Intraseasonal Persistence of European Surface Temperatures

ERIK W. KOLSTAD AND STEFAN P. SOBOLOWSKI
Uni Research Climate, and Bjerknes Centre for Climate Research, Bergen, Norway

ADAM A. SCAIFE
Met Office Hadley Centre, Exeter, United Kingdom

(Manuscript received 10 January 2015, in final form 15 March 2015)

ABSTRACT

Recent periods of extreme weather in Europe, such as the cold winter of 2009/10, have caused widespread impacts and were remarkable because of their persistence. It is therefore of great interest to improve the ability to forecast such events. Weather forecasts at midlatitudes generally show low skill beyond 5–10 days, but long-range forecast skill may increase during extended tropospheric blocking episodes or perturbations of the stratospheric polar vortex, which can affect midlatitude weather for several weeks at a time. Here a simple, linear approach is used to identify previously undocumented persistence in northern European summer and winter temperature anomalies in climate model simulations, corroborated by observations and reanalysis data. For instance, temperature anomalies of at least one standard deviation above or below climatology in March were found to be about 20%–120% more likely than normal if the preceding February was anomalous by 0.5–1.5 standard deviations (with the same sign). The corresponding range for April (i.e., persistence over two months) is between 20% and 80%. The persistence is observed irrespective of the data source or driving mechanisms, and the temperature itself is a more skillful predictor of the temperatures one month ahead than the stratospheric polar vortex or the NAO and even than both factors together. The results suggest potential to conditionally improve the skill of long-range forecasts and enhance recent advancements in dynamical seasonal prediction.

1. Introduction

In recent years, severe weather anomalies in Europe have received considerable attention, mostly because of their detrimental impacts on human and natural systems but also because of the apparent persistence of weather patterns over weeks and even months. The cold winter of 2009/10 (Fereday et al. 2012) and the warm summer of 2003 (Black et al. 2004) are cases in point.

Temperature fluctuations in northern Europe (NE) are clearly associated with anomalies in the North Atlantic Oscillation (NAO; Walker 1924) and the related Northern Hemisphere (NH) annular mode (NAM; Thompson and Wallace 2001). The NAO exhibits persistence (Barnes and Hartmann 2010), but this is largely due to interannual variability (i.e., persistence over a whole season) rather than processes intrinsic to intraseasonal time scales (Keeley et al. 2009).

Persistence is also found in the stratosphere (Gerber et al. 2010) and downward propagation of NH polar stratospheric geopotential height and zonal wind anomalies (Kodera et al. 1990; Baldwin and Dunkerton 1999), which are dynamically coupled to the NAO at the surface (Ambaum and Hoskins 2002; Orsolini et al. 2011; Tripathi et al. 2015), have also been shown to impact the weather over large regions of the extratropical NH over time periods of up to two months (Baldwin and Dunkerton 2001; Thompson et al. 2002; Kolstad et al. 2010; Marshall and Scaife 2010; Mitchell et al. 2013).

In addition, anomalies in the NAO and European surface climate have been linked to Arctic sea ice variability (e.g., Yang and Christensen 2012; Liptak and Strong 2014), Eurasian snow cover (e.g., Cohen and Entekhabi 1999), and tropical phenomena such as the Madden–Julian oscillation (e.g., Cassou 2008) and El Niño events (e.g., Toniazzo and Scaife 2006). On longer time scales, slowly varying North Atlantic sea
surface temperature (SST) fluctuations also yield predictability (Keenlyside et al. 2008). The assumption is that, if large, global ensemble dynamical modeling systems can predict these mechanisms, they will also be able to predict the surface climate. This has been shown to be possible during sudden stratospheric warmings (Marshall and Scaife 2010; Sigmoid et al. 2013).

Forecast systems have indeed recently shown improved forecast skill of the NAO and seasonal near-surface temperature anomalies for Europe (Alessandri et al. 2011; Riddle et al. 2013; Baehr et al. 2014), and a forecast correlation skill score of the NAO exceeding 0.6 was recently reported (Scaife et al. 2014). However, the NAO in turn only explains a similar share of the variance of NE winter temperatures (Hurrell et al. 2003), meaning that the forecast skill of surface temperature remains limited.

