INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2011)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.2314

RMetS
Royal Meteorological Society

How potentially predictable is northern European winter climate a season ahead?

C. K. Folland,[a*,†] A. A. Scaife,[a,†] J. Lindesay[b] and D. B. Stephenson[c]
[a] *Met Office Hadley Centre for Climate Change, Exeter, Devon, UK*
[b] *The Australian National University, Canberra, Australia*
[c] *University of Exeter, Exeter, Devon, UK*

ABSTRACT: We estimate the potential predictability of European winter temperature using factors based on physical studies of their influences on European winter climate. These influences include sea surface temperature patterns in different oceans, major tropical volcanoes, the quasi-biennial oscillation in the tropical stratosphere, and anthropogenic climate change.

We first assess the predictive skill for winter mean temperature in northern Europe by evaluating statistical hindcasts made using multiple regression models of temperature for Europe for winter and the January–February season. We follow this up by extending the methodology to all of Europe on a 5° × 5° grid and include rainfall for completeness. These results can form the basis of practical prediction methods. However, our main aim is to develop ideas to act as a benchmark for improving the performance of dynamical climate models. Because we consider only potential predictability, many of the predictors have estimated values coincident with the winter season being forecast. However, in each case, these values are predictable on average with considerable skill in advance of the winter season. A key conclusion is that to reproduce the results of this paper, dynamical forecasting models will require a fully resolved stratosphere. Copyright © 2011 Royal Meteorological Society and British Crown copyright, the Met Office

KEY WORDS North Atlantic Oscillation; sea surface temperature; volcanoes; El Nino; QBO; seasonal forecasting; two-stage linear regression; stratosphere

Received 11 March 2010; Revised 1 December 2010; Accepted 7 February 2011

1. Introduction

Forecasts for European winter are currently carried out operationally within the Met Office using a combination of dynamical and statistical methods (Graham *et al.*, 2005; Folland *et al.*, 2006). A key component of the forecasts is the use of the GloSea forecast system which initializes a climate model with data from both the oceans and the atmosphere (Graham *et al.*, 2005). The skill of such models is currently very limited (van Oldenborgh *et al.*, 2005) especially in the extratropics. This paper is concerned with developing a multivariate method involving a number of factors that have been shown to influence European winter temperature and atmospheric circulation using physically based studies. Some of these factors are not accurately represented in the current generation of seasonal forecasts and so the potential sources of extra skill are appreciable. We only discuss potential predictability and assume (and justify below) that the winter mean numerical values of our influencing factors are skilfully predictable prior to a winter forecast.

Predictability of winter atmospheric circulation and surface climate over Europe has long been regarded as quite low. A very good review of knowledge at the beginning of the twenty first century of physical processes that might contribute to extratropical winter predictability, with emphasis on new insights into extratropical sea surface temperature (SST) influences, was published by Kushnir *et al.* (2002). They particularly emphasized a larger likely role of extratropical sea surface temperature (SST) processes than was picked up by operational models at that time. At about the time this paper was published, a statistical method of forecasting winter conditions over Europe using a mixture of extratropical and tropical SST was published by Rodwell and Folland (2002) (RF2002). This was based on studies of interactions between the North Atlantic Ocean and the atmosphere. Using the HadCM3 coupled climate model (Gordon *et al.*, 2000), evidence was found that a particular SST pattern over the Atlantic measured in May skilfully modulated the North Atlantic Oscillation (NAO) in the following winter. The pattern is influenced by prior conditions of the NAO and is called the North Atlantic tripole. The tripole is known to feed back on the NAO (e.g. Rodwell *et al.*, 1999). This May pattern tends to be

* Correspondence to: C. K. Folland, Met Office Hadley Centre for Climate Change, FitzRoy Road, Exeter, Devon, EX1 3PB, UK.
E-mail: chris.folland@metoffice.gov.uk

† The contribution of these authors was written in the course of their employment at the Met Office, UK, and is published with the permission of the Controller of HMSO and the Queen's Printer for Scotland.

present throughout the upper subsurface ocean in phase with the SSTs, and the subsurface component tends to be preserved under the near surface summer thermocline that develops from May. The thermocline is destroyed around the beginning of the following winter and the underlying tripole pattern often reemerges.

The RF2002 forecasting method was limited to predicting the NAO. Here we use indices of the May SST tripole derived by RF2002 to help predict north European temperature rather than the NAO. This is reasonable because there is a strong relationship between the winter NAO and north European temperature (Hurrell, 1995; Junge and Stephenson, 2003; Scaife *et al.*, 2005). Thus, a positive (westerly) NAO pattern tends to create positive north European temperature anomalies. However, we have not made any attempt to optimize a version of this SST pattern that is the most highly correlated with north European temperature.

The influence of the El Nino-Southern Oscillation (ENSO) on winter North Atlantic atmospheric circulation has only recently been partly elucidated. Early work (e.g. Fraedrich, 1992) indicated weak effects on North Atlantic atmospheric circulation and suggested that El Nino gave a higher probability of a negative, blocked, index of the NAO. In a detailed analysis of individual winters over the last century, Toniazzo and Scaife (2006) explained these weak results. They found that moderate El Ninos, as measured by Nino3, had a consistently strong effect on the NAO, making it more negative than usual, in qualitative agreement with previous results. However, they discovered that strong El Ninos gave a different atmospheric circulation pattern that did not project significantly onto the NAO and have a considerably smaller effect on European temperature. They also provided a mechanism for the strong El Nino effect on winter North Atlantic atmospheric circulation; this causes a Rossby wave forcing from the tropics through the troposphere. Recently, Ineson and Scaife (2009) and Cagnazzo and Manzini (2009) have confirmed the influence observed during moderate El Ninos on the NAO; extratropical planetary waves drive downward propagating westward wind anomalies from the stratosphere. Both studies found that the effects of ENSO on North Atlantic atmospheric circulation are strongest in later winter, essentially January–March. So for the full conventional winter as defined here, December–February, we should expect a weaker effect on European temperature with a slightly stronger effect in January–February.

