Contr. Atmos. Phys., February 1999, p. 105–111
0005-8173/99/01 0105-07 $ 3.00/0

Identification of Regional Persistent Patterns Through Principal Prediction Patterns

MATTHIAS DORN AND HANS VON STORCH

Meteorologisches Institut der Universität Hamburg, Germany

(Manuscript received January 05, 1998; accepted June 05, 1998)

Abstract

Most persistent atmospheric circulation patterns in terms of air pressure distributions are derived for the N Atlantic / NW European region. They are identified through Principal Prediction Pattern Analysis and through Principal Oscillation Pattern Analysis. With the help of these patterns the "degree of persistence" in a given pressure map can be determined. The larger the initial degree of persistence, the better is the skill of the persistence forecast.

1 Background

When we began this study, we wanted to construct an empirical simple and efficient forecast scheme, which should be superior to persistence. The basic idea, originally proposed by K. Hasselmann in the mid 1980s in an unpublished manuscript, was to identify predictand/predictor patterns (namely *Principal Prediction Patterns*, PPP) with maximum correlation. In this way we were after the "most predictable" patterns. The same approach has been used by Barnston (1994) in his efforts predicting monthly and seasonal means.

In the present case daily data were considered, and it turned out that the predictand patterns are very similar to the predictor patterns, indicating that the "most predictable" patterns represent a persistent situation – and, indeed, the forecast based on the scheme is not better than (damped) persistence. So, instead of experimenting with a simple forecast superior to persistence, we actually tested a scheme for the determination of the degree of persistence of a given regional circulation pattern.

After we found out that we had not dealt with a forecast technique superior to persistence but rather a technique for the determination of persistent states, we asked ourselves how we could determine persistent patterns in other ways. Following an advice by Grant Branstator, we used real Principal Oscillation Patterns (POPs) to this end, and found that the "most predictable" PPPs are indeed very similar to the longest-memory real POPs.

The analytical tools, PPPs and POPs, are linear techniques so that the degree of persistence depends only on the prevailing patterns and the amplitude of the patterns, but not on the sign of the patterns. Thus, they are not tools for the identification of *blockings* as these features are always high-pressure systems.

The paper is organized as follows. The empirical techniques, PPP and POPs, are briefly described in Section 2, the data (approximately hundred years of daily North Atlantic/NW Europe air-pressure maps) are introduced in Section 3, where also the Principal Prediction Patterns are presented and compared to POPs. In Subsection 3.3 the predictive skill, in an overall sense as well as sorted after the "initial degree of persistence", is presented. The paper is concluded with a discussion in Section 4.

2 Diagnostic Tools

The diagnostic tools employed in this study are linear pattern-oriented techniques, i.e., a field such as the air pressure map $\vec{\mathbf{S}}_t$ at a given time t in a given region is expanded into a short series with some *patterns* $\vec{p}^{\,k}$:

$$\vec{S}_t = \sum_{k=1}^{K} \alpha_k(t)\vec{p}^{\,k} + \text{residual} \tag{2.1}$$

with K being considerably smaller than the dimension of $\vec{S}$ (which is the number of grid points if $\vec{S}$ is a gridded field). One advantage of (2.1) is that the often high-dimensional vector $\vec{S}$ is reduced to a short vector $\vec{\alpha} = (\alpha_1 \dots \alpha_K)$. Typically, the low-index patterns are of larger scale and the residual of small scale. The low-indexed patterns are most efficient in representing the variance of $\vec{S}$ and, in our set-up, their coefficients should be best predicted. The residual accounts for little variance of $\vec{S}$ and is hardly predictable. If $\hat{\alpha}_k(t+\tau)$ is a prediction of the coefficient $\alpha_k(t+\tau)$ of $\vec{S}_{t+\tau}$, then

$$\hat{\vec{S}}_{t+\tau} = \sum_{k=1}^{K} \hat{\alpha}_k(t+\tau)\vec{p}^{\,k} \tag{2.2}$$

is a forecast of the field $\vec{S}_{t+\tau}$.

