rather warm days. There are clear deviations from the pattern predicted by the R-model. Southerlies tend to be damped more strongly than northerlies. The velocity potential exhibits just one center (Fig. 4b) and the distribution of χ is rather similar to the pdf in Fig. 2, as predicted by eq. (3.19). The velocity potential is stronger for southerlies than for northerlies, but the amplitude of these deviations is less than 10% of the total amplitude (Fig. 4c).

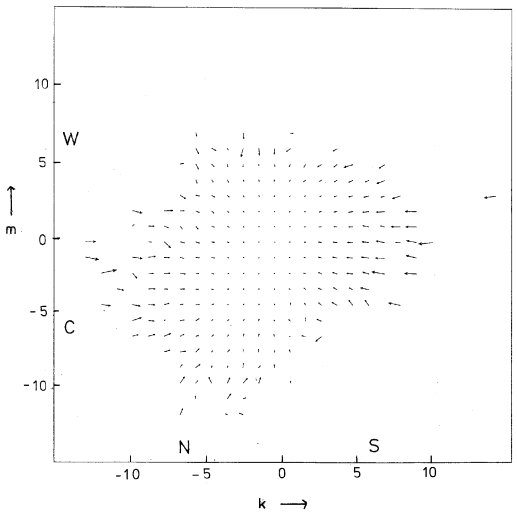

Fig. 3. Drift velocity $\boldsymbol{v}_d = (A_x, A_y)$ according to eq. (3.16). Maximum flow velocity 7.1. Grid as in Fig. 2.

The statistical significance tests reveal that high significance can be attributed to the negative deviations from the streamfunction of the R-model. On the other hand, the positive deviations do not pass the significance test. The deviations in the velocity potential are significant.

The diffusion coefficients D_{xx} and D_{yy} as evaluated according to eq. (3.16) are far from being uniform (Fig. 5). By and large the stochastic forcing is stronger near the boundaries, where deviations are larger, than in the center. Obviously D_{xx} in Fig. 5 contains features of grid scale. These are hardly reliable. Given a cloud of states centered in one grid cell it takes about three days with $D_{xx} \leqslant 0.1$ to smear it out over the adjacent boxes. The term D_{xx} represents mainly the impact of weather systems on the Icelandic surface pressure field. Anomalous heat fluxes in the boundary layer contribute to D_{yy}, as do stochastic variations of cloudiness and of other components in the radiation budget. Stochastic variations from dynamical processes are also important.

All in all, we find that the R-model portrays the basic features of the flow reasonably well. Decay by damping is clearly dominating. Deviations from the R-model point towards an asymmetry in the sense that situations with positive temperatures and southerlies are not just the reverse of those with negative temperatures and northerlies.

4.1.2. Grid-point meridional surface wind and relative topography; MZRT1. The east–west surface pressure difference M_s is the first variable x in this case, and y is the relative topography RT as given in the grid-point data *MZRT*. This example differs from the forgoing insofar as the pressure difference is evaluated over about double the distance and that the relative topography is a temperature variable of the free atmosphere.

Both variables are strongly lag-correlated (Table 3), somewhat in contrast to the foregoing case. The decay of pressure differences with increasing lag is not as rapid as in the set RD. This indicates that the Dalatangi–Reykjavik