Evidence for La Nina influences, generally favouring a positive NAO for stronger La Ninas, comes from the observational analysis of atmospheric circulation and La Nina, e.g. Moron and Gouriand (MG2004). They show that the strongest 30 La Ninas over the period 1873–1996 tend to create a positive NAO in January and February, though in December the pattern is somewhat different, with high pressure near the UK. The effects appear weaker than for moderate El Ninos but do reflect the weak modelling results of Davies *et al.* (1997) concerning La Nina influences on North Atlantic atmospheric circulation. MG 2004 found relatively little difference between the 10, 20, and 30 strongest La Ninas. The different influences of El Nino and La Nina in the North Atlantic region in November and December compared to January and February have recently been confirmed in an observational cluster analysis by Fereday *et al.* (2008). El Nino and La Nina SSTs are regularly predicted in seasonal forecasts and skill is high in advance of the Northern Hemisphere winter season, as shown in dynamical seasonal forecast models (Jin *et al.*, 2008). Our use of a contemporaneous ENSO predictor with the winter or January–February periods is realistic for predictability studies.

Observational results suggest that major tropical volcanoes cause a strong positive NAO in the winter following the eruption, and thereby affect European winter temperature (e.g. Robock and Mao, 1995; Fischer *et al.*, 2007). The effect in winter is counter-intuitive in that such volcanoes cause a winter warming signal over northern Europe due to advective warming from the induced positive, westerly, phase of the NAO. This overwhelms the direct surface cooling effects of the stratospheric aerosols created by the volcano (Stenchikov *et al.*, 1998). We have used lists of major tropical volcanic eruptions compiled by Robock and Mao (1995), Fischer *et al.* (2007) and Stenchikov *et al.* (2006) to select an estimate of those European winters likely to be most affected. The physics of this dynamical volcanic effect continues to be debated after initial apparently skilful attempts to simulate it (e.g. Graf *et al.* (1994) have not always been successfully repeated. An excellent review is provided by Robock (2000). We have not included extratropical volcanoes as their effects on European temperature are weaker (Robock and Mao, 1995). Evidence from recent climate modelling research by Marshall *et al.* (2009) shows that skilful modelling of volcanic effects depends considerably on the initialisation of observed conditions at the beginning of winter. This, of course, is standard procedure in seasonal prediction. For the large tropical volcanic eruptions starting with that of Krakatau (in 1883) used here, it will be known quite quickly whether such an eruption has created substantial amounts of aerosol in the stratosphere. This is because routine monitoring of volcanic eruptions is carried out, for example, through the Smithsonian Global Volcanism Program http://www.volcano.si.edu. The Program is linked to data from nine Volcanic Ash Advisory Centres: http://www.meteo.fr/vaac/. (both URLs as at Nov 2010).

The physics of the effects of the quasi-biennial oscillation (QBO, Ebdon and Veryard, 1961) on climate, and the mechanisms of the QBO, are extensively discussed by Baldwin *et al.* (2001). A statistically significant effect on surface atmospheric circulation in January was first seen in observations by Ebdon (1975), which can be interpreted as an effect on the AO and the NAO, followed by broadly similar results at 50 hPa (Holton and Tan, 1980). Later, it was shown that tropospheric anomalies lagged the stratospheric anomalies by 2–3 weeks (Baldwin and Dunkerton, 1999). As indicated by Ebdon, the NAO/AO tends to be in a negative phase when the QBO is easterly (Thompson *et al.*, 2002). In this phase, there tends to be widespread winter cold anomalies and extremes throughout much of Europe, Asia, and North America (Hurrell, 1995; Thompson and Wallace, 2001). Studies of existing seasonal forecasts suggest that the QBO could be an important source of missing predictability in current forecast systems (Boer and Hamilton 2008). Model hindcast experiments also provide evidence that the QBO in the lower stratosphere influences the NAO and thus European temperature and that this effect is partly predictable (Marshall and Scaife, 2009). These studies showed statistically significant and practically useful differences in modelled north European temperature between sets of winters with easterly and westerly phases of the QBO. Surface temperature differences exceeded 1.5 °C in some areas.

Finally, although the signal is only now clearly emerging, it is necessary to allow for anthropogenic warming over Europe. This is not straightforward as nearby SSTs and European temperatures have been influenced by interdecadal changes in the NAO which may be natural in origin. Thus the winter NAO index increased strongly between about 1965 and 1995 and strongly influenced European surface temperature (Scaife *et al.*, 2005, 2008), though the NAO index has reduced since then (Trenberth and Jones, 2007). Our approach has been to use the HadCM3 climate model forced with all major anthropogenic influences including sulphate aerosols (similar to experiments described by Stott *et al.*, 2000) to estimate the net anthropogenic warming effects over Europe, deliberately choosing a model that failed to pick up winter NAO trends in its simulations over the last 50 years or more. This deficiency in HadCM3 has been turned to good effect to estimate the European anthropogenic warming signal in detail across the continent in the absence of a strong circulation change. The overall effect is not strictly monotonic as the model picks up some cooling in the mid-twentieth century due to a sharp increase of anthropogenic aerosols over Europe not offset by the then slow anthropogenic greenhouse gas-induced warming (van Oldenborgh *et al.*, 2009). Since then, European aerosols have declined while anthropogenic greenhouse gas forcing has accelerated (van Oldenborgh *et al.*, 2009). Modelled recent winter warming also increases eastwards, away from the moderating influence of the North Atlantic Ocean. However, because the average of six model integrations used still contains some significant effects of internal interdecadal to decadal variability over Europe, we have fitted a quadratic curve to the data. This simplification is not a serious one, remembering we are not trying to recreate the observed winter temperature observations but rather a representation of them without the substantial observed interdecadal NAO effects. However, the fitted curve may underestimate the slowing of warming between the 1940s and 1960s due to increasing sulphate aerosol effects, so this needs to be borne in mind. Nevertheless, as can be seen from the results presented in Section 3 below, this is not a serious deficiency. The model also picks up some European cooling between 1861 and 1910. The latter is similar in shape to the small decline in Northern Hemisphere surface temperatures over the latter part of this period (Brohan *et al.*, 2006) and to a marked drop in annual and winter half-year temperatures between these dates in central Europe described by Böhm *et al.* (2010) using new methods of bias adjustment to surface temperature data. Folland (2005), showing the results of using the older surface temperature data of Jones *et al.* (1999), indicates a rather smaller observed drop for Europe as a whole between 1872 and 1910. Overall, we regard the observed data as adequately supporting the modelled late nineteenth century cooling. We are also encouraged to use these model results because, globally, HadCM3 simulates the observed interdecadal global mean surface temperature variations over the last 140 years very well (Stott *et al.*, 2000; IPCC 2007).