While numerical models are improving and may be our best tools for long-range forecasting, they do not necessarily represent all of the low-frequency processes important for long-range prediction. It has even been suggested that there is unexploited predictability because dynamical modeling systems have signal-to-noise ratios that are too low relative to observations (Eade et al. 2014; Scaife et al. 2014). Potential sources of empirical skill are therefore valuable. In this study, we use data from preindustrial climate model runs and historical observations to detect previously undocumented month-to-month persistence in NE temperatures.

2. Data and methods

a. Data sources

We computed time series of monthly mean temperature from the Twentieth Century Reanalysis, version 2 (20CR; Compo et al. 2011), and from preindustrial control runs from 18 climate models from phase 5 of CMIP (CMIP5; Taylor et al. 2012), each with a minimum duration of circa 400 years, area weighted over the region, using land points only. The models that were used were BCC_CSM1.1(m), CanESM2, CCSM4, CESM, FIO-ESM, GFDL CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadGEM2-ES, INM-CM4.0, IPSL-CM5A-LR, MIROC-ESM, MIROC5, MPI-ESM-MR, MRI-CGCM3, and NorESM1-M (Expansions of model name acronyms are available at http://www.ametsoc.org/PubsAcronymList). We also used monthly mean temperatures for 19 Global Historical Climatology Network (GHCN; Lawrimore et al. 2011) stations in northern Europe, each with more than 140 years of data. All the temperature data were standardized so that deviations from climatology are referred to in terms of standard deviation.

b. Atmospheric indices

As a measure of the stratospheric polar vortex strength, we compute a simplified NAM index as $-1$ times the area-averaged 50-hPa geopotential height anomalies poleward of 65°N. The correlation between daily values of this index and the zonal-mean NAM index in the stratosphere exceed 0.99 (Baldwin and Thompson 2009). The NAO index was computed as the difference between area-averaged mean sea level pressure (MSLP) anomalies for two regions: 40°–50°N, 30°W–0° and 60°–70°N, 30°W–0°. These generic regions were used to allow for small variations in the spatial distribution of variance between climate models. Both indices were standardized.

c. Linear correlation and multivariate regression

We computed lag autocorrelation coefficients for time series from the data sources mentioned above. We used two-sided Student’s $t$ tests to compute the minimum absolute values that the correlation coefficients must take in order to be significantly different from zero at the 5% level. These values are straightforward to compute and only depend on the number of elements in the time series. There were 141 yr for 20CR, 203 yr on average for the GHCN stations, and 660 yr on average for the CMIP5 model runs. The lower thresholds for the absolute values of the correlation coefficients in terms of significance are 0.17, 0.14, and 0.08, respectively.

When two standardized variables $X$ and $Y$ are correlated with correlation coefficient $\rho$, it is trivial to show that the conditional mean of $Y$ given $X = x$ is $E(Y \mid X = x) = \rho x$. Similarly, the conditional variance is $\sigma^2_{Y|x} = 1 - \rho^2$. If $Z$ is the standard Gaussian distribution and we assume that both $X$ and $Y$ are normally distributed, the conditional probability of $Y$ exceeding a threshold $y$, given that $X = x$, is

$$P(Y > y \mid X = x) = P \left[ Z > \frac{(y - \rho x)}{(1 - \rho^2)^{1/2}} \right].$$

The correlation coefficient $\rho$ is the square root of the coefficient of variance $R^2$, obtained from a linear regression. This simple approach, when applied using more than one predictor, assumes that the predictors contain distinct, independent, information. This is not, strictly speaking, the case when it comes to NE surface temperatures. Thus, it is useful to employ a multivariate regression framework to evaluate the relative contributions of more than one predictor $X$ to the predictand $Y$. We used a linear multivariate regression model to evaluate the contribution of the aforementioned indices and leading-month temperature anomalies, where we
estimated the regression coefficients by means of an ordinary least squares (OLS) method. As a measure of multicollinearity between the predictors and to assess how much information is redundant because of relationships between predictors, we computed the \( R^2 \) value for each of the regression models.