Statistical analysis offers algorithm for the determination of sets of patterns which satisfy certain favorable conditions (for an overview, see von Storch and Zwiers, 1998). The following two subsection briefly introduce two of these techniques, namely PPP and POPs.

2.1 Principal Prediction Patterns (PPPs)

Principal Prediction Patterns are pairs of patterns for expanding the predictor (i.e, the field $\vec{S}$ at time t) and the predictand (the same field $\vec{S}$ but at τ days in advance):

$$\begin{aligned} \vec{S}_t &= \sum_{k=1}^{K} \alpha_k^0(t)\vec{p}_0^{\,k} + \text{residual} \\ \vec{S}_{t+\tau} &= \sum_{k=1}^{K} \alpha_k^\tau(t)\vec{p}_\tau^{\,k} + \text{residual} \end{aligned} \tag{2.3}$$

where the additional index 0 marks an expansion of the predictor and the index τ the expansion of the predictand. The patterns are determined such that the correlation ρ_k between α_k^0 and α_k^τ are maximum for $k=1$, second maximum for $k=2$ and so forth. Also, the coefficients are orthogonal. This set of conditions is satisfied by Canonical Correlation Analysis. The patterns $\vec{p}_0^{\,k}$ and $\vec{p}_\tau^{\,k}$ are normalized so that $\mathrm{VAR}(\alpha_k^0) = 1$ and

$$\mathrm{VAR}(\alpha_k^\tau) = \rho_k(\tau)^{-2} \tag{2.4}$$

With this normalisation, the optimal prediction (marked by a hat) of α_k^τ is simply

$$\hat{\alpha}_k^\tau = \alpha_k^0 \tag{2.5}$$

The pattern $\vec{p}_0^{\,k}$ and $\vec{p}_\tau^{\,k}$ occur typically together. When $\vec{S}_t = \vec{p}_0^{\,k}$ then, on average, $\vec{S}_{t+\tau} = \vec{p}_\tau^{\,k}$. The argument is linear, so that the more general statement is valid as well: τ days after $\vec{S}_t = \beta\vec{p}_0^{\,k}$ we have on average $\vec{S}_{t+\tau} = \beta\vec{p}_\tau^{\,k}$ for any β, including $\beta = -1$.

Thus, $\vec{p}_0^{\,k}$ is a predictor for $\vec{p}_\tau^{\,k}$. By construction it is even the best linear predictor – therefore the name *Principal Prediction Patterns* (PPP) was coined by K. Hasselmann for the pairs of patterns $(\vec{p}_0^{\,k}, \vec{p}_\tau^{\,k})$

A-priori nothing is known about the combinations of patterns which may emerge. For instance, $\vec{p}_\tau^{\,k}$ may be like $\vec{p}_0^{\,k}$ but shifted by some degrees to the east – in that case the pair of pattern would describe an eastward propagation feature. However, such combinations hardly emerge from our analysis. Instead, $\vec{p}_\tau^{\,k}$ is mostly similar to the predictor pattern but reduced in magnitude: $\vec{p}_\tau^{\,k} \sim \rho_k(\tau)\vec{p}_0^{\,k}$ (see below). Thus, in the present application of CCA, the correlation coefficient may be understood as a kind of "memory rate" μ (i.e., the ratio of an anomaly after τ days to an initial anomaly one; $-1/log(\mu)$ is a damping time).

The complete PPP forecast is given by

$$\hat{\vec{S}}_{t+\tau} = \sum_{k=1}^{K} \alpha_k^0 \vec{p}_\tau^{\,k} \tag{2.6}$$

In the present study, the CCA is done with an a-priori EOF truncation which enhances the robustness of the estimated quantities because of a reduction of degrees of freedom. The mathematics are worked out by Barnett and Preisendorfer (1987) and von Storch and Zwiers (1999).