We use these model results to choose suitable data sets to create forecasts of European winter temperature over the whole continent. Our emphasis is on the well attested atmospheric circulation relationships described above. Thus, although we carry out some statistical tests on the multivariate methods we create for predicting north European temperature, they only provide a guide. Note that Fletcher and Saunders (2006), and references therein, have explored empirical correlations between the NAO and various predictors in the preceding summer half year and provide some observational hypotheses for their results. Except for the RF2002 predictor, these predictors are not independently verified with numerical model experiments so are not included below and indeed differ greatly. However, in the discussion of our results in Section 5, we review one of their predictors, Northern Hemisphere snow cover.

2. Development of the methods

2.1. Data

We use the HadCRUT3v (Brohan *et al.*, 2006) dataset of surface temperature. The data are available on a 5° × 5° latitude–longitude grid. We use northern Europe area-averaged temperatures over the region 15 °W, 45 °N; 30 °E, 65 °N as a test-bed for our methods. Temperatures are calculated as anomalies from a 1971–2000 average. Later in the paper a complete 5° × 5° grid covering Europe is used for forecasts of the observed pattern of surface air temperature over the whole continent. Winters, here December to February, are allocated the year in which January falls, and winters 1876–2008 or 1954–2008 are used in the regression equations described below. The May North Atlantic SST tripole index, used here for 1954–2008 is that calculated by RF2002 using the GISST3.0 dataset, updated and improved from that in Parker *et al.* (1995). Standardized indices of Nino 3.4 SST have been taken from HadISST (Rayner *et al.*, 2003) back to1876 because there is evidence of reliability of

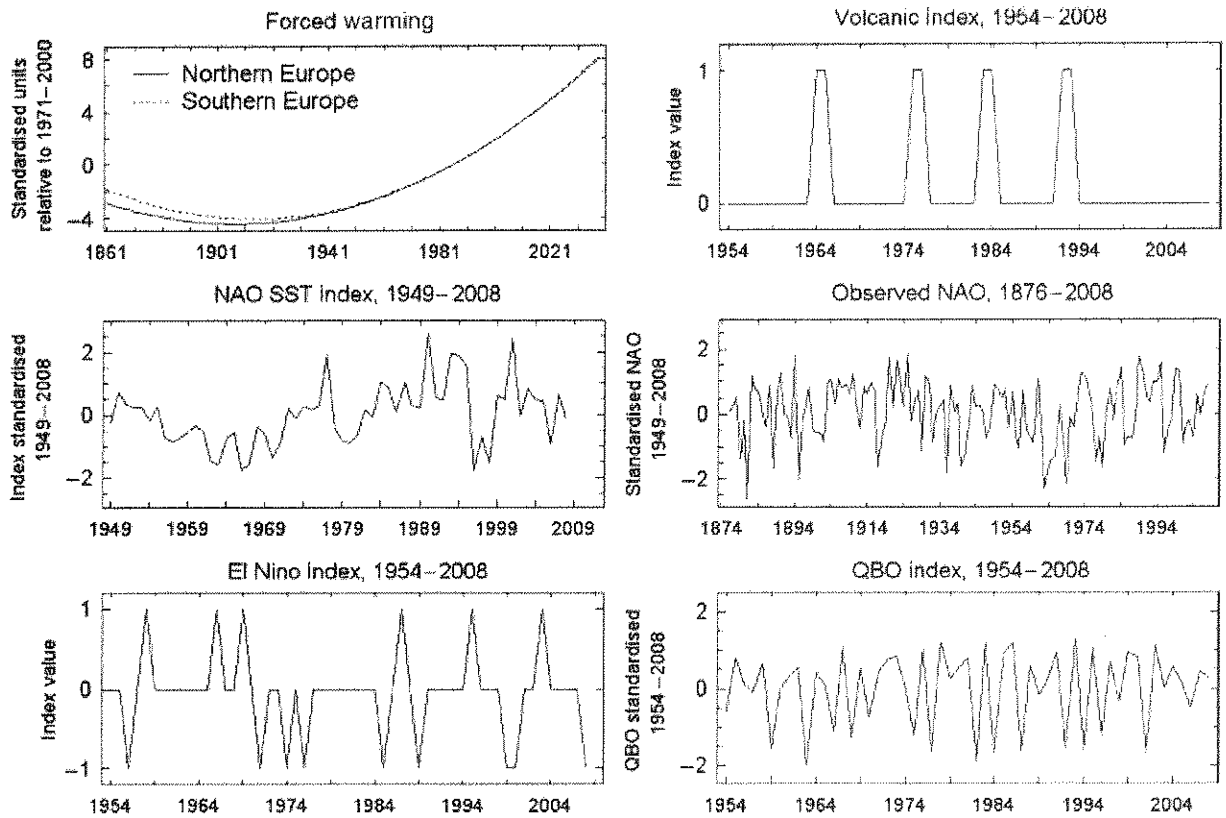

Figure 1. Predictor data time series including the complete series of smoothed modelled surface temperatures for northern and southern Europe.

this index back to that date as discussed for Nino3 in Folland and Karl (2001). Because climate warming has been very small in the Nino 3.4 region according to HadCRUT3 and HadISST, we do not adjust the raw data for climate trends. A test case where long-term warming was removed showed only small changes in the standardized index in any year.

