Interannual variability of Central European mean temperature in January–February and its relation to large-scale circulation

Peter C. Werner\textsuperscript{1}, Hans von Storch\textsuperscript{2}

\textsuperscript{1}Potsdam-Institut für Klimafolgenforschung, D-11412 Potsdam, Germany
\textsuperscript{2}Max-Planck-Institut für Meteorologie, D-20146 Hamburg, Germany

ABSTRACT: The Central European temperature distribution field, as given by 11 stations [Fanø, Hamburg, Potsdam, Jena, Frankfurt, Uccle, Hohenpeißenberg, Praha (Prague), Wien (Vienna), Zürich and Genève (Geneva)], is analyzed with respect to its year-to-year variability. January–February (JF) average temperatures are considered for the interval 1901–1980. An Empirical Orthogonal Function (EOF) analysis reveals that the JF temperature variability is almost entirely controlled by one entirely positive EOF. The second EOF represents only 7\% of the total variance and describes a north-south gradient. The time coefficient of the first EOF is almost stationary whereas the second pattern describes a slight downward trend at the northern stations and a slight upward trend at the southern stations. The relationship of the temperature field to large-scale circulation, represented by the North Atlantic/European sea-level pressure (SLP) field, is investigated by means of a Canonical Correlation Analysis (CCA). Two CCA pairs are identified which account for most of the temperature year-to-year variance and which suggest plausible mechanisms. The CCA pairs fail, however, to consistently link the long-term temperature trends to changes in the large-scale circulation. In the output of a 100 yr run with a coupled atmosphere-ocean model (ECHAM1/LSG), the same CCA pairs are found, but the strength of the link between Central European temperature and North Atlantic SLP is markedly weaker than in the observed data.

INTRODUCTION

The statistics of regional distributions of climate variables are regarded as important for a number of reasons. Historically, this type of information has been of value to users in agriculture and hydrology who require information on the range of natural variability. Traditionally, the time series has been regarded as stationary. Today, with the threat of the anthropogenic greenhouse effect apparent to all, regional climate data are being used to investigate the possible implications of man-made climate change. Two questions are currently being asked: can a trend be discovered in recent history? Do General Circulation Models (GCMs) reproduce realistically the statistics of the present-day regional climate?

In the present study we consider the space-time variability of Central European temperature in winter in the interval 1901–1980. The area ‘Central Europe’ is represented by 11 stations: Potsdam, Hohenpeißenberg, Frankfurt am Main, Jena and Hamburg (Germany), Uccle (Belgium), Wien (= Vienna) (Austria), Genève (= Geneva), Zürich (Switzerland), Fanø (Denmark) and Praha (= Prague) (Czech Republic) (Fig. 1). Unfortunately no site-specific information is available to us indicating to what extent the time series are homogeneous or affected by urbanization. A visual inspection of the raw data did not reveal any apparent inhomogeneities. To estimate the importance of the urbanization effect we compared the time series of the urban stations against the 2 rural stations Fanø and Hohenpeißenberg. These differences appear stationary for most stations, with the exceptions of Zürich and Genève (not shown). In Zürich temperature increases monotonically relative to the Hohenpeißenberg series, whereas the Genève temperature seems to suffer from a minor inhomogeneity at around 1920. We conclude that urbanization does not influence our data except for Zürich.

We define ‘winter’ as the January–February (JF) mean because the monthly mean temperatures in January and
February are more highly correlated than December is with January or February (e.g., for Potsdam station temperature the correlations ($\rho$) are: $\rho$(Jan, Feb) = 0.48, $\rho$(Dec, Jan) = 0.22 and $\rho$(Dec, Feb) = 0.20). Also, in December zonal circulations are more frequent than in January and February (Smirnov & Kazakova 1966, Miller et al. 1967, Hess & Brezowsky 1977). In the next section, we derive Empirical Orthogonal Functions (EOFs) from the JF mean temperatures at the 11 stations.

In the subsequent section, we analyze the relationship between the Central European temperature field and the large-scale circulation by means of a Canonical Correlation Analysis (CCA). As a parameter to represent the large-scale circulation we chose the sea-level pressure (SLP) field on a 5° X 5° grid from 35° N to 75° N and from 30° W to 40° E. The SLP data were prepared by the National Center for Atmospheric Research, Boulder, Colorado, USA. This data set was checked critically by Trenberth & Paolino (1980) who found no substantial data problem for the area of the North Atlantic area. Also, von Storch et al. (1993) used this data set to relate Iberian rainfall to the large-scale SLP field. They found that the substantial changes which took place in Iberian rainfall since the beginning of the century could be described by similar changes in the Atlantic SLP field. Our conclusion that the SLP data set is not contaminated by serious data problems is also supported by the study of Hense et al. (1990) who found the SLP changes from the beginning to the middle of the century to be consistent with the simultaneously observed sea-surface temperature (SST) changes.

