Impact of global sea surface temperature on summer and winter temperatures in Europe in a set of seasonal ensemble simulations

By HENRIK FEDDERSEN

Danish Meteorological Institute, Denmark

(Received 2 December 1998; revised 4 February 2000)

SUMMARY

In the PRediction Of climate Variations On Seasonal to interannual Time-scales (PROVOST) ensemble of seasonal simulations, performed with the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric general-circulation model forced by observed sea surface temperature (SST) for the 1979–1993 period, the consistency between individual members of the ensemble of simulated temperature at 850 hPa (T850) in Europe varies considerably from year to year. This interannual variability is assumed to be a consequence of slowly varying lower-boundary forcings, most notably the SST forcing, plus a random part from the internal, chaotic atmospheric variability. In those years when the ensemble members show a particularly consistent pattern of T850 anomalies in Europe, the consistency is significantly higher than would be expected from the interannual variability, indicating that the SST forcing has a significant impact on the model simulated temperature. Composites of global SST indicate that the El Niño/Southern Oscillation (ENSO) might be linked to the high internal consistency of some of the winter season ensembles of T850 anomaly patterns in Europe. However, in some ENSO years the ensemble shows no consistent signal in Europe, so the occurrence of an ENSO event is not sufficient to ensure high internal ensemble consistency in Europe.

In a separate approach, an attempt to linearly relate global SST anomaly patterns to the ensemble mean of T850, using canonical correlation analysis, also indicates a link between simulated T850 anomaly patterns in Europe and ENSO events. Mature El Niño (La Niña) conditions are associated with positive (negative) T850 anomalies in the model in southern Europe and negative (positive) anomalies in northern Scandinavia and over the Barents Sea in the January–February season. A weaker relation is found between onsetting ENSO conditions and T850 anomalies in the model in the south-eastern and westernmost parts of Europe in the July–August season.

Validations against re-analysed observations of T850 show poor skill in Europe for the ensemble mean of simulated T850, and no obvious relation is found between internal ensemble consistency and skill. The link that is found in the model between ENSO and simulated T850 in Europe, is not evident between ENSO and re-analysed T850.

KEYWORDS: El Niño Ensemble simulations Seasonal predictability in Europe

1. INTRODUCTION

Seasonal forecasting based on general-circulation models (GCMs) has become feasible during the last decade. Experimental seasonal forecasts based on GCMs from several climate and weather research centres are published regularly on the Internet. While research has demonstrated that the models have skill in the tropics (Brankovic et al. 1994; Shukla 1998), it remains to be seen to what extent extra-tropical seasonal climate can be predicted. In some extra-tropical regions, such as North America, the influence from the El Niño/Southern Oscillation (ENSO) ensures some predictive skill (e.g. Palmer and Anderson 1994; Trenberth et al. 1998) whereas in other regions, such as Europe, unpredictable, internal atmospheric variability appears to be a dominant component of the climate signal (e.g. Rowell 1998). Predictive skill on seasonal time-scales is believed to be closely connected to the lower-boundary forcing of the atmosphere (Charney and Shukla 1981; Palmer and Anderson 1994), most notably the sea surface temperature (SST).

Ensemble simulations of the atmosphere with prescribed forcing from observed SST have been used extensively to compute potential predictability defined as the fraction of the total variance in each grid box, which can be attributed to the known SST variability (Zwiers 1996; Rowell 1998). Potential predictability is the same as the anomaly correlation that can be obtained if the model is perfect (Rowell 1998). In the tropics

* Corresponding address: Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen, Denmark.
the SST-forced variability dominates on seasonal time-scales (i.e. high potential predictability), whereas in the extra-tropics internal atmospheric variability is of the same magnitude or higher than the SST-forced variability (i.e. low potential predictability). For extra-tropical regions of moderate potential predictability the consistency between individual members of an ensemble normally varies considerably from season to season. The most likely reason is that some of the observed SST patterns have an impact on the model seasonal climate, while others have little or no impact. It is of interest to identify those SST patterns that act as a source of predictive skill if the model simulations agree with observations, or as a boundary forcing to which the atmospheric model responds erroneously if the model simulations do not agree with observations.

The extent to which European seasonal climate is connected to variations in SST is an open question. European climate has in previous work been linked to SST anomalies in both the North Atlantic and the tropical Pacific. Ratcliffe and Murray (1970) found lagged correlations between SST anomalies in an area off Newfoundland and surface pressure over Europe. In a modelling study Palmer and Sun (1985), forcing their atmospheric GCM with enhanced SST anomalies in the area off Newfoundland, were partly able to reproduce the observational results of Ratcliffe and Murray (1970). Palmer and Sun (1985) suggested that a positive feedback between atmosphere and ocean could be the factor that would allow anomalies to persist over long periods, as is frequently observed. Recently, Colman (1997) described a lagged correlation between January–February SST anomalies in the North Atlantic and observed July–August surface air temperature in central England.