Tropical volcanic data are based on the list of Stenchikov *et al.* (2006). We assume, following Robock and Mao (1995), that the influence of tropical volcanoes on Northern Hemisphere winter climate occurs in the first and second winters after the eruption. In addition, in 1902 there was an earlier major tropical eruption in May 1902 (Soufrière), which argues further for choosing the following winters 1903 and 1904 rather than 1904 and 1905. The first tropical volcano used is Krakatau, 1883, and the last is Pinatubo, 1991. Because differences in the effects of these volcanoes on European atmospheric circulation and temperature are not currently well understood, we use a simple index to identify these volcanic effects. Otherwise, the index is set to zero when the volcanoes are not operative.

The QBO data are those of the Free University of Berlin, based on an update of Naujokat (1986) and taken from http://www.geo.fu-berlin.de/met/ag/strat/produkte/qbo/index.html (as at November 2010) though few data are available after 2008. We use 30 hPa equatorial zonal wind speed following the dynamical modelling results of Marshall and Scaife (2009). Hamilton (1998), in a dynamical modelling study of the QBO, noted that quite modest amplitudes of the QBO at 40hPa affected the Arctic stratospheric circulation, negative QBO values (corresponding to tropical easterlies at 40 hPa) weakening that circulation and positive QBO values strengthening it. So we have used a continuous QBO index in the form of the 30 hPa near-equatorial wind speed. Several months before a given winter, QBO values can usually be well predicted for the winter average using the method described by Gray *et al.* (1992).

Finally, the estimates of European warming due to anthropogenic effects have been taken from the mean of an ensemble of six HadCM3 integrations with all forcings as discussed in Section 1, to which smoothing has been applied to remove spurious sampling fluctuations of temperature on time scales of less than about three decades. As described in Section 1, this coupled climate model had no significant modelled trends in the NAO or ENSO, so its estimates of winter warming over Europe are virtually independent of the effects long-term changes in the NAO pattern. Such estimates have been made for every $5^{\circ} \times 5^{\circ}$ area of Europe for each winter. The forced warming estimates for the north European region and slightly different values for the south European region are the weighted average of constituent $5^{\circ} \times 5^{\circ}$ boxes where the weights are the relative areas of these $5^{\circ} \times 5^{\circ}$ boxes. Figure 1 shows all the predictor time series including the anthropogenically forced climate warming estimates separately for northern and southern Europe.

2.2. Formulation and computation of linear regression equations

The basic methodology uses ordinary least squares multiple regression where the predictor data from model results are used to predict north European temperature anomalies. We have checked this method against a forward stepwise selection of variable method but prefer the former method based on physical understanding, even where the length of the data is insufficient for statistical significance. We present some statistical tests but do not use them to weed out predictors. An example that shows the need for a physical approach is climate warming. Methods based on data finishing at winter 1998 (not shown) show considerably less significance for climate warming, or even no significance at the 5% level. Data ending in 2008 always show 5% significance, but we consider the background climate change effects are real throughout as continental-scale detection and attribution studies strongly indicate this (Hegerl and Zwiers, 2007). The climate change predictor is inserted as a continuous temperature variable based on the model estimates of north European temperature change.

The El Nino predictor index is a discontinuous variable set to 0 if it is less than ± 1 standard deviation of its variability over 1876–2008. It is set to -1 for all more negative values and to 1 for values between one and 1.75 standard deviations (Toniazzo and Scaife 2006). Above that value the index is again set to zero. The observational and dynamical studies quoted above and longer observational data (Brönnimann *et al.*, 2007) support these choices.

Finally, the volcanic index is set to one for the two winters following large tropical volcanic eruptions, otherwise it is set to zero. The tripole SST index of RF02 has values that vary continuously. Note that no predictor would have been included if its calculated sign was inconsistent with the direction of its expected physical effect, but this is not observed.

One limitation is that the tripole SST index is only available since winter 1949 (Dec 1948–Feb 1949). This is because the index was calculated from a maximum covariance analysis between SST and 500 hPa height (RF02) and the latter data are only available since 1948. However, a greater limitation is that a reasonably homogenous QBO index is only available since 1954, though the remaining indices are available since 1876. To exploit the explanatory variables available before 1954, we perform ordinary least squares regression in two stages: we first regress temperatures on the longer 1876–2008 series (volcanic index, Nino3.4 index and the climate warming series) and then we regress the residuals from this regression model onto the more recently available 1954–2008 SST and QBO series. The final statistical model hindcasts for 1954–2008 are then constructed as the sum of the two linear predictors. These two-stage predictions are compared with the skill of regression predictions developed over 1954–2008 alone. Note that we only carried out this two-stage regression procedure for results involving average north European temperature.

We noted above that ENSO relationships tend to change between December and later in winter. So we have also investigated prediction equations for the last two months alone where we expect a stronger overall ENSO relationship and perhaps some intraseasonal predictability. However, a two-month period suffers from greater internal climate variability. The NAO underlies a considerable part of the various predictor–predictand relationships. Its effects on temperature anomalies are generally of opposite sign in southern compared to northern Europe, so this pattern will appear in many of our mapped results in Section 4.

To avoid a large estimation uncertainty caused by co-linearity, linear regression predictors should not be highly correlated with one another. The only moderately correlated predictors over 1954–2008 for north European temperature are climate warming and the tripole SST index ($R = 0.41$). This is due to the climate warming trend being positively correlated with a general increase in the North Atlantic SST index (and the winter NAO) from 1965 to 1995, with only a moderate decrease in the SST index since. However, they are still not correlated at the 5% level due to the strong serial correlation of climate warming. This lack of correlation between explanatory variables allows us to carry out the regression in two stages to take account of the information in the explanatory variables available before 1954.