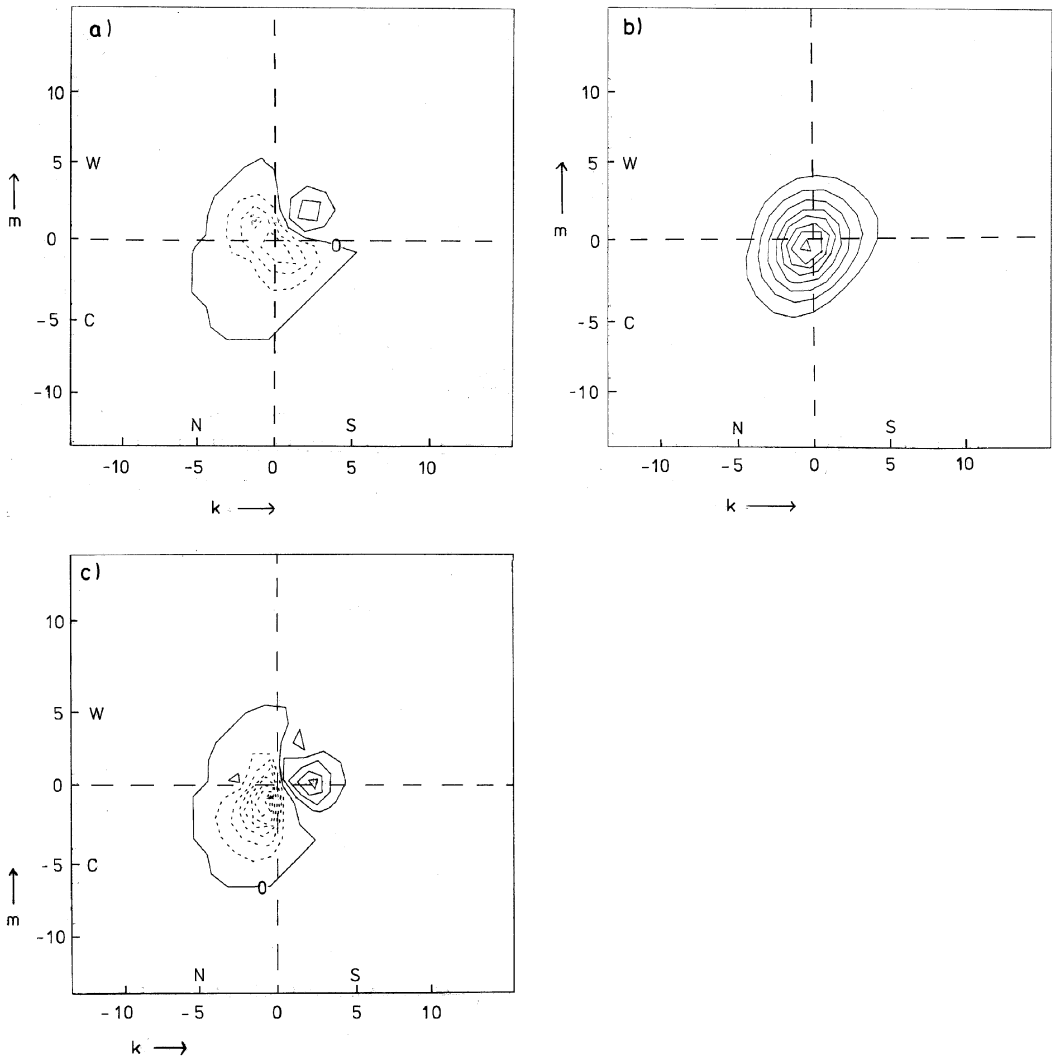

Fig. 4. Streamfunction (a) and velocity (b) potential of the pdf-weighed flow in RD as determined directly from eq. (3.18). Contour interval 0.002 for ψ and 0.08 for χ. Grid as in Fig. 2. Also given (c) is the difference between (b) and the velocity potential of the R-model; contour interval 0.008. Flow components $\bar{f}u$, $\bar{f}v$ etc. can be calculated from the figures by measuring distances in terms of the grid indices. Resulting units are in Dx/(6 h).

pressure difference is affected by local circulations which do not extend sufficiently far out from Iceland to have an influence on M_s. It comes, then, as a surprise that the relative topography anomalies decay slightly faster than do those of the mean temperature of Dalatangi and Reykjavik. This may reflect the stronger impact of the Atlantic ocean on the surface temperatures than on the relative topography. The circulation is weak and clockwise with $\zeta = -0.04$, and the eigenvalues are real. However, Fig. 6 reveals that the R-model is not really appropriate in this case. The rotational motion in the phase plane is clockwise for positive temperatures and southerlies and anticlockwise for negative temperature and northerlies. The R-model captures only a residual of both motions. Both strong southerlies and northerlies go with a tendency towards a cooling of the atmosphere

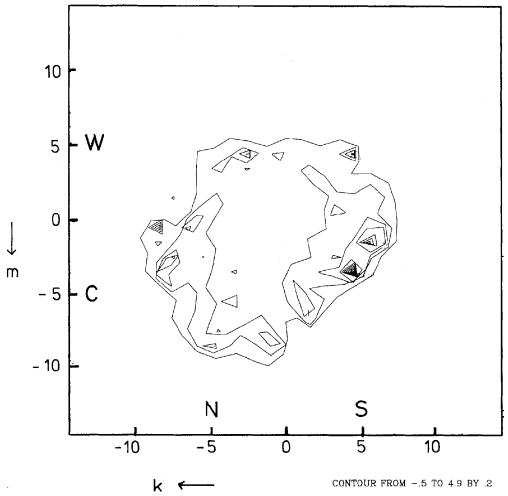

Fig. 5. Diffusion coefficients D_{xx}. The coefficients have been obtained by applying eq. (3.16) to the data of RD. Contour interval 0.02. Diffusion coefficients are computed only for grid boxes with at least two neighbour boxes.