Connections between European climate and ENSO have also been sought. Fraedrich and Müller (1992) used observations, including 26 warm and 22 cold ENSO events, to estimate the influence of ENSO on European climate in winter. They found that warm ENSO events on average led to cold winters in Scandinavia and northern Europe and mild winters in the Mediterranean, whereas Johansson et al. (1998) found no predictability associated with northern Europe for ENSO. Non-zero correlations between the Southern Oscillation and Iberian seasonal rainfall were reported by Rodó et al. (1997).

In a number of modelling studies atmospheric GCMs have been used to identify the impacts of tropical Pacific SST anomalies on global climate on seasonal to interannual time-scales (e.g. Shukla and Wallace 1983; Lau and Nath 1994). In the Pacific/North American region the models all agree well with observational studies of the main features of the large-scale circulation patterns in winter. Although the response in the Atlantic/European sector is much weaker, May and Bengtsson (1998) found a statistically significant response which is not inconsistent with the observational study of Fraedrich and Müller (1992).

In the following sections, nine-member ensembles of simulated summer and winter seasonal 850 hPa temperatures for the 15-year period 1979–93 are analysed in order to study to what extent global SST patterns, such as those associated with ENSO, can be linked to seasonal temperature anomalies in Europe. The simulations were performed with the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric model which was forced by observed SST (Becker 1998; Brankovic and Palmer 2000). With an ensemble of nine simulations, each simulation being forced by identical SST, but started from different initial conditions, it is possible to extract information about potential impacts of SST on seasonal climate, even in areas where the impact is weak. Such weak impacts are likely to remain undetected in observational studies as observations represent only one realization of many possible atmospheric responses to a given SST forcing.
After a brief description of data in section 2, interannual variations in the distribution of anomaly correlations between pairs of simulated T850 in Europe from each of the 15 summer and 14 winter ensembles are studied in section 3. The anomaly-correlation distributions show the consistency of the T850 anomaly patterns of the nine members of each ensemble. Composites of SST anomalies from the years of high T850 consistency in Europe are calculated in order to establish whether any particular SST anomalies are pronounced during those years. A separate approach described in section 4 uses canonical correlation analysis (CCA) (Barnett and Preisendorfer 1987) to linearly relate global SST anomaly patterns to the ensemble mean of simulated T850 in Europe. CCA patterns and CCA time series are compared with the anomaly-correlation distributions and composites of section 3. The agreement between simulated and observed T850 in Europe is discussed in section 5, and concluding remarks are presented in section 6.

2. DATA

All data is taken from the ‘ECMWF Seasonal Simulations CD-ROM’ set which contains selected data from the ECMWF PROVOST (PRediction Of climate Variations On Seasonal to interannual Time-scales) simulations. In the PROVOST project the ECMWF Integrated Forecasting System model (cycle 13r4, horizontal resolution T63, 31 vertical levels) (Simmons et al. 1988; Branković and Palmer 2000) was used to generate ensembles of simulated seasonal climate. Each ensemble consists of nine members each approximately 120 days long and covering the seasons: March, April, May, June (spring); June, July, August, September (summer); September, October, November, December (autumn); and December, January, February, March (winter) during the period March 1979–December 1993. The nine members within an ensemble are initiated from nine consecutive 12 UTC analyses leading up to the first day in the target season (e.g. 22 November, 23 November, . . . , 30 November for the winter simulations). The initial conditions, as well as the prescribed SST forcings, are taken from the ECMWF re-analyses. Data is stored on the CD-ROMs as 10-day means on a 2.5° × 2.5° grid.

The present analysis of T850 in Europe is restricted to ‘high summer’ and ‘mid winter’ which are here defined as the 60-day periods from 1 July until 29 August and from 31 December until 28 February, respectively. Europe is defined as the region between 35°N–70°N and 10°W–42.5°E. SST fields are, for reasons of computational efficiency, interpolated onto a 5° × 5° grid.

3. INTERNAL CONSISTENCY BETWEEN MEMBERS OF AN ENSEMBLE

A systematic model response to prescribed SST forcing leads to a consistent signal in the members of an ensemble which is generated by simulations with identical SST forcing. Conversely, internal consistency between the members of an ensemble indicates that the actual SST has an impact on the seasonal climate in the region in question. Internal consistency can also be caused by memory of the relatively highly correlated initial conditions or by pure chance. However, memory of initial atmospheric conditions is likely to be insignificant as the analysis in the following is applied to months 2 and 3 of the model integration, whereas the simulations of the first month are not considered. The likelihood of the chance option is estimated through statistical tests in the following.