3. Results–north European winter temperature

3.1. December–February one-stage linear regression

For each predictor of DJF European mean temperature, Table I lists linear regression coefficients and their nominal significance for winters confined to 1954–2008 for the model with all predictors. The time series of the predictors are shown in Fig 1. The standardized values indicate the relative importance of the predictors. Values significant at the 5% level or better are starred. The multiple correlation has a value of $R = 0.53$. These relationships have been cross-validated to hindcast each value over 1954–2008, in turn. To achieve this, all data with significant serial correlations with the year being assessed have been removed. Here, the relevant year being assessed together with the two years on either side of it are omitted from both predictors and predictand. The cross-validated correlation reduces R to 0.36. Thus the true variance explained is 13% rather than the 28% derived from the raw multiple correlation. This can be explained as follows. Fitting a set of predictors to a dataset without a cross-validation step gives rise to the phenomenon of 'artificial skill'. Artificial skill increases firstly as the number of predictors increases. This component can be estimated using the 'adjusted multiple correlation' (Draper and Smith, 1998) which decreases relatively further compared to the raw multiple correlation as the number of predictors increases. Here, using 5 predictors, the adjusted multiple correlation for all the data over 1954–2008 falls to 0.45 or 20% of the variance explained. However, this is still not an unbiased measure of the skill of *future* forecasts independent of the data we have used. This is because the full structure of a time series is not picked up with limited data, even though the series may be stationary in the long term. A good estimate of true skill can be made using cross-validation, though it can be difficult to apply cleanly (von Storch and Zwiers, 1999). Cross-validation creates (in principle) a time series uncorrelated with the year being predicted which is then used to create a regression-based hindcast. A separate forecast model is thus created for each year. If data were much more numerous, separate 'training' and 'testing' periods might be better, and may pick up a further loss of skill due to non-stationarity. Here, we have limited data and cannot use this method. Nevertheless, cross-validation still shows that the reduction in variance explained from 28 to 13% is substantially greater than that obtained from discounting the effects of artificial skill within the fixed historical time series alone. It should be a good estimate of the predictability of current and near future winters using the factors we have chosen. In the more distant future, climate warming may change these relationships.

Returning to Table I, factors 2–4, when positive, cause warming of north European temperature through an enhanced frequency of westerlies, while a positive Nino 3.4 index, associated with El Nino, causes cooling through an enhanced frequency of blocking. Table I also shows how important it is to include climate warming. Table I confirms that the May SST index of RF2002, though designed to predict the North Atlantic Oscillation, is still very useful for influencing north European temperature. Although only significant at the 8% level over this period, its standardized regression coefficient is the second highest at 0.25 and close to that of climate warming, the most important factor.

We have compared this method with forward stepwise regression (e.g. Afifi and Azen, 1979) over the same period. The results (not shown) are quite similar except that the volcanic and El Nino Indices are not entered at the chosen level of significance, around the 5–10% level. We conclude that climate warming, the NAO May SST factor (here marginally the most significant factor, now at the 5% level) and, less strongly but interestingly, QBO equatorial winds at 30 hPa are the key predictive factors for north European winter temperature on average.

Table I. Statistics of one-stage linear regression methods for north European temperature anomalies, winters 1954–2008.

Regression parameter	Standardized regression coefficient	Significance regression (fraction)
Climate warming	0.27	0.05
North Atlantic SST index	0.25	0.08
QBO index	0.20	0.12
Volcanic index	0.14	0.25
Nino 3.4 index	−0.11	0.40

Volcanic effects and El Nino or La Nina add only a little additional variance statistically over the period sampled because in many years they are inactive. However, they are important in those years when they do occur. Given more data and better ways of representing their effects on north European temperature they are likely to become statistically significant predictors.

3.2. Reconstructing the relative importance of each factor over 1954–2008

Figure 2(a) shows the one-stage regression time series for 1954–2008 based on all factors (red), observed north European temperature (black), and the time series of each individual predictor contribution. This gives their relative contributions to north European temperature anomalies in each year. Of particular interest are predictor contributions for some extreme years. The very small linear regression constant of −0.10 has to be accounted for when discussing this breakdown. Since it is a constant and the slope of the climatic warming factor is unaffected, we have incorporated the constant in the climatic warming curve in the diagram. The coldest winter in the series, 1962–1963, is reproduced with the correct sign and about half the observed amplitude; it is also the second coldest simulated winter in the period, similar to the coldest simulated, 1966. The QBO, North Atlantic SST index and the cooler climate of the time combine in nearly equal measure to make a very cold north European 1962–1963 winter likely. This leaves half the amplitude of 1962–1963 due to internal variability or unexplained factors. In more recent years, the cold winter 1995–1996 is contributed to by the North Atlantic SST index and a negative QBO index, though again only half the amplitude of the anomaly is explained. Very warm years seem generally less well explained, though the largest positive value in the series of the North Atlantic SST index contributes appreciably to the very warm winter 1990, with no contribution from the other factors. The biggest single contribution to the last two very warm winters, 2007 and 2008, comes from climate warming with a smaller warm contribution from La Nina in 2008. However, much of the amplitude of these two winters is not explained by our predictors.

3.3. Extended period 1876–2008

Table II lists comparable results to the first three columns of Table I. The correlation between the results of the regression equation created by the two-stage process and the observations over 1954–2008 is, by definition, the multiple regression coefficient for that period for the two-stage process. This has a very similar value of 0.52 to the one-stage regression. Thus, two-stage regression does not explain extra variance but is consistent with one-stage regression. However it is useful to show hindcast data compared to the original data (Figure 2(b)) back to 1876. This shows that the climate warming factor is quite realistic when compared to the behaviour of the observations over the whole period and not surprisingly is very significant at near the 0.5% level over this longer period. The high-frequency signals from volcanoes are relatively small since 1876 but also have better statistical significance in Table II than in Table I, though the ENSO signal does not, possibly because of poorer early SST data. The multiple correlation value however is only 0.29 compared with 0.53 for the period 1954–2008, showing the impact on hindcasting interannual variability of not including the QBO and NAO SST predictors.