2.2 Principal Oscillation Patterns (POPs)

The Principal Oscillation Patterns (POPS; for an overview, see von Storch et al., 1995) are eigen solutions of the linear dynamical equation

$$\vec{S}_{t+1} = \mathcal{A}\vec{S}_t + \text{noise} \tag{2.7}$$

with the matrix $\mathcal{A}$ fitted to the data: $\mathcal{A} = \Sigma_1\Sigma_0^{-1}$, where Σ_1 is the covariance matrix of $\vec{S}_t$ and $\vec{S}_{t+1}$. The *noise* is assumed to be independent of $\vec{S}_t$. As in the case of PPP, the fit of the POP model (2.7) is made in a phase space, spanned by the most important EOFs.

The matrix $\mathcal{A}$ may be seen as the operator which maps the circulation $\vec{S}_t$ at time t on the circulation $\vec{S}_{t+1}$ one day later.

The matrix $\mathcal{A}$ is not Hermitian and may have real and complex eigenvectors and eigenvalues. The modulus of the eigenvalues is less than 1. The complex patterns are most useful for the identification and description of spatially propagating features, whereas the real patterns describe standing features. If $\vec{e}^j$ is a real eigenvector with the real eigenvalue λ_j then by definition $\mathcal{A}\vec{e}^j = \lambda_j \vec{e}^j$.

Thus, if $\vec{S}_t = \vec{e}^k$ then the best forecast is $\hat{\vec{S}}_{t+\tau} = \mathcal{A}^\tau \vec{S}_t = \mathcal{A}^\tau \vec{e}^k = \lambda_k^\tau \vec{e}^k = \lambda_k^\tau \vec{S}_t$. Thus, this pattern is a persistent pattern, with a memory rate of λ_k^τ in τ days.

3 Results

3.1 Data

The statistical analyses are done for daily air pressure in the N Atlantic and NW Europe (25°W - 20°E, 40° - 60°N). Gridded analysis have been collected by NCAR from 1899 until today. Even though there may be some inhomogeneities the data are considered mostly "clean" in the selected region (Trenberth and Paolino, 1980). Only winter (DJF) cases are considered, with lags up to 5 days. All data are "centered", i.e., a long-term mean of 30 years is subtracted at each point.

The CCA required for the determination of the PPPs was done with data from 1958–1988, and the skill was determined from all data 1899–1991. All analyses were done for $\tau = 1 \dots 5$, but only results for $\tau = 3$ are shown as these appear as representative. The air pressure field $\vec{S}$ is truncated by the first eight EOFs, explaining 96% of the total variance.

The POP model (2.7) was fitted with data from 1958–1988 as well.

As reference forecast, *damped persistence* is used, i.e.,

$$\vec{S}^p_{t+\tau}(x) = d(x;\tau)\vec{S}_t(x) \tag{3.1}$$

at each grid point x. The spatially varying number $d(x,\tau)$ is the average memory rate at the grid point x calculated for a time lag of τ, which is

$$d(x,\tau) = \frac{\mathrm{COV}\left(\vec{S}_{t+\tau}(x), \vec{S}_t(x)\right)}{\mathrm{VAR}\left(\vec{S}_t(x)\right)} \tag{3.2}$$

Table 1: Correlations $\rho_k(\tau)$ for the first four PPPs and lags $\tau = 1 \dots 5$.

k	$\rho_k(\tau)$				
$\tau =$	1	2	3	4	5
1	0.91	0.77	0.66	0.59	0.52
2	0.91	0.73	0.58	0.45	0.36
3	0.85	0.62	0.46	0.36	0.27
4	0.77	0.51	0.37	0.29	0.21

3.2 Persistent Patterns

The first four PPPs for $\tau = 3$ days are shown in Figure 1, and the correlations $\rho_k(\tau)$ for all lags $\tau = 1 \dots 5$ in Table 1. The patterns are almost the same for different lags, but the predictand patterns are smaller for longer lead times (cf. (2.4)). The predictor and predictand patterns are rather similar (but not identical). A comparison of the patterns yields indeed the approximate relationship $\vec{p}^k_\tau \sim \rho_k(\tau)\vec{p}^k_0$.