The internal consistency between the nine simulations within an ensemble can be described by the distribution of anomaly correlations between the 36 possible pairs of
Figure 1. Distribution of anomaly correlation coefficients (ACC) between members of (a) each of the 15 ensembles of simulated 850 hPa temperature (T850) in Europe in July–August 1979–93; (b) each of the 14 ensembles of simulated T850 in Europe in January–February 1980–93.

the ensemble members. The anomaly correlation coefficient (ACC) between a pair of T850 simulations is given by (Wilks 1995)

\[
ACC = \frac{\sum_i (T850_{i,m} - T850_{i,clim})(T850_{i,n} - T850_{i,clim})}{\sqrt{\sum_i (T850_{i,m} - T850_{i,clim})^2 \sum_i (T850_{i,n} - T850_{i,clim})^2}},
\]

where the summations include all grid points, \(i\), in the European region; indices \(m\) and \(n\) are used for the two ensemble members, and \(T850_{i,clim}\) is the model climatology at grid point \(i\).

If there is little internal consistency, the distribution of correlations can be expected to be centred around a correlation of zero, whereas high internal consistency will lead to a distribution which is skewed towards positive correlations.

High internal ensemble consistency in terms of anomaly correlations indicates a shift of the underlying probability density function (p.d.f.) away from the model climatology. High internal ensemble consistency may be associated with low spread between the ensemble members (narrow p.d.f.), but, this is not always the case, particularly for a region with high internal variability such as Europe.

Figure 1(a) shows the distribution of T850 anomaly correlations in the July–August season for all 15 years, i.e. for 15 \(\times\) 36 ACCs. The distribution is slightly skewed towards positive correlations with an average ACC of 0.12. In order to test the null hypothesis that SST forcing has no impact, a significance test of the Monte Carlo type is applied. In this test each of the average ACCs of 1000 random distributions is compared with the actual average ACC of 0.12. A random ACC distribution is derived by a random reorganization of the T850 simulations into 15 nine-member ensembles and subsequent calculation of the 36 possible ACCs for each ensemble. The result of the test was that none of the 1000 average ACCs of the distributions of reorganized T850 simulations exceeded 0.12, i.e. the null hypothesis is rejected at a significance level < 0.001, and therefore it is concluded that, on average, the simulated T850 in Europe shows a weak systematic response to the prescribed SST forcing. Figure 1(b) shows a distribution which is also slightly skewed towards positive T850 ACCs (with an average ACC of 0.08) for the 14 January–February simulations between 1980 and 1993. Also, in this case, the null hypothesis that SST forcing has no impact is rejected at a significance level < 0.001.
The distributions of correlations between T850 ensemble members for the individual years reveal substantial differences from year to year, both in summer and winter. In some years the SST forcing appears to have a strong impact on T850 in Europe, consequently the individual members of the simulated T850 ensemble for those years show high internal consistency, and the ACC distributions are shifted towards positive ACCs, while in other years the SST forcing appears to have no impact on T850 in Europe, and the ACC distributions are centred around zero.

Examples of years in which the T850 ensemble members were internally consistent include 1982/83 and 1988/89 which, incidentally, were both ENSO years (warm event/El Niño in 1982/83; cold event/La Niña in 1988/89). Figure 2 shows that the distributions of ACCs in both summer and winter of those two years are shifted towards higher ACCs than the distributions covering the entire 1979–93 period shown in Fig. 1.

The mean of a distribution for an individual year, i.e. the average ACC between pairs of simulations within the ensemble, gives an indication of the skewness of the distribution. Figure 3(a) shows increased average ACCs and thus higher internal consistency within the ensemble in July–August of both 1982 and 1988 as well as in 1979, 1980 and 1993. Similarly, Fig. 3(b) shows increased average ACCs for January–February in 1983, 1985 and 1989.
The statistical significance of the shifts in the ACC distributions of the individual years (compared with all other years) is tested using the Wilcoxon–Mann–Whitney test (Wilks 1995). The null hypothesis that the distribution is not shifted is, for the July–August season, rejected at the 5% level for 1979, 1982, 1988 and 1993, and for the January–February season the null hypothesis is rejected at the 5% level for 1983, 1985 and 1989, so in each of those seven seasons the prescribed SST is likely to have a systematic impact on model simulated T850 in Europe.

In an attempt to isolate SST anomaly patterns that might be associated with increased internal ensemble consistency, composite maps of absolute values of SST anomalies are calculated for the above-mentioned years. Absolute values are chosen so that, e.g. warm and cold ENSO events which may both lead to increased internal ensemble consistency of T850 in Europe, will not cancel when composited. The SST anomaly composites are computed for three consecutive two-month periods in order to capture possible lagged connections between SST and T850 (one two-month period coincides with the T850 period, i.e. either July–August or January–February, and the other two two-month periods are immediately prior to the T850 period).

The statistical significance of the composites is tested by comparing the actual composite with composites of all other combinations of four (for July–August) or three (for January–February) years. The actual value of the composite in a grid point is only considered significant if it is greater than 95% of all possible composites of the same number of years.