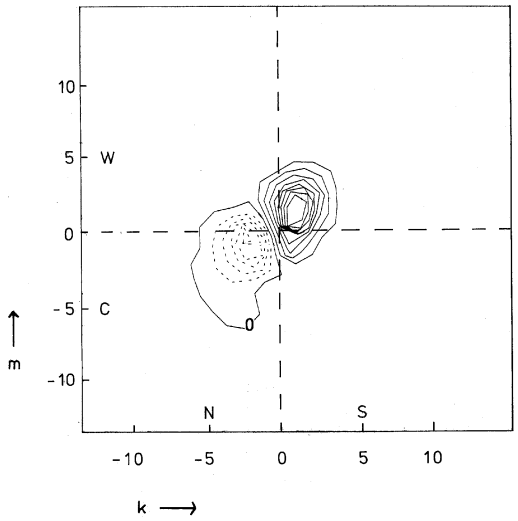

Fig. 6. Streamfunction ψ for the data set $MZRT1$; isoline interval 0.004.

above Iceland. The tests attach significance to both patches. The split of the streamfunction in an anticyclonic part for southerlies and a cyclonic part for northerlies has been found, albeit less pronounced, also in the foregoing case RD. The velocity potential (not shown) indicates that the damping of northerlies is weaker than that of southerlies.

4.1.3. Grid-point zonal surface wind and relative topography; MZRT2. We replace M_s by Z_s but keep the relative topography as second variable. The decay of anomalies of the zonal wind component is slightly slower than of those of the meridional component. The correlation coefficient of both variables is just 0.096. The vorticity is negative but small. The eigenvalues are real. Again there are deviations from the R-model. Warm situations decay faster than cold ones (not shown).

Table 3. *Autocorrelations AC and crosscorrelations CC for the variables $x = M_s$ and $y =$ relative topography; MZRT1.*

Lag (days)	AC_{xx}	AC_{yy}	AC_{xy}	AC_{yx}
0	1·00	1·00	0·57	0·57
1	0·57	0·69	0·41	0·37
2	0·33	0·45	0·27	0·26
3	0·24	0·33	0·22	0·18
4	0·17	0·25	0·18	0·13
5	0·11	0·18	0·15	0·09
6	0·01	0·14	0·12	0·06
7	0·07	0·10	0·09	0·04
8	0·06	0·07	0·08	0·03
9	0·06	0·05	0·07	0·02
10	0·03	0·04	0·04	0·01

4.1.4. Grid-point meridional and zonal surface winds; MZRT3. With $x = M_s$ and $y = Z_s$ we have a data pair with correlation coefficient 0.17, weak anticyclonic vorticity $\zeta = -0.13$ and real eigenvalues. The R-model is not fully satisfactory. Fig. 7 shows the streamfunction with negative values for southerlies and positive ones for northerlies. Thus there is a tendency towards a strengthening of the westerly component both for strong northerlies and southerlies. Again we find this difference of flow regimes with southerly winds and those with northerlies seen so clearly in $MZRT1$.

4.2. Upper level data

4.2.1. Meridional 500 hPa wind and relative topography; MZRT4. The combination $x = M_5$, $y = RT$ is obviously similar to the cases RD and $MZTR1$ considered above. The decay of the autocorrelation of the pressure difference at 500 hPa is slightly slower than at the surface and

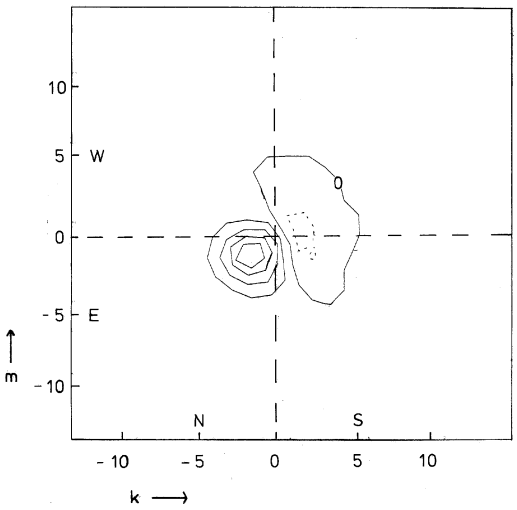

Fig. 7. Streamfunction in *MZRT*3; isoline interval 0.004.