An alternative candidate to represent large-scale circulation would be geopotential height. This parameter has the advantage over SLP of being hardly affected by local factors. Unfortunately, upper-air fields are available only from 1946 onward, and there are some inhomogeneities in the data set. We prefer therefore for our analysis the SLP data set which is fairly homogeneous and available from 1901 onward.

We then examine the consistency of trends in the large-scale circulation and temperature. In the penultimate section, the output of a coupled atmosphere-ocean climate model (ECHAM1/LSG) is screened to discover whether the links between large-scale circulation and Central European temperature found in the observed data are reproduced by the climate model. The final section offers a series of conclusions.

**EMPIRICAL ORTHOGONAL FUNCTIONS OF CENTRAL EUROPEAN TEMPERATURE IN JANUARY/FEBRUARY**

**Method**

EOFs are a powerful tool to identify the dominant coherent spatial patterns in a vector field. EOFs are the eigenvectors of the covariance matrix of the analyzed vector time series; the input data for the EOF analysis are anomalies of the meteorological parameters, which in the present study are the winter temperatures at the Central European stations.

If $C_i$, $i = 1, ... , m$ is a set of EOFs of a field $F(x,t)$, then the field

$$\sum_{i=1}^{m} c_i(t) \cdot C_i$$

with the coefficients $c_i$ as dot product of the $i$th EOF and the field $F$

$$c_i(t) = \sum_x F(x,t) \cdot C_i(x)$$

is an optimal representation of $F$ by $m$ orthogonal patterns. The coefficients $c_i(t)$ are called the EOF coefficients. The coefficients are normalized to variance 1, i.e., $\text{Var}[c_i(t)] = 1$, so that the information about the relative strength is in the patterns.

**Results**

We derived EOFs of Central European temperature for the complete interval 1901–1980 (Fig. 2) as well as for the 2 subintervals 1901–1940 and 1941–1980 (Fig. 3).
The percentages of explained variance are also given in Figs. 2 & 3 in the upper left corner of the panels. In all cases, the first EOF accounts for more than 80% whereas the second mode represents only 7% of the total variance.

The first EOF is almost uniform with typical anomalies of 2 K. The numbers are slightly larger in the northeast (2.5 K) and smaller at Genève and Wien (1.2 K) (Fig. 2). Thus, a 'normal winter' is everywhere cold or everywhere warm with typical January/February mean anomalies of ± 2 K. The first EOF pattern does not change if it is derived from the 2 sub-periods 1901–1940 and 1941–1980 (Fig. 3).

The second EOF describes a north-south contrast with positive values of 0.3 K along the northern edge of the Alps and negative values in the northern part of the region (–0.6 at Fanò and Hamburg; Fig. 2). That is, the most frequent deviations from the EOF-1 concept 'everywhere warm or everywhere cold' are 'cold in the northern lowlands and warm in the Alps' and vice versa. The pattern of the second EOF is also almost independent of the period from which it is derived (Fig. 3). Only Wien and Fanò show large differences in value between the 2 sub-periods 1901–1940 and 1941–1980.

The coefficients of the first two 1901–1980 EOFs are shown as time series in Fig. 4 and as frequency distributions in Fig. 5. The first EOF coefficient has mostly stationary variance between 1 and –1 but exhibits marked negative extremes in the years 1929, 1940, 1942, 1947, 1956 and 1963 (Fig. 4). These years are known to have been notable 'cold winters'. The distribution function is not Gaussian but markedly skewed (Fig. 5).

The coefficient of the second EOF is Gaussian distributed (Fig. 5). The greatest contrasts between the northern lowlands and the Alps were in 1912, 1936 and 1940 (Alps warmer than lowland) and in several winters in 1904–1914 and in the early 1930s (Alps colder than lowland; Fig. 4). There is a slight trend in this EOF coefficient that describes a gradual warming (0.006 K yr⁻¹) of the Alpine region and a slight cooling (–0.01 K yr⁻¹) in the lowlands.