Examples for predictions prepared for two significantly different regimes are given in Figure 2, displaying both the predictor and the predictand.

The POPs are similar to the PPPs (see Figure 3). However, the memory rates are considerably smaller than those obtained in the PPP analysis. The difference between the PPPs and the POPs is that the POPs are strictly persistent patterns, whereas the PPPs allow for minor modifications in time; strictly speaking, the PPPs are not really persistent patterns but quasi-persistent patterns. We conclude that the PPPs shown in Figure 1 are not only the most predictable but also the most (quasi-) persistent patterns.

For the evaluation of the skill of the PPP forecast (2.2), we calculated from the nearly hundred year record of regional weather maps the anomaly correlation coefficients for the PPP forecasts and, as a reference, of the damped persistence for lead times up to 5 days (Figure 4). It turns out that the anomaly correlation coefficient is almost identical to that found for damped persistence. For $\tau = 1$ day its mean value is about 0.75, and it decreases to about 0.3 for $\tau = 5$ days.

3.3 Predicting Skill

As "degree of persistence" we use the proportion of variance of the predictor field $\vec{S}_t$ represented by the first four PPPs, i.e.,

$$\varepsilon = 1 - \frac{\mathrm{VAR}\left(\vec{S}_t - \sum_{k=1}^4 \alpha_k^0 \vec{p}_0^k\right)}{\mathrm{VAR}\left(\vec{S}_t\right)} \tag{3.3}$$