It is useful to compare these results with what may be much closer to the best possible hindcasts. These have been calculated by regressing an observed NAO index on winter north European temperature over 1876–2008. The observed NAO index uses the difference in PMSL between Ponta Delgada, Azores, and Stykkisholmur, Iceland (Figure 2 for this NAO time series) rather than the Rodwell and Folland 500 hPa NAO in order to estimate the best possible relationship between north

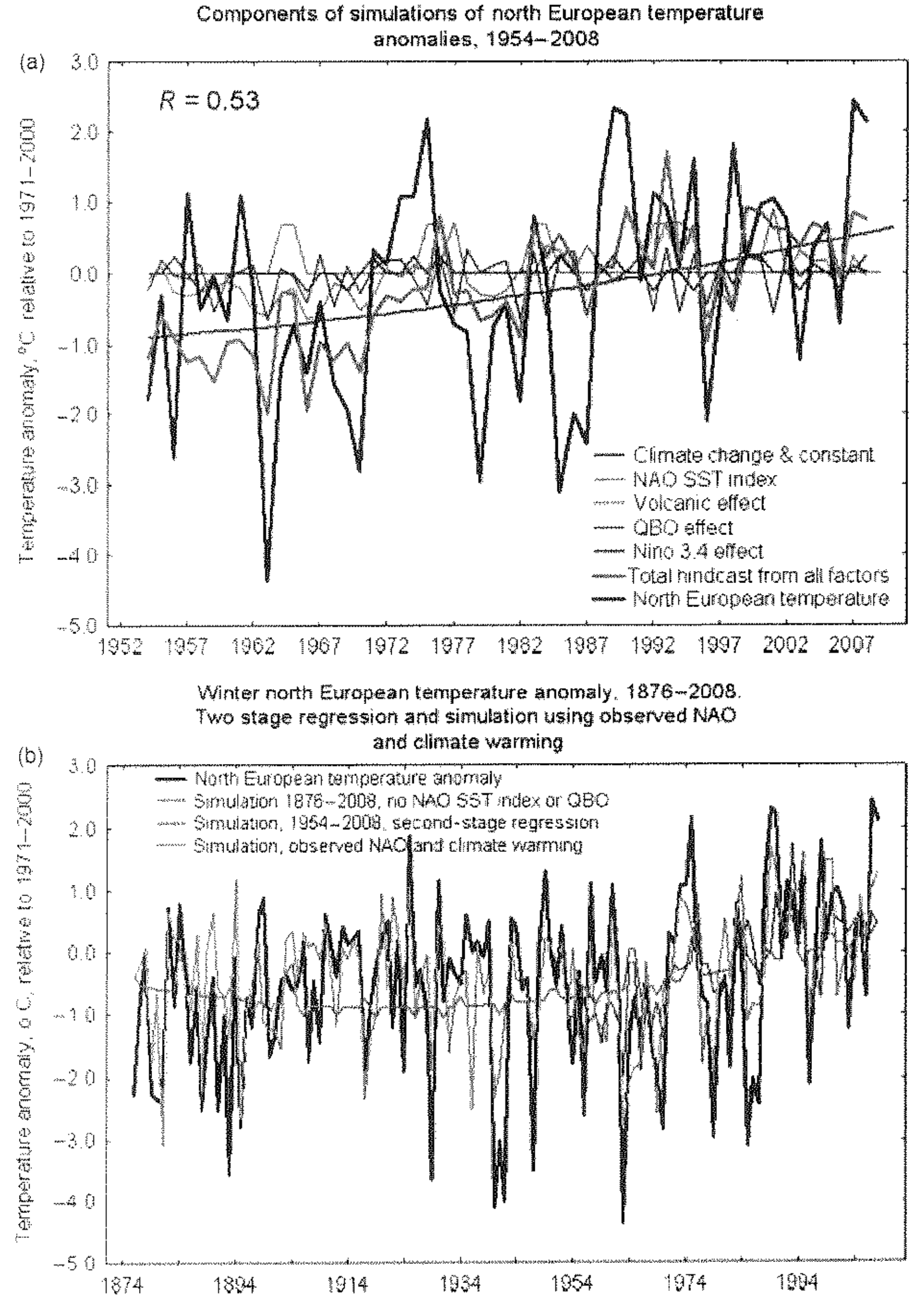

Figure 2. (a) Simulation (no cross-validation) of north European temperature anomalies, 1954–2008, using one-stage regression. The contribution to the simulated temperature of each predictor is shown together with observed north European temperature (black). (b) Two-stage regression simulations of winter mean north European temperature anomalies, 1876–2008. The red line is the simulation using predictors from 1876, the green line using the observed NAO and climate warming and the blue line using predictors available. (c) Cross-validated multiple regression hindcasts, 1954–2008, for north European winter mean temperature anomalies. Each hindcast is based on a set of predictors calculated from all data except the winter being hindcast and the two years before and after. d) Residual errors (observed north European temperature minus hindcast) from the cross-validated model and the model including the observed NAO, both for 1954–2008. The latter is based on a regression over the whole period 1876–2008. Note large failures such as 1985.