The SST anomaly composites show that the four years of internally consistent July–August T850 in Europe are not associated with systematic SST anomalies that are significant in the manner described above, whereas the composite of the three years of internally consistent January–February T850 in Europe is dominated by a strong ENSO signal with relatively large SST anomalies in the central, tropical Pacific (Fig. 4). This is not surprising as all three years in the composite are ENSO years, but note that internal ensemble consistency was not increased during other ENSO events such as those of 1986/87 and 1991/92. Thus, while high internal consistency in winter was only found during ENSO events, the occurrence of an ENSO event is not a sufficient condition for increased internal ensemble consistency of T850 in Europe.
Figure 4. Composites of absolute sea surface temperature anomalies (degC) from September to February 1982/83, 1984/85 and 1988/89 where the ensembles of model simulated 850 hPa temperature in Europe were internally consistent in January–February. Only statistically significant values are shown.
4. LINEAR PATTERN ANALYSIS

SST anomaly patterns that can be associated directly with simulated T850 in Europe can be found using a variety of linear methods (Bretherton et al. 1992). In the following, coupled patterns of SST and T850 are found by applying CCA (Barnett and Preisendorfer 1987) to fields of quasi-global (40°S–75°N) SST anomalies and ensemble means of simulated T850 in Europe. The SST field is divided into two two-month periods prior to, and one two-month period coinciding with, the T850 target period in order to capture possible lagged connections between SST and T850.

In the notation of Bretherton et al. (1992), the left data field, \( s(t) \), is a vector which is formed by stacking SST anomaly fields from the three consecutive two-month periods, i.e.

\[
s(t) = [\text{SST}_1(t) \text{SST}_2(t) \text{SST}_3(t)],
\]

where the indices 1–3 refer to the periods March–April, May–June and July–August or to September–October, November–December and January–February. The corresponding right data field, \( z(t) \), is given by the simulated T850 ensemble-mean anomaly field in Europe in July–August or in January–February, i.e.

\[
z(t) = \text{T850}_3(t).
\]

The left and right data fields are normalized so that data in every grid point has zero mean and unit standard deviation and thus contributes equally to the CCA.

Following Barnett and Preisendorfer (1987) the CCA is not applied directly to the left and right data fields, but to pre-filtered fields that are obtained as the first few principal components of the normalized anomaly fields, \( s(t) \) and \( z(t) \). The number of principal components to include in the CCA should, on the one hand, be less than the number of years in the time series and small enough to avoid overfitting; on the other hand, the number of principal components should not be so small that a potentially important part of the signal is left out. In practice, the number of principal components that were retained, was decided after significance tests and cross-validation of canonical correlations.

CCA transforms the principal components into pairs of time series, \( u_i(t) \) and \( v_i(t) \), and vectors (or patterns), \( \alpha_i \) and \( \beta_i \), which are linearly related through

\[
u_i(t) = \alpha_i \cdot \text{PC}_s(t)
\]

\[
u_i(t) = \beta_i \cdot \text{PC}_z(t),
\]

where \( \text{PC}(t) \) is a vector of the retained principal components. The time series \( u_i(t) \) and \( v_j(t) \) are uncorrelated for \( i \neq j \), and the canonical modes, denoted by indices \( i \) and \( j \), are ordered such that the correlation between \( u_i(t) \) and \( v_i(t) \), the so-called canonical correlation, is decreasing with increasing mode number \( i \).

CCA maximizes the canonical correlation for the first mode. For short time series and many spatial degrees of freedom CCA will tend to find highly correlated linear relationships between any pair of noisy fields (Cherry 1996). Therefore, a Monte Carlo type test has been applied to test whether the correlation found is significantly greater than the canonical correlation that could be expected by chance. In this test the order of the T850 ensemble means are randomly scrambled and the CCA, including the pre-filtering, is repeated. If any one of the years remains in the same location after the random scrambling, the scrambling is repeated before proceeding with the CCA, in order not to get contributions to the random canonical correlations from genuine correlations between SST and T850 anomalies.
Five hundred such random canonical correlations are calculated, and the null hypothesis that the SST and T850 anomaly fields are not correlated, is rejected at a level of significance which is estimated as the fraction of the number of times the random canonical correlation is greater than the actual canonical correlation.

The robustness of the CCA was checked further using cross-validation where one year is removed, and CCA is applied to the remaining years. The retained principal components for the year that was removed, are projected onto the new canonical patterns. This procedure is repeated for all years to obtain cross-validated time series. The correlation between the cross-validated time series gives a good indication of the robustness of the CCA (Feddersen et al. 1999). Additionally, the pattern correlations between each pair of the first-mode cross-validation canonical patterns and the actual first-mode canonical patterns are examined.