d. Conditional probability ratios

We are interested in how much the probability of an anomalously warm or cold month changes if we know something about the previous month. This could be that the previous month was anomalous in temperature or in the NAO or NAM indices. A useful measure of this change is the ratio between the conditional probability (given the information that we have) and the unconditional probability. If we assume that both \( X \) and \( Y \) are normally distributed and standardized and that their correlation coefficient is \( \rho \), we can define a linear version of the probability ratio (LPR) as

\[
\text{LPR}(x, y, \rho) = \frac{P[Z > (y - \rho x)/(1 - \rho^2)^{1/2}]}{P(Z > y)}.
\]

where we inserted the expression for the conditional probability from Eq. (1). We now have a simple expression for the probability ratio, where the only variables are the correlation coefficient, the threshold value \( y \), and the value \( x \). Figure 1 shows how the LPR varies for a fixed \( y \) value of 1 when both \( x \) and \( \rho \) are allowed to vary.

An example of how to use the LPR is provided. If we were asked at the end of February to estimate the probability that the temperature anomaly during March will be greater than one standard deviation, our uninformed forecast would be 16%, because \( P(Z > 1) \approx 0.16 \). However, if we had known that the long-term correlation coefficient between monthly mean temperatures in February and March is 0.45 and that February had been warm with a standardized temperature anomaly of one standard deviation, we could use Eq. (1) to compute \( P(Y > 1 \mid X = 1) = P[Z > (1 - 0.45)/(1 - 0.45^2)^{1/2}] \approx 0.27 \), where \( X \) and \( Y \) represent the standardized temperature anomalies in February and March, respectively. The corresponding LPR is about 1.7, which can be read from Fig. 1. This means that, statistically speaking, the chances of a warm March increase by about 70% if we include the information about the February temperature, relative to the uninformed forecast. We might of course have other reliable information that indicated that March would not be warm but, based on the autocorrelation alone, an increase of 70% would be our best forecast.

Because the assumption of linearity in \( X \) and \( Y \) may not always be reasonable, we also define an empirical probability ratio (EPR) based on the actual data, where no assumptions about the distributions are necessary. We compute this ratio separately for warm and cold anomalies and for display purposes employ a scaling constant \( \varepsilon = +1 \) when we consider warm pairs of months and \( \varepsilon = -1 \) when we consider cold anomalies. The ratio is defined as

\[
\text{EPR}(x_0, x_1, y, \varepsilon) = \frac{P(\varepsilon Y > y \mid x_0 < \varepsilon X < x_1)}{P(\varepsilon Y > y)},
\]

where the probabilities are derived empirically.

3. Results

a. Linear approach

In Fig. 2a the overall January–February lag autocorrelation of 20CR near-surface temperature anomalies (i.e., deviations from monthly climatological means) is shown for land areas in the NH extratropics. The autocorrelation for the area inside the box, which we refer to as the NE domain from now on, is higher than for most other regions in the NH, although some other regions, such as parts of eastern Asia and the central United States, also have high autocorrelation. Although not shown here, the autocorrelation is relatively high in summer and spring as well, but it is low in fall; it is only in winter that the NE domain stands out compared to other regions.

The area-averaged 1-month lag autocorrelation for the NE domain is shown for all months of the year in Fig. 2b. The circles show the significant autocorrelation values for each of the three data sources. In general, the reanalysis has the highest lag correlation, whereas the mean of the stations falls between the reanalysis and the climate models. This could have been at least partly because there is a warming trend in the reanalysis and observational data (and trends introduce autocorrelation). However, the green plus signs, which show the
correlation coefficients of detrended 20CR data, indicate that the trends have little effect on the overall results. It generally did not change the ranking of the data sources. Another possible reason for the stronger lag correlation in the reanalysis and the observations than in the climate models is that there are mechanisms for persistence that are not captured by the models.

The black curve in Fig. 2b shows the mean of the three data sources shown with circles, weighted equally. The black circles indicate values for which the correlation coefficients for all the three data sources are significant. The persistence is highest in winter (February) and summer (July), with minima in spring (May) and fall (October), similar to what has been shown for the NAO (Keeley et al. 2009). There could be a simple explanation for the semiannual dips in the equinox seasons. As shown later, westerly flow into NE leads to mild conditions in winter but cool conditions in summer, so the persistence of circulation anomalies are likely to become masked during the spring and fall transition seasons.