**CANONICAL CORRELATION ANALYSIS OF REGIONAL TEMPERATURE AND THE SEA-LEVEL PRESSURE FIELD**

**Method**

In this section we analyze jointly the North Atlantic/European SLP and the regional Central European winter temperature using the CCA technique (Mardia
Mean temperature in JF ($10^{-2}$ K)

1901–1940

1941–1980

1st EOF

2nd EOF

7%

6%

Fig. 3. First 2 EOFs of temperature at 11 Central European stations derived from winters (JF) 1901–1940 (left) and 1941–1980 (right). Percentage numbers in the upper left corner of the diagrams are percentage of total variance accounted for by the pattern.
et al. 1979, Anderson 1984, Barnett & Preisendorfer 1987). The CCA yields 2 sets of patterns, \( \{C_i^{\text{SLP}}\} \) and \( \{C_i^T\} \) for the circulation parameter SLP and the temperature field \( T \). Any temperature field \( T(x) \) or sea-level pressure field \( SLP(x) \) may be expanded

\[
T(x) = \sum_{i=1}^{m} \gamma_i^T \cdot C_i^T(x) \quad \text{and} \quad SLP(x) = \sum_{i=1}^{m} \gamma_i^{\text{SLP}} \cdot C_i^{\text{SLP}}(x)
\]

so that the coefficients \( \gamma_i^T \) and \( \gamma_i^{\text{SLP}} \) share a maximum correlation if \( i = j \) and are orthogonal otherwise (if \( i \neq j \)). The coefficients \( \gamma_i^T \) and \( \gamma_i^{\text{SLP}} \) are normalized to variance 1 so that the information on the relative strengths is in the CCA patterns \( C_i^{\text{SLP}} \) and \( C_i^T \). The coefficients are given as the dot product of the field to be represented and the adjoint patterns \( \hat{C}_i^{\text{SLP}} \) or \( \hat{C}_i^T \), i.e.

\[
\gamma_i^{\text{SLP}}(t) = \sum_x SLP(x,t) \cdot \hat{C}_i^{\text{SLP}}(x)
\]

\[
\text{and} \quad \gamma_i^T(t) = \sum_x T(x,t) \cdot \hat{C}_i^T(x)
\]

A parameter to measure the strength of the link between a circulation pattern and the temperature distribution is the rate of explained variance:

\[
e_i(x) = 1 - \frac{\text{Var}[T(x,t) - \gamma_i^{\text{SLP}}(t) \cdot C_i^T(x)]}{\text{Var}[T(x,t)]}
\]

The CCA is done in subspaces spanned by the first few EOFs. This procedure has the double advantage that the input-data sets are independent (orthogonal functions) and that much of the noise in the data field is eliminated. The estimated correlation coefficients obtained in a CCA are positively biased: the more EOFs, the larger the correlation coefficients become. We try to avoid this bias by applying the following concept: we expect the largest correlation coefficient to grow rapidly as the number of EOFs increases as long as the added EOFs represent additional information. If the extra EOFs represent mostly noise, however, the correlation coefficient is expected to grow slowly. This procedure leads us to the choice of 5 EOFs (or less) of temperature and sea-level pressure.
The CCA was carried out with the sea-level pressure and temperature data from the 'estimation' interval 1901–1940; the explained variances $e_i(x)$ and the correlation coefficients are calculated from the complete interval 1901–1980 including the 'test' interval 1941–1980.

**Results**

The first CCA pair (Fig. 6) has a correlation [between $\gamma^T(t)$ and $\gamma^{SLP}(t)$] of 70% in the estimation period 1901–1940. The SLP pattern describes an anomalous southwesterly flow into Central Europe. This southwesterly flow advects anomalously warm maritime air so that the temperature pattern is positive everywhere with maximum values of almost 2 K along the Alps and minimum values of 0.7 K at Fans. In Fig. 6b the rate $e_1$ of explained local variance, as derived from the test sample 1941–1980, is also given. This rate is greater than 40% everywhere except for Zürich (31%), Hamburg (37%), and Fans (22%). Apparently, the first CCA pair determines mostly the southern part of the analysis area.

The second CCA pair (Fig. 7) specifies the northern part of the analysis area. The SLP field represents a strong northwesterly flow that affects mainly the northern part of Central Europe. There the typical temperature anomalies are almost 2 K (at Potsdam, Hamburg and Fans), and the rates $e_2$ of explained local variance are more than 40%. Along the Alps, at Genève and Zürich, the typical anomalies are 0.6 K with explained variances of less than 10%.