CCA was applied to the SST and T850 anomaly fields and the pre-filtering was done with a systematically increasing number of retained principal components, and it was found, using the methods described above, that retaining the first five principal components of both fields gave a very robust CCA in July–August, whereas only four principal components of the SST anomaly field and three principal components of the T850 anomaly field should be retained in January–February. For July–August the canonical correlation was significant at the 6% level for this choice; the cross-validated canonical correlation was 0.73—compared with the actual canonical correlation of 0.97—and the correlations between the first-mode cross-validated canonical patterns and the actual canonical patterns ranged from 0.76 to 1.00 for SST and from 0.93 to 1.00 for T850. For January–February the canonical correlation was significant at the 2% level; the cross-validated canonical correlation was 0.90—compared with the actual canonical correlation of 0.95—and the correlations between the first-mode cross-validated canonical patterns and the actual canonical patterns ranged from 0.86 to 1.00 for SST and from 0.95 to 1.00 for T850. The results are not very sensitive to small changes in the number of retained principal components, but if too many or too few principal components are retained, then the CCA results are no longer robust.

Figure 5(a) shows the leading-mode time series, \( u_1(t) \) and \( v_1(t) \), for the July–August season. The corresponding heterogeneous correlation patterns are shown in Fig. 6 (SST) and Fig. 8(a) (T850). Similar time series and heterogeneous correlation patterns for the January–February season are shown in Figs. 5(b), 7 and 8(b).

The left and right heterogeneous correlation patterns, which are obtained as correlations in each grid point between \( s(t) \) and \( u_i(t) \) and between \( z(t) \) and \( u_i(t) \), respectively, indicate how strongly one field is linked to the time series of the \( i \)th CCA mode of the other field. A measure of the total significance of a heterogeneous correlation pattern is given by the fraction of variance which is explained by the time series corresponding to the other field. This fraction of explained variance is easily computed as the area average of the squared values in the heterogeneous correlation pattern (Lau and Nath 1994).

Figure 8 shows correlations between the leading-mode time series for the SST anomaly field and global T850 although CCA is only applied to T850 in Europe (inside the box on the maps). Areas of high correlations (positive or negative) in the global patterns in Fig. 8 are areas where seasonal T850 variations are linearly related to those SST variations which, per construction, are linearly related to seasonal T850 variations in Europe.

The patterns in Fig. 6 show a temporal development in most of the Pacific Ocean which agrees well with SST anomaly patterns in the onsetting stages of a generic warm ENSO event (Harrison and Larkin 1998). Only in the eastern equatorial Pacific is there no sign in Fig. 6(c) of the strong warming which is seen in the generic warm ENSO
event. The associated SST time series in Fig. 5(a) shows that the leading CCA mode is most pronounced in 1982, 1988, 1992 and 1993. But only in three (1982, 1988 and 1993) out of those four years does the SST forcing of the atmospheric model result in a T850 response in Europe which has significantly higher internal consistency than average within the ensemble (Fig. 3). Thus, the model fails to respond consistently in Europe in July–August 1992.

The area-averaged SST anomalies shown in Table 1 for July–August 1982, 1988, 1992 and 1993 from the three tropical areas of high correlations in Fig. 6(c) (the Indonesian region, the central tropical Pacific and the tropical Atlantic) may give a hint
Figure 6. Temporal correlations (%) between prescribed sea surface temperature (SST) in each grid box and the time series, $v(t)$, of the first canonical correlation analysis (CCA) mode for July–August 850 hPa temperature in Europe. The first CCA mode accounts for 11% of the total, unfiltered SST variance. See text for further explanation.

as to why the model fails. Two entries in Table 1 show particularly weak anomalies: July–August 1992 in the Indonesian region and July–August 1993 in the Atlantic region, but only in 1992 does the model fail to produce an internally consistent T850 response in Europe. This suggests that SST anomalies in the Indonesian region might be critical to the model’s atmospheric response in Europe whereas SST anomalies in the tropical Atlantic might be of minor importance. Integrations of the same model are needed for
years that are not included in the 1979–93 period in order to confirm or reject this hypothesis.

In the extra-tropics correlations between SST anomalies and the leading CCA mode of T850 in Europe in July–August are generally weaker than in the tropics (Fig. 6). In particular, SST anomalies in the oceans immediately adjacent to Europe, the northeastern Atlantic and the Mediterranean, do not correlate with T850 in Europe.
Figure 8. Temporal correlations (%) between simulated 850 hPa temperature (T850) (ensemble mean) in each grid box and the time series, $u_1(t)$, of the first canonical correlation analysis (CCA) mode for sea surface temperature (SST) (a) July–August T850, March–August SST; (b) January–February T850, September–February SST. The first CCA mode accounts for 23% of the total, unfiltered T850 variance in Europe (the box) in July–August and for 36% of the total, unfiltered T850 variance in Europe in January–February. See text for further explanation.
TABLE 1. MEAN SEA SURFACE TEMPERATURE ANOMALIES IN DEGC