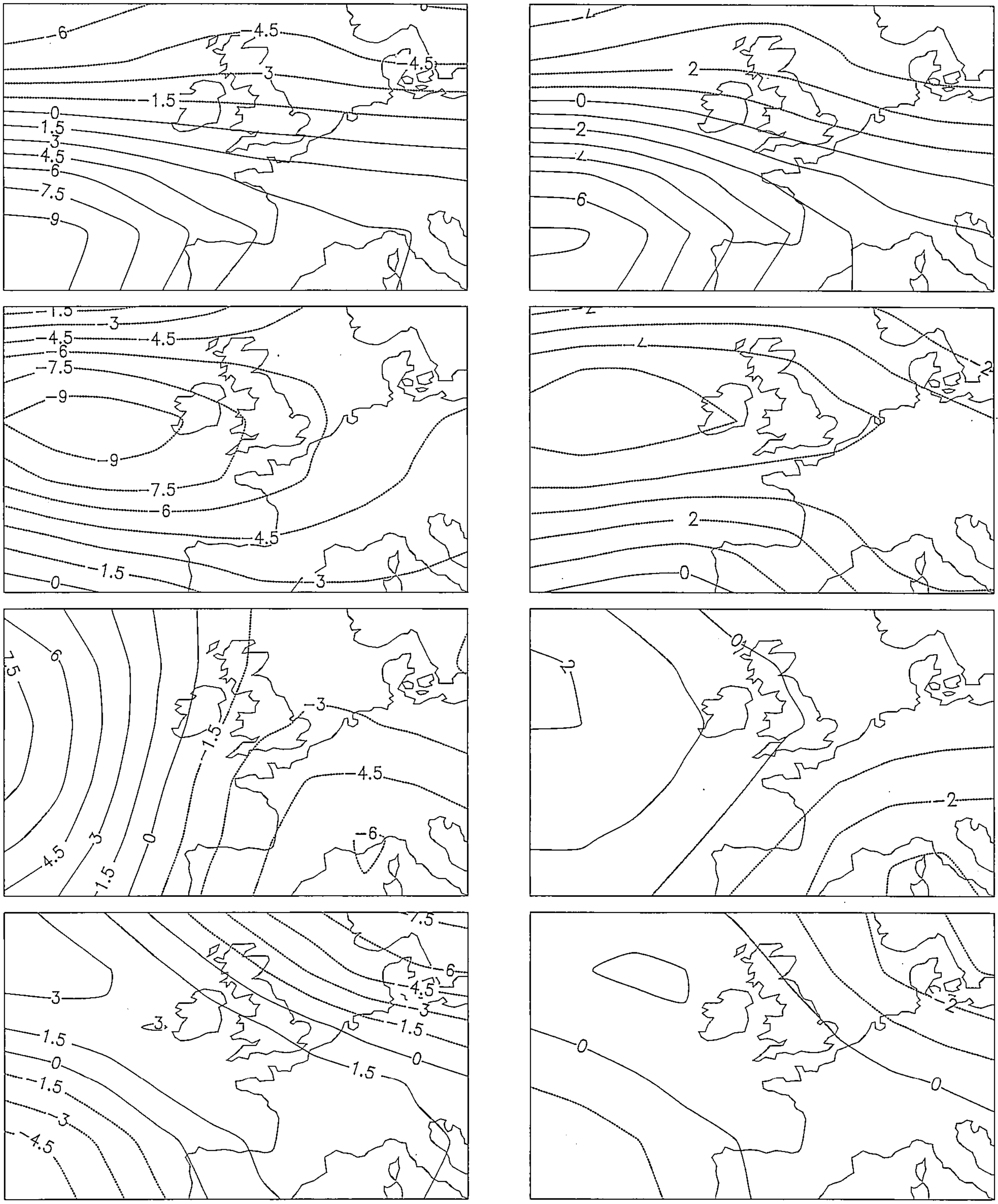

Figure 1: First four pairs of Principal Prediction Patterns $\vec{p}_0^{\,k}$ (left) and $\vec{p}_\tau^{\,k}$ (right) for $\tau = 3$ *days*. The correlations amount to 0.66, 0.58, 0.46, 0.37.