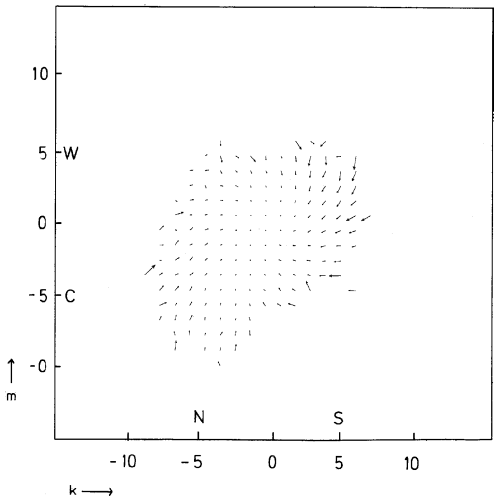

Fig. 8. Drift velocity in *MZRT*4; maximum velocity 7.2 Dx.

equals almost that of the relative topography. According to Table 4, southerlies tend to stay if there is a warm day with a positive deviation of the relative topography. The correlation coefficient is 0.4. The convergence is smaller than in RD (Table 1) and the vorticity is negative with $\zeta = -0.28$. The eigenvalues of the POP analysis are complex so that we obtain rotating eigenvectors. The period of their rotation is 51 days. This period is rather long. The drift velocity field as displayed in Fig. 8 shows distinct deviations from the symmetries assumed by the R-model. The flow pattern for $x, y > 0$ is strongly anticyclonic with rapid flow back to the center. Velocities are smaller and the curvature is less distinct for negative x, y. Correspondingly, the difference between the velocity potential and that given by the R-model is positive in the first quadrant and negative in the third one. This result is statistically significant. Altogether we find here again clear indication that southerly flow regimes are not the linear counterparts of northerly ones.

Table 4. *Autocorrelations AC and crosscorrelations CC for the dataset MZRT4 where* $x = M_5$ *and* $y = RT$

Lag (days)	AC_{xx}	AC_{yy}	CC_{xy}	CC_{yx}
0	1·00	1·00	0·39	0·39
1	0·66	0·69	0·15	0·38
2	0·42	0·45	0·12	0·24
3	0·30	0·33	0·14	0·14
4	0·22	0·25	0·13	0·08
5	0·15	0·18	0·11	0·05
6	0·12	0·14	0·09	0·03
7	0·09	0·10	0·07	0·01
8	0·07	0·07	0·06	0·00
9	0·06	0·05	0·04	0·00
10	0·04	0·04	0·02	0·00

The asymmetry of the flow field depicted in Fig. 8 must have an impact on the evolution of pdfs. As an example, we compare in Fig. 9 the pdfs obtained in two integrations of the FP equation (3.14) where all states were centered in the grid box (4, 4) in the first run and in (−4, −4) in the second. If the R-model were completely satisfactory the evolution of both "clouds" of states would be the same except for a change of sign of the coordinates. Obviously this is not the case. The cloud starting at (4, 4) moves more rapidly towards the center than the other one and is elongated towards large x, while the other cloud is more elliptic, with the larger axis being parallel to the y-axis. Table 5 documents these differences in a more quantitative way. It gives the mean position

$$(\bar{x}, \bar{y}) = \int f(x, y)\, \mathrm{d}x\, \mathrm{d}y$$

as well as the mean deviations therefrom as a function of time for both clouds. The differences