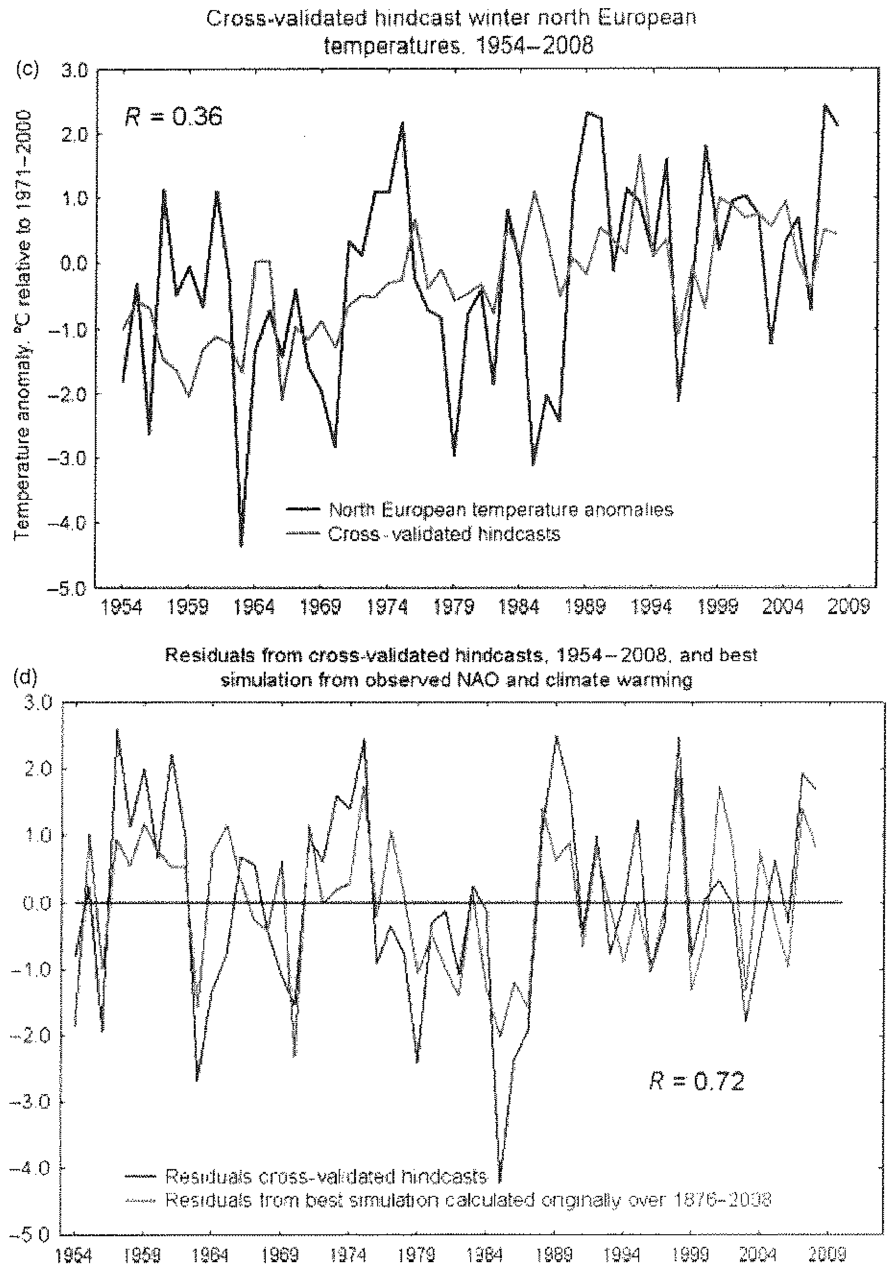

Figure 2. (*Continued*).

Table II. Statistics of two-stage linear regression method for north European temperature anomalies, winters 1876–2008.

Regression parameter	Standardized regression coefficient	Significance regression (fraction)
Climate warming	0.24	0.006 (1876–2008)
North Atlantic SST index	0.30	0.02 (1954–2008)
QBO index	0.23	0.08 (1954–2008)
Volcanic index	0.13	0.12 (1876–2008)
Nino 3.4 index	−0.06	0.50 (1876–2008)

European temperature and the NAO over a much longer period. The fitted correlation over 1876–2008 gives $R = 0.71$. The best possible regression also needs to include the climate warming signal. When this is included, $R = 0.74$ and the resulting regression curve is shown in green in Figure 2(b). This indicates that a perfect knowledge of the NAO plus climatic warming (which is only estimated in this paper) would explain just over 50% of the variance of north European winter temperature variability. However, some failures and a reduced variance overall show that the NAO is by no means a complete description of the influences on north European winter temperature. In a few winters over 1876–2008, the observed NAO can be seen to be far from an adequate predictor. There is little evidence in these diagrams that any deficiencies in our climate warming predictor are having much effect. The other climatic influences on north European winter temperature may or may not be predictable, of course.

3.4. Cross-validated hindcasts

Figure 2(c) shows the cross-validated time series from 1954 to 2008 from the one-stage method. Some extreme winters like 1963 (cold), 1990 (warm) and 1996 (cold) are simulated with some skill though with less skill than for the fitted hindcast model. The general warming over the period is picked up well. There is some evidence of a cold bias in the hindcasts before 1975, possibly larger than can be explained by any cold bias in the climate change predictor alone.

Figure 2(d) shows the residual errors for the cross-validated hindcasts for 1954–2008 (blue) compared to those from the hindcasts using the observed NAO and climatic warming for the same period (red). The residual errors are defined as the observed north European temperature minus the hindcasts. The cross-validated residual errors are strikingly well correlated with those from the observed NAO model with $R = 0.72$, though the standard deviation of the cross-validated model errors at 1.47 °C is naturally substantially greater over the same period than those of the model using the observed NAO at 1.06 °C. Figure 2(d) indicates that an important component of the errors in both models must arise from aspects of the winter atmospheric circulation or SSTs that influence north European winter temperature that are not picked up by factors directly related to the NAO or the climate warming signal. Measurement noise in the data will, of course, contribute to the errors as will internal climate variability (Section 5). Nevertheless, Figure 2(d) indicates that predictive factors not included in this paper are likely to be worth investigating.

Finally, we compare the cross-validated winter hindcasts over 1954–2008 with persistence. The correlation between the observations and persisted values from one winter before is 0.31. Therefore, the cross-validated hindcast system correlation of 0.36 just beats persistence. Another forecast strategy sometimes suggested is the persistence of the average of the previous two years though this is disadvantaged here by the QBO signal, making it more likely that our forecast system will beat this skill level. Indeed the correlation is 0.27, lower than for the persistence one year ahead.