There are indications of persistence on longer time scales in Fig. 2c, where the 2-month lag correlation for NE is shown. The time scale of this persistence is well beyond the time scale of synoptic weather, but the correlations are significant for all the data sources from December to March and from May to July, with a clear maximum for February.

In Fig. 3, the LPR values for each month are shown. They were computed by means of Eq. (2), with $X$ and $Y$ representing the first and last of the two months, respectively. The mean correlation coefficients shown along the black curves in Figs. 2b,c; a fixed threshold of $y = 1$; and $x$ values of 0.5, 1, and 1.5 were applied. In Fig. 3a, the LPR value for February is about 1.7 when $x = 1$; this was the example that was used earlier to illustrate the meaning of the LPR. The top curve in Fig. 3a shows that, if the temperature anomaly in February had been 1.5 standard deviation instead of 1, the LPR had increased from 1.7 to about 2.2. In other words, a warm anomaly in March is more than twice as likely as normal if February is unusually warm with an anomaly of at least 1.5 standard deviation. Since the Gaussian distribution is symmetric, the interpretation is the same for cold anomalies. Similarly, the bottom curve in Fig. 3a shows that, even if the temperature in February is anomalously warm (cold) by just 0.5 standard deviation, a one standard deviation warm (cold) anomaly is about 20% more likely than normal. Figure 3b shows that it is possible to make a forecast two months ahead for some months of the year; for instance, it is about 80% more likely than
normal that April will be one standard deviation warmer (colder) than normal if the temperature in February is 1.5 standard deviation higher (lower) than normal or 20% more likely with a February anomaly of 0.5 standard deviation.

b. Composite approach

The LPR values in Fig. 3 were computed under the assumption of linearity. We now derive probability ratios by means of a composite approach, where the underlying distributions of the temperatures are arbitrary. Looking at cold anomalies first, we used \( e = -1, \ y = 1, \ x_0 = 0.5, \) and \( x_1 = 1.5 \) in Eq. (3) to compute the EPR values shown in Fig. 4a. The values for each of the data sources are shown with circles, the mean of the three is shown along the black curve, and the LPR values that were shown as black circles in Fig. 3a (i.e., the probability ratios that were obtained with the linear approach) are repeated here as black squares for easy comparison. Note that the LPR computations were based on a fixed value of \( x = 1 \), while the EPR curves in Fig. 4 represent a (carefully chosen) interval of \( x \) values, so the curves are not directly comparable, although they give a good indication of the differences and similarities between the linear and empirical approaches. As the main purpose of the EPRs is to gauge the validity of the linear approach used earlier, we did not estimate the significance of the individual EPR values. In Fig. 4b, the EPRs for warm anomalies \([e = +1 \text{ in Eq. (3)}]\) are shown.

One of the crucial assumptions of the linear approximation used to compute the LPR is symmetry for cold and warm anomalies. Figure 4 indicates that there is indeed a fair amount of symmetry, in that the seasonal cycle for both warm and cold anomalies show peaks in the EPR during winter and summer, as was found with the linear approach. The magnitudes of the EPRs are also quite symmetric for warm and cold anomalies. Another issue that can break a linear approximation is the magnitude of extremes. The black squares in both panels of Fig. 4 are sufficiently similar to the values along the curve that, along with the symmetry discussed above, we may conclude that the linear approach yields a very good representation of even extreme warm and cold anomalies.

In Fig. 5, a composite approach was used to illustrate the large-scale atmospheric circulation patterns.

---

**Fig. 3.** (a) LPR values computed for the 1-month lag correlation values from the black curve in Fig. 2b, with \( y = 1 \) and with \( x \) values of 0.5, 1.0, and 1.5 for the black curves from bottom to top, respectively. As in Fig. 2b, the black dots indicate significant correlations, and these are only shown for the middle curve (but are valid for all the curves). (b) As in (a), but using the 2-month lag correlation coefficients shown on the black curve in Fig. 2c.