**CONSISTENCY OF TRENDS IN CIRCULATION AND TEMPERATURE**

**Method**

We may use the results of the CCA analysis of the previous section to derive temperature anomalies indirectly using only anomalous circulation. This 'CCA model' is given by:

$$\hat{T}(x,t) = \sum_{i=1}^{m} \gamma^i_{SLP}(t) \cdot C_i^T(x)$$

(von Storch et al. 1993) where $m$ is the number of CCA pairs used.

The power of the model in Eq. (6) can be monitored either by the rate of the $T$-variance explained by $T$

$$\epsilon(T, \hat{T}, x) = 1 - \frac{\text{Var}[T(x,t) - \hat{T}(x,t)]}{\text{Var}[T(x,t)]}$$

or by the correlation between $T$ and $\hat{T}$
The power of the model (Eq. 6) is assessed by calculating the 'explained variance' (Eq. 7) and the correlations (Eq. 8) in the 1941–1980 test subset. All variances and correlations are calculated with anomalies relative to the 1901–1940 means.

\[
\rho(T, \hat{T}, x) = \frac{\text{Cov}[T(x,t) - \hat{T}(x,t)]}{[\text{Var}[T(x,t)] \text{Var}[\hat{T}(x,t)]]^{1/2}} \quad (8)
\]

We computed the rates of explained variance at the 11 locations for \(m = 1, 2 \& 3\) (Fig. 8). Not surprisingly, the first CCA pair is sufficient for the southern stations (e.g. Genève), whereas the second CCA pair is needed for the northern stations (e.g. Fane). The addition of the third CCA pair does not increase the capacity of the CCA model (Eq. 6). Therefore we have chosen \(m = 2\).

With the \(m = 2\) model, more than 50% of the variance at the northern stations is explained. The maximum rate (\(e = 64\)% is at Hamburg, the minimum (\(e = 18\)% is at Zürich. The correlation is high everywhere, with average values of 80% (Fig. 9). This high level indicates that the year-to-year fluctuations are represented quite well by the CCA model, whereas the variability of \(e\) indicates that the CCA model has, at least at some locations, difficulties in reproducing the actual size of the anomalies.

The different success of the CCA model (Eq. 6) on the year-to-year timescale and on the multidecade timescale is documented in Fig. 10 for the 11 Central European stations considered. The similarity of the high-frequency variations is good not only for the estimation interval 1901–1940 but also for the test interval 1941–1980. That the CCA model does not describe the trends correctly is obvious: the differences between observations and reconstructions are large at all stations (Fig. 10).

According to in situ observations, winters became cooler in the north (\(-0.010\) \(\text{K yr}^{-1}\) at Hamburg) and warmer in the south (\(0.022\) \(\text{K yr}^{-1}\) at Zürich and \(0.011\) \(\text{K yr}^{-1}\) at Genève; Fig. 11a). The reconstructed trends (Fig. 11b) deviate markedly from the observed trends (Fig 11a). According to the changes in circulation, winters should have become cooler everywhere in Central Europe by \(-0.01\) to \(-0.03\) \(\text{K yr}^{-1}\) over the last 80 winters. The largest difference between observed (\(0.022\) \(\text{K yr}^{-1}\)) and estimated (\(-0.016\) \(\text{K yr}^{-1}\)) change is at Zürich (Fig. 10f), where we had found that the urbanization effect significantly contributed to the local temperature time series.
In the context of 'regionalization' of climate model output (von Storch et al. 1993) it is of interest to know if climate models do reproduce the observed link between large-scale circulation and regional climate. If this link is not reproduced then the regional climate is probably not simulated realistically. If the link is modelled realistically, then the regional climate might also be modelled realistically. With this in mind we examined the output of a 100 yr run with the coupled atmosphere-ocean climate model ECHAM1/LSG (Cubasch et al. 1992).

The SLP field is modelled on a grid similar to the grid on which the analyses of observed data were available. The Central European temperature field in the model is represented by 6 grid points between 53° N, 6° E and 48° N, 17° E. The CCA was carried out as with the observed data, in particular with 5 EOFs of SLP and temperature. The first 2 CCA pairs derived from the modelled data (Fig. 12) are similar to the first 2 CCA pairs inferred from the observed data (Figs. 6 & 7).

The first pairs, from simulated data as well as from observed data, represent an anomalous southwesterly flow into Central Europe associated with an overall warming in this area. Thus, the model reproduces the main aspects of the first CCA pair.