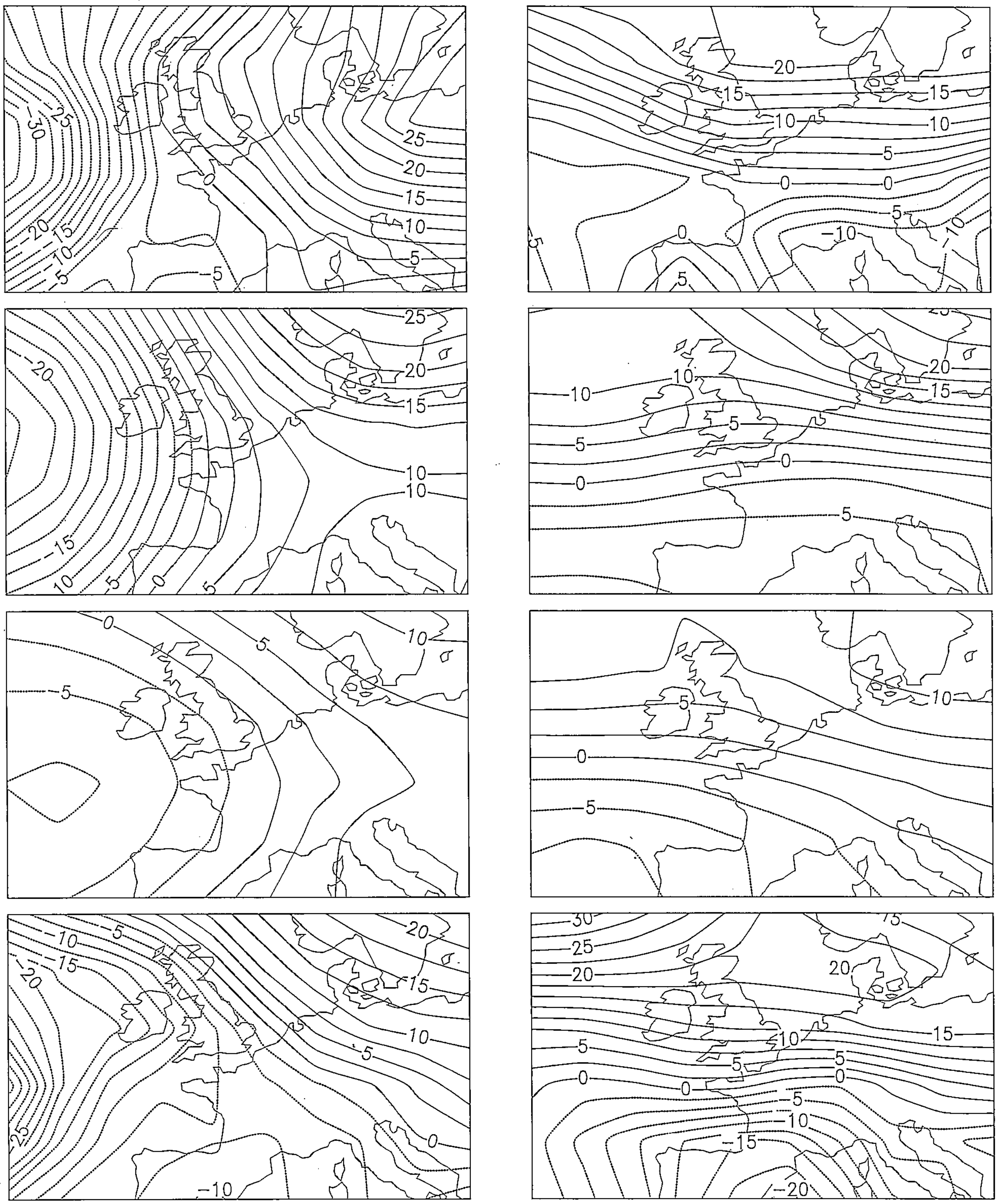

Figure 2: Two examples of successful 3-day forecasts prepared by the PPP method (2.6) initiated at two widely different states (Jan. 04th, 1947, left, and Feb. 14th, 1947, right).

a) initial state $\vec{\mathbf{S}}_t$.

b) truncated initial state $\sum_{k=1}^{4} \alpha_k^0 \vec{p}_0^{\,k}$

c) forecast $\hat{\vec{\mathbf{S}}}_{t+3} = \sum_{k=1}^{4} \alpha_k^0 \vec{p}_3^{\,k}$

d) observed state $\vec{\mathbf{S}}_{t+3}$

All maps represent anomalies, i.e. deviations from the long-term mean. Units: hPa.