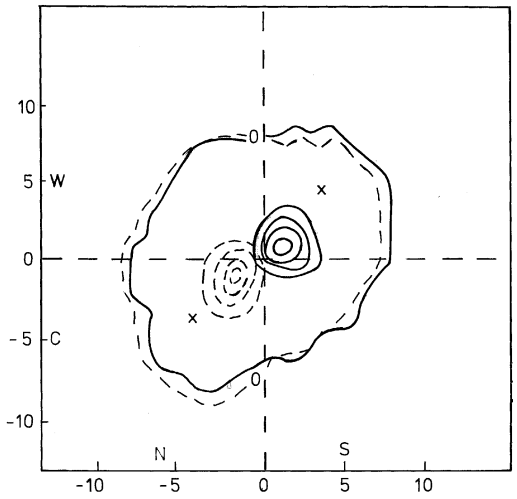

Fig. 9. Probability density distribution after three days as obtained in integrations of the FP equation (3.14) with initial states centered in box (4, 4) (solid) and in box (−4, −4) (dashed). Flow data from *MZRT*4; isoline interval 0.02. The crosses give the position of the center of the initial distribution.

of the mean position are largest on the second day. The second cloud has a larger spread in the *y*-direction, as can be seen directly from Fig. 9. This example documents clearly that states with strong southerlies and high temperature have less persistence than those with opposite sign. Moreover, they approach the origin along a mean trajectory with pronounced anticyclonic curvature, while the "cold" cloud moves almost along a straight line towards the center.

Table 5. *Intercomparison of two integrations of the FP equation (3.14) for the data set MZRT4. In the first integration (first entry) all initial states are concentrated in the grid box (4, 4), in the second (second entry) the starting box is (−4, −4). The mean position (x, y) of both clouds are given, as well as the related variances for the first three days of the runs*

Day	$-\bar{x}$	$-\bar{y}$	$\overline{(x-\bar{x})^2}$	$\overline{(y-\bar{y})^2}$
1	3.26/−2.89	2.15/−2.51	0.95/0.91	1.55/1.44
2	1.22/−2.23	1.20/−1.78	1.23/1.38	1.67/2.22
3	1.73/−1.68	0.77/−1.26	1.31/1.56	1.58/2.51

4.2.2. Meridional and zonal 500 hPa winds; MZRT5. The relative topography is replaced by the meridional pressure differences across Iceland in *MZRT*5, so that x (y) corresponds with the meridional (zonal) geostrophic wind. These variables are uncorrelated and the observed pdf is almost circular. The R-model indicates that the flow in the phase plane is weakly anticyclonic but the eigenvalues are real. The R-model appears to capture the overall situation quite well, but again northerlies decay slower than southerlies.

4.2.3. Zonal 500 hPa winds and relative topography; MZRT6. This case fits the general impression that real eigenvalues must be expected whenever a zonal wind component is part of the pair. The R-model is satisfactory.

4.3. Combination of upper-air and surface data

4.3.1. Meridional winds at 500 and 1000 hPa; MZRT7. In *MZRT*7 we look at the pair $x = M_5$, $y = M_s$. This way we relate the meridional geostrophic wind components at the upper level and at the surface. As revealed by Tables 2 and 4 the surface winds decay only slightly faster than the 500 hPa meridional wind deviations. There is a strong and positive correlation with wind variations at 500 hPa lagging those at the surface (see also Hartmann, 1974).

The coefficients of the R-model (Table 1) reveal a strong clockwise rotation with $\zeta = -0.48$. The eigenvalues are complex with a rotation period of 33 days, which is still relatively long compared to cyclone timescales.

The pdf is elliptic (not shown) with a strong tilt. This reflects the high correlation of both variables. The velocities in the phase plane reveal clearly that there is clockwise circulation (Fig. 10) reflecting the tilt of the axis. The related streamfunction (Fig. 11) shows a pronounced maximum, but there is a small lobe of negative values to the right. This indicates that the 500 hPa winds do not lag the surface winds if both components are positive and relatively strong.

4.3.2. Zonal winds at 500 hPa and at 1000 hPa; MZRT8. The pair to be analyzed here are $x = Z_5$, $y = Z_s$. These variables are strongly correlated with $CC_{xy}(0) = 0.73$ and there is anticyclonic rotation, i.e. upper level zonal winds lag the surface winds. As can be seen from Table 3, $\beta_2 = 0$, i.e. the