With respect to details, however, the model deviates markedly from reality. Firstly, the strength of the link in the observed data, as given by a correlation of 0.64, is underestimated by the climate model (0.53). Second, the relative strengths of the SLP pattern and of the temperature patterns deviate. In the observed data the maximum SLP (temperature) anomaly accounts for −5 mb (1.48 K; Fig. 6) whereas the modelled numbers are −3.5 mb and 2.3 K (Fig. 12a). Thus, in the model a considerably weaker SLP anomaly excites a temperature anomaly which is markedly larger than in the observed data. A third point worth mentioning refers to the patterns. The SLP patterns in Figs. 6 & 12a are, apart from their strengths, similar. The temperature patterns, however, are different. In the modelled data the main feature is a west-east gradient whereas in the observed data the dominant feature is a north-south gradient.

The patterns for the second CCA pairs derived from model output are shown in Fig. 12b and those derived...
Fig. 10. *In situ* observations and reconstructions, by means of concurrent SLP anomalies and model Eq. (6) with \( m = 2 \), of local JF mean temperatures at 11 Central European stations. The time interval 1901–1940 was used to fit the model, and the time interval 1941–1980 represents independent data. A linear trend derived from all data is added. Top: difference between reconstructions and *in situ* observations; bottom: anomalies derived from *in situ* observations (---) and reconstructed anomalies (—). Anomalies are defined as deviations from the 1901–1940 average.

This page: Fanø, Hamburg, Potsdam, Uccle, Genève

*Overleaf:* Zürich, Wien, Hohenpeißenberg, Frankfurt, Jena, Praha
Fig. 10 (continued)
from observations are in Fig. 7. With respect to spatial patterns the model does a credible job. An SLP difference between Northern Europe and the Atlantic of 5 to 6 mb in the model is associated with maximum temperature anomalies of 1.2 K in the northeast part of the considered area. In the observed data a SLP difference of about 10 mb yields a maximum temperature increase of almost 2 K. Therefore the strength of the signal in terms of magnitudes of the anomalies is successfully reproduced. The strength of the signal in terms of explained local variances and canonical correlation is, however, notably underestimated by the climate model. The correlation in the modelled data is only 0.20 compared to 0.61 in the observed data. In the simulated data the second SLP CCA pattern controls up to 20–25% of the winter mean temperature in the northeast part of Central Europe. In the real world the second SLP CCA pattern has its strongest effect also in the northeast part of Central Europe where it explains more than 40% of the local variance.

**CONCLUSIONS**

The main conclusions to be drawn are:

1. Two patterns are sufficient to describe the spatial variability of winter (JF) mean Central European temperature. With the first 2 EOFs over 90% of the variance of winter mean temperature is described.

2. The variations in winter mean large-scale North Atlantic/European circulation control large parts of the variations in Central European temperature. This is particularly true for the year-to-year variations.

3. The 80 yr trends of circulation and the regional temperature field are not linked to each other by the CCA patterns. Thus, the year-to-year variations and the century time-scale of Central European temperature are controlled by different mechanisms (see below).

4. The ECHAM1/LSG climate model does reproduce the broad features of the observed links between large-scale circulation and Central European temperature. The details of this link, however, are not well simulated. The strength of the ‘typical’ temperature anomalies that go with a ‘typical’ SLP anomaly are overestimated by the model. Also, the geographical details of the regional temperature signal are not reproduced.

There are 2 possible explanations for the discrepancy of the circulation and temperature trends in (3). One explanation might be that on the low-frequency timescale, circulation alone is not responsible for the temperature variations; instead large-scale thermal dis-
Fig. 12. First 2 pairs of CCA patterns inferred from the 100 yr output of a climate model (ECHAM4/LSG). Correlations are (a) 0.53 for the first pair and (b) 0.20 for the second pair. Left diagrams show SLP patterns with the 6 model grid points representing Central Europe. The associated temperature distribution at these 6 grid points, together with the percentage of local temperature variance, is shown in the right panels. Panels (a) should be compared with Fig. 6 and panels (b) with Fig. 7.

A wild-card in the present analysis is the urbanization effect: it is possible that the readings of temperature have systematically increased through the past 80 yr simply because of the increasing size and density of the towns in which many of the thermometers are placed. The urbanization effect would imply an artificial positive temperature trend. The comparison of the temperature time series from urban stations with those from the 2 rural stations Fanø and Hohenpeissenberg led us, however, to the conclusion that this effect does not interfere with our results, apart from Zurich.

We propose as a more likely hypothesis that the variability of the regional temperature on timescales of several decades is controlled not only by the circulation but also by low-frequency variations of Atlantic sea-surface temperature.

LITERATURE CITED


Editor: G. Esser

Manuscript first received: February 17, 1993
Revised version accepted: August 11, 1993