**Fig. 4.** (a) EPR values for \([e = -1, y = 1, x_0 = 0.5, \) and \( x_1 = 1.5 \) in Eq. (3). The values are shown with circles for the 20CR reanalysis (green), the mean of the GHCN stations (blue), and the ensemble mean of the CMIP5 models (purple), where the black curve represents the mean of the three. The black squares show the LPR values that were marked with black circles in Fig. 3a. (b) As (a), but with \( e = +1 \).
associated with warm and cold anomalies in the climate models. In this context, all the months for which the standardized NE temperature exceeded 1 (−1) standard deviation are counted as anomalously warm (cold). The ensemble-mean 50-hPa geopotential height during all the anomalously cold Januarys is shown in Fig. 5a (all the models were weighted equally). The anomaly pattern has high pressure over the polar cap: the signature of the negative phase of the NAM. Figure 5b shows the difference between the anomalously cold Januaries that were followed by an anomalously cold February and the ones that were not. The former months are more likely to project on the negative NAM pattern than the latter. Figures 5c,d show the corresponding anomaly patterns during anomalously warm Januaries, and these are roughly equal in magnitude and opposite in sign to the patterns that were found for cold anomalies. This symmetry lends further support to the use of a linear approach.

Although not shown here, the stationary wave pattern for 50-hPa geopotential height for January is a dipole consisting of a large-scale trough centered over Siberia and stretching into the Atlantic and a ridge over Alaska. The anomaly pattern in Fig. 5b (Fig. 5d) is consistent with a decrease (increase) in the amplitude of the Siberian trough.

Figures 5e–h show anomaly patterns for sea level pressure. Figure 5e shows that cold anomalies in January are, on average, associated with anomalously high pressure north of about 50øN and anomalously low pressure farther south in the Atlantic sector both at the surface and aloft. This pattern clearly projects well on the negative phase of the NAO. It is also practically identical to the pattern found for extreme cold days in northern Europe by Diao et al. (2015) and is conducive to anomalous easterly flow. Similarly, the main consequence of the difference pattern in Fig. 5f (and, although not shown, difference patterns for geopotential height at 250 and 500 hPa) is a suppression of the amplitude of the quasi-stationary Icelandic low. Figure 5g shows that the surface pattern in Fig. 5e is more or less inverted for anomalously warm Januaries, but no clear differences emerge in Fig. 5h.

Figure 6a shows that anomalously cold Julys are associated with an anomalous 50-hPa low centered over Scandinavia, which extends the summertime stationary low centered west of Greenland and suppresses the western part of the stationary high over Eurasia (stationary waves not shown). There are no significant differences between persistent and nonpersistent cold Julys (Fig. 6b). For anomalously warm Julys (Fig. 6c), the inverse of the pattern in Fig. 6a emerges, and again there is no separation between persistent and non-persistent months (Fig. 6d). The surface pattern in
Fig. 6e shows that anomalously cold Julys are associated with anomalous northerly and northwesterly near-surface flow over NE. The differences between persistent and nonpersistent months are negligible (Fig. 6f).

For warm anomalies (Fig. 6g), an opposite pattern to the one in Fig. 6e is found. Recalling from Fig. 5e (Fig. 5g) that cold (warm) anomalies during winter were due to anomalous easterly (westerly) near-surface flow, the patterns in Figs. 6e and 6g indicate that the circulation differences between winter and summer are a likely explanation for the annual cycle in the persistence, as discussed earlier. Finally, we note that Fig. 6h joins Figs. 6b,d,f in showing no significant differences between the Julys that show persistence and months that do not. This suggests that, in contrast to winter, summertime persistence from month to month is not enhanced by location-specific atmospheric anomalies.

c. Role of atmospheric circulation

As the wintertime spatial patterns in Fig. 5 resembled the spatial patterns of the NAO and the NAM, it is natural to explore the association between temperature persistence and wintertime (October–April) monthly time series of the NAO index and the NAM index. Perhaps these indices can add predictive skill if they are included in a multivariate regression model along with the NE temperature itself.