<table>
<thead>
<tr>
<th></th>
<th>Indonesian region</th>
<th>Central Pacific region</th>
<th>Atlantic region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jul–Aug 1982</td>
<td>−0.48</td>
<td>0.81</td>
<td>−0.75</td>
</tr>
<tr>
<td>Jul–Aug 1988</td>
<td>0.52</td>
<td>−1.13</td>
<td>0.71</td>
</tr>
<tr>
<td>Jul–Aug 1992</td>
<td>−0.19</td>
<td>0.42</td>
<td>−0.77</td>
</tr>
<tr>
<td>Jul–Aug 1993</td>
<td>−0.41</td>
<td>0.51</td>
<td>−0.17</td>
</tr>
</tbody>
</table>

Indonesian region (5°S–5°N, 125°E–155°E), central tropical Pacific region (5°S–5°N, 180°W–150°W) and the tropical Atlantic region (5°S–5°N, 30°W–0°). 1988 was a La Niña year; the other three were El Niño years—hence the opposite signs in 1988.

The patterns in Fig. 7 agree very well with the SST patterns in both the Pacific and the Indian Oceans of a generic warm ENSO event (Harrison and Larkin 1998). The patterns are most prominent in 1982/83 and in 1988/89 (Fig. 5(b)), and as the T850 internal ensemble consistency in Europe is high in January–February in both 1983 and 1989 (Fig. 3(b)), ENSO-type SST anomaly patterns resembling those of Fig. 7 appear to have an impact on T850 in Europe in January–February in the ensemble simulations.

The SST anomalies evolve differently during different ENSO events. Therefore, some events are not evident in Fig. 5(b), e.g. the 1991/92 El Niño. This, and other ENSO events (e.g. 1986/87), do not lead to increased T850 internal ensemble consistency in Europe either. Based on the present material it has not been possible to conclude whether this is because these events are weaker than the 1982/83 and 1988/89 events, or whether the model response depends critically on certain localized SST anomalies that are not present in every ENSO event, or, possibly, a combination of both mechanisms.

Figure 8 shows T850 correlation patterns. In the July–August season correlations are negative in all of Europe (Fig. 8(a)). The most negative correlations are found in the south-eastern and westernmost parts. Thus, simulated T850 in Europe tends to be below normal for ENSO-like SST anomaly patterns that resemble those of Fig. 6 in March–August. The leading CCA mode for SST is inversely related to simulated T850 in Europe, large parts of North America, the northern Atlantic and the western part of Asia. The general agreement over most of the Pacific Ocean between the T850 and SST correlation maps shows that in this region the simulated T850 tends to follow the prescribed SST.

In the January–February season the T850 correlation pattern (Fig. 8(b)) shows high positive values in southern Europe and contrasting high negative values in northern Scandinavia and the Barents Sea. That is, during warm (cold) ENSO events the model atmosphere responds with mild (cold) winters in southern Europe and cold (mild) winters in northern Scandinavia and the Barents Sea. The high correlations in southern Europe are part of a band of high correlations which stretches from the central and eastern equatorial Pacific north-eastwards across the northern, tropical Atlantic, north-west Africa and the Mediterranean to parts of central Asia. The well established connection between ENSO and a north–south temperature-anomaly dipole in North America (Ropelewski and Halpert 1986; Livezey et al. 1997) is also evident in the global T850 correlation pattern in Fig. 8(b).

European seasonal climate is known to be strongly influenced by the North Atlantic Oscillation (NAO) which in winter can be recognized as a dipole in sea-level pressure (SLP) anomalies roughly between Iceland and the Azores (Van Loon and Rogers 1978; Barnston and Livezey 1987). In summer the dipole centres are shifted to positions
Figure 9. As Fig. 8, but for sea-level pressure instead of 850 hPa temperature.
approximately over northern Greenland and the British Isles (Hurrell and van Loon 1997). Figure 9 shows how global SLP anomalies correlate with the SST time series of Fig. 5. In the July–August season (Fig. 9(a)) there is no sign of the NAO. But in January–February Fig. 9(b) shows that SLP around Greenland tends to vary in phase with the SST time series of Fig. 5(b) (positive correlations), while SLP in most of Europe and adjacent parts of the North Atlantic tends to vary 180° out of phase with the SST time series (negative correlations). Thus, the ENSO-like SST patterns of Fig. 7 are associated with an NAO-like SLP pattern in the North Atlantic.

5. Comparing simulated and observed temperatures in Europe

The results described in the previous sections are based entirely on model simulations. In this section an attempt is made to illustrate how well the simulated T850 in Europe compares with observed, i.e. ECMWF re-analysed, T850.