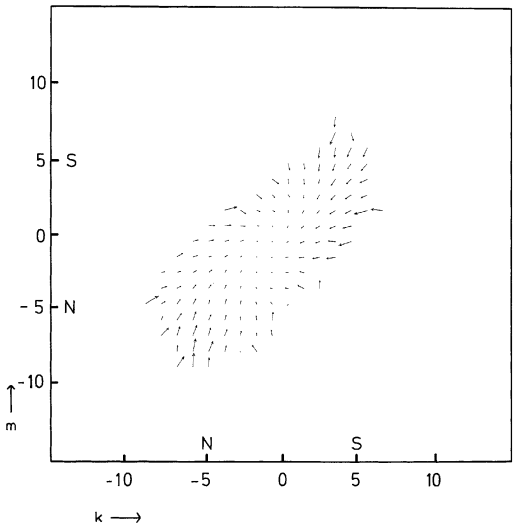

Fig. 10. Drift velocity in *MZRT*7; maximum velocity 6.

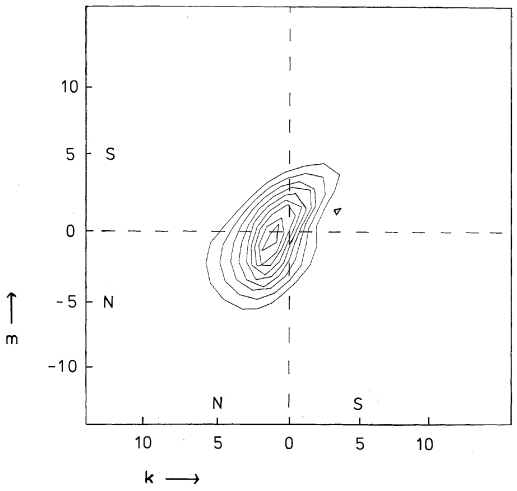

Fig. 11. Stream function in *MZRT*7; contour interval 0.04.

actual value of Z_{1000} has little impact on the tendency of Z_{500}. It is, therefore, not surprising that the eigenvalues are real in this case.

The velocity pattern shows a distinct asymmetry with westerlies decaying more rapidly than easterlies, so the R-model is not really satisfactory in this case.

5. Conclusions and discussion

The following conclusions can be drawn:

(1) The regressive model provides a good estimate of the drift velocity in many cases. Exceptions are the data combinations M_s/RT (*MZRT*1) and M_s/Z_s (*MZRT*3), where the R-model yields even qualitatively incorrect flow patterns. Moreover, there are many cases with pronounced deviations from the linearity assumed by the R-model. These deviations have been found using the FP equation.

(2) The decay times are always shorter than the times linked to the rotation. The return to normal is always faster than internal oscillations.

(3) Eigenvalues are invariably real if zonal winds are part of the data pair. Complex eigenvalues are found in many other cases. However, the related oscillation periods are rather long when compared to cyclone timescales.

(4) Temperatures and meridional wind components are always positively correlated as expected. However, the related drift velocity is hardly linear as assumed by the R-model. There is a clear separation of southerly and northerly flow regimes. The streamfunction tends to be positive for southerlies but negative for northerlies. This separation extends to the combination of meridional and zonal winds at the surface but not at 500 hPa.

(5) There is strong coupling of the respective wind components at the ground and at 500 hPa, with the upperlevel winds lagging those at the ground.

It has been demonstrated in the foregoing that the evaluation of velocities and related diffusion coefficients in the phase plane of the parameters yields additional insight beyond what follows from the correlations and from the autoregressive POP model. However, the demands on record length are clearly larger for the more advanced methods to yield significant results. Correspondingly, some of the deviations of FP results form the R-model turned out not to be significant.

Although the annual cycle has been removed it should be stressed that one would have to stratify the data at least by seasons in order to remove additional aspects of the annual cycle. For example, the variability in winter is larger than in summer. This effect is, of course, not removed here.

Given the length of the data record one might expect to see some signatures of climatic change in

the Icelandic region. However, the analysis procedures used here assume statistical stationarity and are, therefore, not suited to detect such changes.