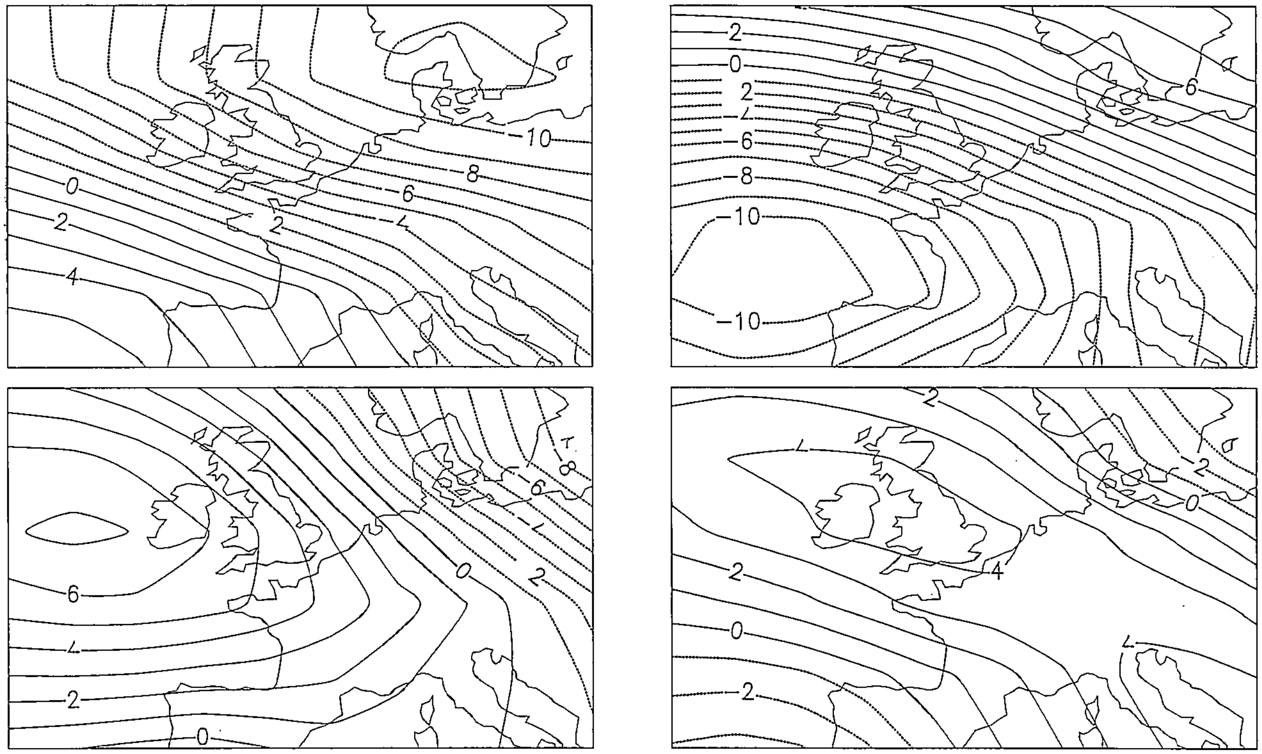

Figure 3: First (longest memory) four POPs with eigenvalues $\lambda_1 = 0.45, \lambda_2 = 0.43, \lambda_3 = 0.38, \lambda_4 = 0.37$ (corresponding to a 3-day memory rates of 0.09, 0.07, 0.05, 0.05).

In principle, $\varepsilon \in [-\infty, +1]$, but practically $\varepsilon \in [0,1[$. Seven intervals of degree of persistence were formed, the relative number of cases per interval is also listed in the following table:

class (i)	interval	frequency (%)
1	[0.0, 0.4[	7.0
2	[0.4, 0.5[	6.8
3	[0.5, 0.6[	11.0
4	[0.6, 0.7[	16.8
5	[0.7, 0.8[	25.1
6	[0.8, 0.9[	25.6
7	[0.9, 1.0]	7.7

In about 8% of all winter days, the regional circulation has a degree of persistence of more than 0.9, and in almost 60% of the time, the degree of persistence is larger than 0.7.

The anomaly correlations and root mean square errors, together with 25%–75% confidence intervals, calculated separately for the seven classes are shown for the lag $\tau = 3$ days in Figure 5. Again, the damped persistence and the PPP forecast exhibit about the same scores. The significant aspect is, however, that the skill is becoming considerably larger when the degree of persistence is larger. For the 8% of cases with minimum degree of persistence (≤ 0.4) the average correlation anomaly correlation coefficient is only about 0.2. For class $i = 4$, the score is about 0.4, but for 8% of cases in the top class $i = 7$ the skill is almost 0.6 (as compared to 0.4 averaged over all cases, see Figure 4.)