Skill scores in terms of anomaly correlations are shown in Fig. 10. Here the simulation anomalies are calculated with respect to the model climatology, and the analysis anomalies are calculated with respect to the analysed T850 July–August or January–February 1979/80–1993 mean. There are large interannual variations in the skill of the ensemble mean, and large spread almost every year in the anomaly correlations for the individual ensemble members. The apparent lack of a relation between the ensemble-mean anomaly-correlation skill and the internal ensemble consistency (or perfect model skill; Fig. 3) indicates that the model simulations of T850 in Europe are not perfect, and that ensemble forecasts that are made with this model may not be totally reliable.

With only nine members in the ensembles it is to be expected that none of the individual ensemble members match the corresponding observation perfectly. The question is whether this is simply a sampling problem, or whether the model, for some reason, is in error so that the ensemble of simulations does not encompass the observation.

The location of the observation relative to the ensemble of simulations can be described by the probability distribution of anomaly correlations between the observation and each of the individual ensemble members. This distribution can then be compared with the probability distributions of anomaly correlations between any two of the ensemble members. If the probability distribution that includes the observation is significantly different from the other distribution, then the difference between the observation and the simulations is unlikely to be due only to sampling problems.

The null hypothesis that the two sets of obtained anomaly correlations are realizations of the same distribution, is tested using the Wilcoxon–Mann–Whitney test (as in section 3). The resulting position of the test statistic in probability space is shown in Fig. 11 for each year in July–August and in January–February. Values near zero occur when the ACCs between analysed T850 and the ensemble of simulated T850 tend to be less than the ACCs between pairs of simulated T850 within the ensemble, i.e. when the actual model skill is significantly less than the perfect model skill. Values near one occur when the opposite is the case, i.e. when the actual model skill is significantly greater than the perfect model skill. The null hypothesis is rejected at the 5% level if the test statistic exceeds one of the dashed lines in Fig. 11.

Significant differences are found in July–August 1979 (Fig. 11(a)) and in January–February 1983 and 1985 (Fig. 11(b)). For another three seasons (July–August 1982 and January–February 1984 and 1989) the null hypothesis is almost rejected at the 5% level. Except for the January–February 1984 seasons, the significant and near-significant differences are all found in seasons where the ensemble of simulated T850 in Europe is
Anomaly correlation coefficient (ACC) skill scores for simulated 850 hPa temperature in Europe in (a) July–August and (b) January–February. Crosses show ACCs for each ensemble member; squares show ACCs for the ensemble mean.

Figure 10. Anomaly correlation coefficient (ACC) skill scores for simulated 850 hPa temperature in Europe in (a) July–August and (b) January–February. Crosses show ACCs for each ensemble member; squares show ACCs for the ensemble mean.

internally consistent. This suggests that the model responds too strongly to certain SST forcings, or that the form of the response is wrong.

An alternative way in which to illustrate the performance of the ensemble simulations is to plot each member of an ensemble of simulated T850 in Europe in a plane spanned by the first two principal components (which together account for more than 70% of the total variance) and to compare this with the corresponding analysed T850 anomaly field projected onto the first two empirical orthogonal functions of the simulations. In the three above-mentioned cases, where the model simulations were found to be most different from observations, based on the test of anomaly-correlation distributions, the analyses are also found to lie outside the ensemble of simulations in the principal component plane (Figs. 12(a), (b) and (c)). The analyses may, however, also lie outside
Figure 11. Position of Wilcoxon–Mann–Whitney test statistic (p) in probability space for each year in July–August (a) and January–February (b). Null hypothesis is that anomaly correlation coefficients (ACCs) between analysed and simulated 850 hPa temperature (T850) in Europe, and ACCs between pairs of simulated T850 from one ensemble are realizations of the same ACC distribution. Dashed lines indicate the 5% rejection level.

Figure 12. Each ensemble member of simulated 850 hPa temperature (T850) anomalies (×) in Europe in a plane spanned by the first two principal components (PC1 and PC2), explaining about 70% of the total variance, as well as the corresponding ECMWF re-analysed T850 anomaly field projected onto the first two empirical orthogonal functions of the simulations (■). July–August 1979 (a), January–February 1983 (b), January–February 1995 (c) and January–February 1990 (d).
the ensemble of simulations in the principal component plane in seasons when the ACC distributions are not significantly different; an example is shown in Fig. 12(d).

6. Concluding Remarks

The present analysis of the ECMWF PROVOST simulations shows that global SST has an impact on model simulated T850 in Europe (i.e. simulated T850 anomalies in Europe are not purely random). An attempt was made to identify SST patterns that have a systematic impact on simulated T850 in Europe. With only 15 years of data it is hard to draw any firm conclusions, but at least some ENSO events appear to have an impact on model simulated T850 in Europe. The model responds to one particular El Niño (La Niña)-type SST pattern (Fig. 7) with mild (cold) conditions in January–February in southern Europe and cold (mild) conditions in northern Scandinavia and over the Barents Sea.