First we computed the same-month correlation coefficients between the NE temperature and the indices. Figure 7a shows that the NAO and the temperature, as well as the NAO and the NAM, are more strongly correlated than the temperature and the NAM index. Figure 7b, which shows the adjusted $R^2$ values for linear regression models where NE temperature is predicted by all possible combinations of these three indices during the previous month: the NE temperature (denoted by $T$ in the legend), the NAO index, and the NAM index. The best results are always obtained when all three indices are used as regressors, but what we are interested in is how much the skill increases when the NAO and NAM indices are added, compared to a univariate model where the NE temperature is the only regressor. The answer is that little is gained and that, of the two atmospheric indices, the NAM adds more skill than the NAO if only one of them is added. The reason for this is that the NAO index is more strongly correlated with the NE temperature than the NAM index (Fig. 7a). It is interesting to note that the NE temperature alone is a better predictor than the NAO and NAM indices together.

4. Discussion

There are two important issues to address. The first is that our results were derived from preindustrial climate model control runs, which are not influenced by external factors such as volcanoes (Briffa et al. 1998) or greenhouse gas–induced global warming. They are therefore not necessarily representative of the current, warming climate. In addition, some climate models have problems
reproducing storms moving north into the Nordic seas (Chapman and Walsh 2007) as well as North Atlantic tropospheric blocking (Scaife et al. 2010, 2011), with obvious consequences for the representation of the surface climate in Europe. Many climate models also have substantial Arctic sea ice biases (Stroeve et al. 2012) and too weak stratospheric variability (Charlton-Perez et al. 2013). However, our results are well supported by historical observations and the large sample size of the model runs allows detection of a robust signal. We also contrasted lag correlations in temperature based on detrended reanalysis data with those based on the raw data (see Figs. 2b,c), and those results indicate that the persistence occurs irrespective of the underlying warming trend over the past centuries.

The second potential concern derives from the use of consecutive monthly mean data. The persistence from one month to the next could be mostly due to strong anomalies during the transitions between months rather than mechanisms on longer time scales. Positive serial dependence is well known; for instance, when the end of one month is cold, the beginning of the next month is also likely to be cold. However, although not all of the correlations in Fig. 2c are significant, they are all positive, and this suggests that temperature anomalies tend toward persistence well beyond the synoptic time scale. Additionally, the black curve in Fig. 8a, which shows the average day-to-day evolution for all the pairs of anomalously cold months in the 20CR data (irrespective of season and where “anomalously cold” means that the standardized temperature anomaly was less than $-1$ standard deviation), does not have a minimum near the center, as would be the case if a large proportion of the cold pairs were due to synoptic-scale dependence during the transitions between months. Similarly, there is no peak near the center of the black curve for pairs of anomalously warm months in Fig. 8b.

5. Concluding remarks

Multicentennial preindustrial climate model runs, reanalysis data, and observations all indicate persistence in the month-to-month near-surface temperature in northern Europe. We believe that our results can be
used to inform long-range and seasonal forecasts for northern Europe, especially if the physical mechanisms associated with the demonstrated persistence can be better understood.

The strength of the polar vortex in the stratosphere and the phase of the near-surface NAO contribute to the persistence, but the best predictor for the temperature during a given month is the temperature itself during the immediately preceding month. Since the NAO index is strongly correlated with the same-month temperature for our study region, especially during winter, it yields less added value to a simple, linear, multivariate prediction model for temperature during the next month than the NAM index when the two indices and the temperature itself are the predictors. Also, a model with the NAO and NAM indices as predictors explains less of the next-month temperature’s variance than a univariate model with the temperature itself during the preceding month as the sole predictor. It is obviously not the temperature itself that represents the persistence; it is probably a proxy for a combination of several physical features, such as fluctuations in North Atlantic SSTs, Arctic sea ice extent, or Eurasian snow cover, perhaps on longer time scales than the intraseasonal scale examined here. This raises new and exciting possibilities for future investigations.

Acknowledgments. The authors thank three anonymous reviewers for their comments, which led to an improved manuscript. We also thank Prof. Tor Eldevik for his helpful comments on an earlier version of the paper. A. A. Scaife was supported by the Joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101) and the European Union Framework 7 SPECS project.

REFERENCES