6. Appendix

The first step in the design of a significance test is to reduce the number of variables (e.g. von Storch and Zwiers, 1999). It would be hopeless to test the significance of all details of Fig. 4c, for example. It appears sufficient to attribute significance to the gross feature of this figure. To that end we introduce as test variables the sums

$$S_\varphi^+ = \frac{1}{2}\sum_{km}(\varphi + |\varphi|),$$
$$S_\varphi^- = \frac{1}{2}\sum_{km}(\varphi - |\varphi|) \qquad \text{(A.1)}$$

over all points with positive and with negative grid point values of the variable φ,

Difference maps like Fig. 4c pose a significance problem because of the finite length of the data. We use a Monte Carlo approach to resolve this issue. Given, for example, a velocity field $\boldsymbol{v}_\mathrm{d}$ from the FP equation and the corresponding set of α_i, β_i from the R-model a large number of runs of length N is performed with the R-model. The forcing is white in each run and has the standard deviation known from the data. Each run produces fields ψ_r, χ_r so that the related test variables $S_{\psi_\mathrm{r}}^+$, $S_{\psi_\mathrm{r}}^-$, $S_{\chi_\mathrm{r}}^+$, $S_{\chi_\mathrm{r}}^-$ can be determined. This way obtain for each run four numbers which characterize the basic structure of the velocity fields. Next we compute the mean and standard deviation of all four series of numbers. The $S_{\psi_\mathrm{r}}^+$ have a normal distribution. Given all this information the statistical significance of the difference map $\delta = \psi_\mathrm{d} - \psi_\mathrm{r}$ is tested by comparing, for example, the mean of S_δ^+ to the standard deviation of the $S_{\psi_\mathrm{r}}^+$.

7. Acknowledgements

Most of this work has been completed while one of us (J.E.) spent his sabbatical at Vedurstofa Islands. He would like to thank this institution for its hospitality and, in particular, H. Olafsson for continuous help and support. We are grateful to Tomas Jóhansson for useful comments. The detailed comments by Klaus Fraedrich and by an anonymous referee helped to improve the paper.

REFERENCES

Barnston, A. and Livezey, R. 1987. Classification, seasonality and persistence of the low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.* **115**, 1083–1126.

Demaree G. and Nicolis, C. 1990. Onset of Sahelian drought viewed as a fluctuation induced transition. *Q. J. R. Meteorol. Soc.* **116**, 221–238.

Ditlevsen, P. 1999. Observation of alpha-stable noise induced in millenial climate changes from an ice-core record. *Geophys. Res. Lett.* **26**, 1441–1444.

Egger, J. 2001. Master equations for climatic parameter sets. *Climate Dynam.* **17**; in press.

Einarsson, M. 1984. The climate of Iceland. In: *World survey of climatology, 15. The Climate of the Oceans* (ed. H. van Loon). Elsevier, 673–679.

Einarsson, M. 1991. Temperature conditions in Iceland 1901–1990. *Jökull* **41**, 1–19.

Eythorsson, J. and Sigtryggson, H. 1971. The climate and weather of Iceland. In: *Zoology of Iceland*, Vol. 1,3. Munksgaard, 62 pp.

Friedrich, R. and Peinke, J. 1997. Description of a turbulent cascade by a Fokker–Planck equation. *Phys. Res. Lett.* **78**, 863–866.

Gardarsson, H. 1999. *Saga Vedurstofa Islands* (History of the Icelandic Meteorological Office). Mal og mynd, 417 pp.

Gardiner, C. 1983. *Handbook of stochastic methods.* Springer, 1–442.

Hartmann, D. 1974. Time spectral analysis of mid-latitude disturbances. *Mon. Wea. Rev.* **102**, 348–362.

Jónsson, T. 1994. Precipitation in Iceland 1857–1992. In: *Climate variations in Europe* (ed. R. Heino). Academy of Finland, 183–188.

Jónsson, T. 1997. Regional climate and simple circulation patterns. Vedurstofa Islands Report; VI-G97022-UR17. Available upon request from Vedurstofa islands, Bustadavegur 9, 150 Reykjavik, Iceland.

Siegert, S., Friedrich, R. and Peinke, J. 1998. Analysis of data sets of stochastic systems. *Phys. Lett. A* **243**, 275–280.

Schnur, R., Schmitz, G., Grieger, N. and Storch, H. V. 1993. Normal modes of the atmosphere as estimated by principal oscillation patterns and derived from quasigeostrophic theory. *J. Atmos. Sci.* **50**, 2386–2400.

Serreze, M., Carse, F. and Barry, R. 1997. Icelandic low cyclone activity: climatological features, linkage with the NAO and relationship with recent changes in the Northern Hemisphere circulation. *J. Climate* **10**, 457–464.

von Storch, H. and Zwiers, F. 1999. *Statistical analysis in climate research.* Cambridge University Press.