The ensemble mean of simulated T850 in Europe appears to be statistically related to ENSO in January–February and possibly also in July–August, but not every ENSO event appears to have an impact on the internal T850 ensemble consistency in Europe. Also, the high internal T850 ensemble consistency in some of the non-ENSO years, particularly in July–August, suggests that there are SST patterns not related to ENSO that also have a systematic impact on model simulated T850 in Europe. Those SST patterns have not been identified, probably because they are less frequent than ENSO-related SST patterns and, therefore, are not captured by compositing or CCA applied to only 15 years of data.

The simulated changes in T850 and SLP in Europe during strong ENSO events are in qualitative agreement with the observational study of Fraedrich and Müller (1992) and with the modelling study of May and Bengtsson (1998). In all cases a warm ENSO event is associated with a reduction of the Icelandic low in winter which leads to an approximate north–south dipole of lower-tropospheric and surface temperature anomalies in Europe with cold anomalies in the north and warm anomalies in the south. However, there is not general agreement on the exact location of the anomalies. Both in observations (Fraedrich and Müller 1992) and in the ECHAM3 model simulations the centre of cold anomalies is located in central Scandinavia (May 1995), whereas in the ECMWF PROVOST simulations the centre of cold anomalies is located further northeast over the Barents Sea (Fig. 8).

A comparison between the ECMWF PROVOST simulations and re-analyses of T850 in Europe shows poor skill in the 1979–93 period, and there appears to be no obvious relation between the internal ensemble consistency and the model skill. In some cases the ensemble of T850 simulations is significantly different from the re-analyses. Thus, while the model responds systematically to the prescribed SST forcing, the T850 response does not agree with observations in Europe. Therefore, one should not have complete confidence in forecasts from coupled models that have the ECMWF atmospheric model as one component.

An example is the experimental seasonal forecasts computed with the ECMWF coupled system (Stockdale et al. 1998). During the strong warm 1997/98 ENSO event the predictions (not shown) were in good agreement in January–February with the global temperature and pressure correlation patterns shown in Figs. 8(b) and 9(b), but the results presented in the present paper suggest that ensemble predictions that are made with a strong ENSO-like SST forcing are less likely to encompass analysed T850 than ensemble predictions that are made with more neutral SST in the tropical Pacific. So when the sign of the anomaly of predicted T850 for January–February 1998 agreed
with that of analysed T850 (not shown) in most of continental Europe as well as in the
north-eastern part of Scandinavia, it could very well be for reasons other than ENSO.

ACKNOWLEDGEMENTS

Comments from Wilhelm May on an early version of the manuscript are much
appreciated as are comments and suggestions from the anonymous referees. The ‘ECMWF
Seasonal Simulations CD-ROM’ set was provided by ECMWF. The work was partly
funded by the Commission of the European Union (contract ENV4-CT95-0109).

REFERENCES

Barnett, T. P. and Preisendorfer, R. 1987 Origins and levels of monthly and seasonal forecast skill for
Barnston, A. G. and Livezey, R. E. 1987 Classification, seasonality and persistence of low-frequency
Becker, B. D. 1998 ECMWF seasonal ensemble simulations on CD-ROM. WMO
Bull., 47, 276–280
Brancković, Č and Palmer, T. N. 2000 Seasonal skill and predictability of ECMWF PROVOST ensem-
climate data. J. Climate, 5, 541–560
J. Lighthill and R. Pearce. Cambridge University Press
Colman, A. 1997 Prediction of summer central England temperature from preceding
Feddersen, H., Navarra, A. and Ward, M. N. 1999 Reduction of model systematic error by statistical correction for
dynamical seasonal predictions. J. Climate, 12, 1947–1989
J. Climatol., 12, 25–31
Hurrell, J. W. and van Loon, H. 1997 Decadal variations in climate associated with the North Atlantic
oscillation. Climatic Change, 36, 301–326
Europe. J. Atmos. Sci., 55, 103–127
Lau, N.-C. and Nath, M. J. 1994 A modeling study of the relative roles of tropical and extratropical
SST anomalies in the variability of the global atmosphere–

onean system. J. Climate, 7, 1184–1207
atmosphere to large central equatorial Pacific SST anom-
aliess. J. Climate, 10, 1787–1829
May, W. 1995 ‘Der Einfluss des El Niño/Southern Oscillation Phänomens auf
die atmosphärische Zirkulation in den mittleren Breiten der
Nordhalbkugel: Beobachtungen und Modellsimulationen’. PhD thesis, Max-Planck-Institut für Meteorologie, Hamburg
May, W. and Bengtsson, L. 1998 The signature of ENSO in the Northern Hemisphere midlatitude
seasonal mean flow and high-frequency intraseasonal vari-
Meteorol. Soc., 120, 755–793
Palmer, T. N. and Sun, Z. 1985 A modelling and observational study of the relationship between
sea surface temperature in the north-west Atlantic and the
atmospheric general circulation. Q. J. R. Meteorol. Soc., 111,
947–975