4 Conclusions and Comments

We have tested a simple forecast scheme based on Canonical Correlation Analysis. The basic idea is to determine pairs of predictor/predictand patterns of the same variable, whoses coefficients are maximum correlated in time. These patterns are named *Principal Prediction Patterns.*

It turned out that the Principal Prediction Patterns represent mostly persistent patterns, and it is therefore not surprising to find that their forecast skill compares to that of straight forward damped persistence. Therefore, the PPP technique can not be considered a useful forecast scheme *per se*, but its merits comes from its ability to determine the *persistent patterns*. The knowledge of these persistent patterns allows for a skillful assessment whether a given atmospheric state likely will be persistent or not.

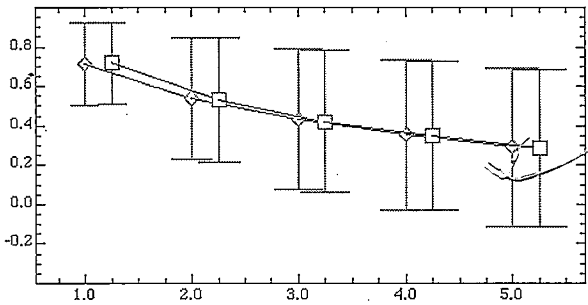

Figure 4: Means of anomaly correlation coefficients of the PPP forecast (diamonds; equation (2.6)) and of the damped persistence forecast (squares; equation (3.1)) for time lags $\tau = 1$ to 5. The vertical bars represent 25%-75% confidence limits, derived from forecasts of all winter maps from 1899 until 1991.

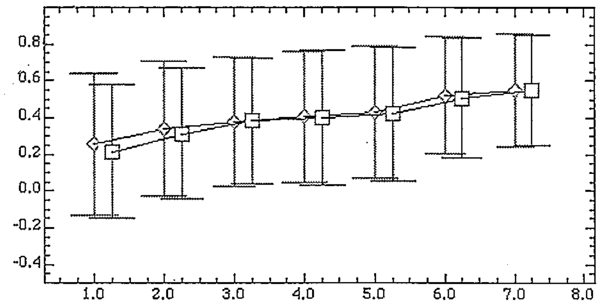

Figure 5: Means of anomaly correlation coefficients of the PPP forecast for 3 days (diamonds; equation (2.6)) and of the damped persistence forecast (squares; equation (3.1)) for the seven classes with different degrees of persistence. The vertical bars represent 25%–75% confidence limits, derived from forecasts of all winter maps from 1899 until 1991.

When a new advanced forecast scheme is tested then it is usually compared with the performance of a simpler scheme, for instance with damped persistence (Livezey, 1995). It is suggested to refine such a comparison by sorting the initial states according to their degrees of persistence and thereby by the expected skill of damped persistence.

We have repeated our analysis for other (larger) areas and found the results mostly insensitive to the choice of the area.

Acknowledgments

With most of the work presented in this paper, one of the authors (M.D.) earned under the supervision of the other author (HvS) his Diploma degree from the University of Hamburg in 1997. Useful discussions and helpful advice were offered by Grant Branstator, Gerhard Schmitz and Klaus Hasselmann.

References

Barnett T.P. and Preisendorfer R., 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperature determined by canonical correlation analysis. Mon. Wea. Rev. **115**, 1825–1850.

Barnston A., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere. J. Climate **7** 1513–1564.

Livezey R.E., 1995: The evaluation of forecasts. In: v. Storch H. and Navarra A. (eds), Analysis of climate variability: Applications of statistical techniques, Springer Verlag, 177–196.

Trenberth K.E. and Paolino D.A., 1980: The Northern hemisphere sea-level pressure data set: Trends, errors and discontinuities. Mon. Wea. Rev. **108**, 855–872.

von Storch H., Bürger G., Schnur R., and von Storch J., 1995: Principal oscillation pattern: A review. J. Climate **8**, 377–400.

von Storch H. and Zwiers F.W., 1999: Statistical analysis in climate research. Cambridge University Press (in press).