3.5. January–February one-stage linear regression

The time series of JF north European temperature is unsurprisingly quite similar to that of DJF. If the DJF series consisted of the JF series and an uncorrelated D series, the correlation between JF and DJF would be 0.67, assuming all months had equal length. The actual correlations are 0.94 (1876–2008) and 0.96 (1954–2008), showing substantial persistence (due probably to interannual forcing factors) from D to JF. In fact, Keeley *et al.* (2009) estimate from winter NAO data that it is likely, in agreement with this result, that a substantial fraction, perhaps 70%, of the interannual variability of the winter NAO is externally forced which would contribute to persistence of north European surface temperature through the winter. We use the same values of the predictors as for DJF because there is little extra skill in estimating these for JF alone, even close to the JF period. Table III shows the results in the same format as Table I. The multiple correlation value is slightly less at 0.50, and the cross-validated value of R is now 0.31 calculated in the same way as for December to February. Thus, just under 10% of the JF north European temperature variance is explained by one-stage linear regression over 1954–2008. A stepwise regression done as for DJF gives only climate warming as a significant predictor. So despite the similarity in the JF and DJF datasets, predictors other than climate warming have lower predictive power, particularly the North Atlantic SST index.

Table III. Statistics of one-stage linear regression methods for north European temperature anomalies, January and February 1954–2008.

Regression parameter	Standardized regression coefficient	Statistical significance (fraction)
Climate warming	0.35	0.02
North Atlantic SST index in previous May	0.14	0.33
QBO index	0.14	0.27
Volcanic index	0.15	0.25
Nino 3.4 index	−0.12	0.34

Thus, the most important change in JF from DJF is the loss of the skill of the NAO SST predictor, which suggests that it has more skill for early than late winter temperature. This would be expected as the re-emerging SST pattern from the previous May will be progressively modified during the following winter. Compared to Table I, the ENSO predictor is a little stronger in Table III as expected from its more consistent effects on European climate in later winter (e.g. Fereday *et al.*, 2008), though significance remains low. Climate warming has clearly become an important predictive factor in both JF and DJF, an important result for European winter temperature forecasting that agrees with the results of Liniger *et al.* (2007) and Boer (2009).

3.6. January–February two-stage linear regression

Table IV shows the results. The January–February multiple correlation between the simulated values and the observations is again slightly lower than for the one-stage regression at 0.48. Otherwise the relative weights of the predictors are similar, though the significance of the volcanic index is greater, as expected from the longer data series. This is also seen for DJF in Tables I and II.

3.7. January–February cross-validated hindcasts

We do not show a figure but recap that the cross-validated correlation skill is, as expected, lower in the shorter period January–February than in winter at 0.31. Nevertheless, this value is still significant despite the

Table IV. Statistics of two-stage linear regression method for north European temperature anomalies, January and February mean, 1876–2008.

Regression parameter	Standardized regression coefficient	Statistical significance (fraction)
Climate warming	0.23	0.008 (1876–2008)
North Atlantic SST index in previous May	0.22	0.10 (1954–2008)
QBO index	0.18	0.18 (1954–2008)
Volcanic index	0.14	0.11 (1876–2008)
Nino 3.4 index	−0.06	0.47 (1876–2008)

substantial loss of skill from the May North Atlantic SST predictor. Note that this factor, appropriately, has lower weight than in DJF, so it really does have less influence on the hindcasts. Again, we compare the cross-validated JF hindcasts over 1954–2008 with persistence. The correlation between the observations and persisted values from one winter earlier is 0.33. Therefore, the hindcast system correlation of 0.31 does not beat persistence. Persistence of the average of the previous two years gives a correlation of 0.30, again comparable to the cross-validated hindcasts.

4. Results–spatially varying predictability of temperature and rainfall

A similar regression methodology is now used to construct spatially varying predictions of European winter temperature and European winter rainfall. For temperature, we use the CRUTEM3 dataset on a 5° × 5° grid (Brohan *et al.*, 2006), and for rainfall, the dataset of New *et al.* (2000). The same factors of anthropogenic warming, ENSO, volcanoes, and the QBO are used to produce a hindcast set of European winter temperature anomalies for the period 1954–2008. The method first identifies the regression pattern associated with each of the predictors by regressing observed values of the 500 hPa NAO used by RF2002, the occurrence of a moderate El Nino or moderate to strong La Nina, the QBO and the occurrence of volcanic eruptions, as described above, against DJF surface temperature anomalies and DJF rainfall anomalies. Anomalies are calculated relative to a climatology for 1971–2000. The observed regression patterns are weighted by the predicted values for each of the predictors for a given year. In this way, the spatially varying method mirrors the hindcasts for the simple area mean predictions above. Thus, for the NAO SST method the pattern weighting is the loading of the May predictor pattern in the observed May tripole SST anomalies and for a moderate El Nino and moderate or strong La Nina index we take the corresponding average temperature or rainfall pattern for these indices (otherwise it is absent). For the QBO we take the observed QBO anomaly, assume this is predictable and weight it in a comparable way to the NAO index. For climate change we use the comparable regression weights, and for volcanic eruptions the pattern is fixed and is applied if the volcanic index is regarded active as in Section 3. The pattern weighting can of course be positive or negative for the NAO method, the El Nino/La Nina method, the QBO and for climate change. It is always the same sign for a volcano.

The patterns are cross-validated by removing the year in question when calculating the regression patterns for each of the predictors. This allowed us to verify that the predictor patterns varied only slightly over the 55 winters due to the addition or removal of single events. We have not modified the cross-validation method used for the map patterns to use the more severe method of Section 3. This is because it was found that omitting just the year in question, or the two years before and after in addition, made almost no difference to cross-validated skill in Section 3. This was found to be $R = 0.36$ in winter for both cross-validation methods.

4.1. December–February

Regression patterns for each of the predictors are shown in Figures 3–7. Anthropogenic climate change generates warming throughout the European region in these estimates (Figure 3) and is strongest in the continental interior. Note that Figure 3 is again based on the HadCM3 results and that the standardized values are based on the period 1861–2039 for both temperature and rainfall. Thus the pattern of Figure 3 includes some projected data well into the twenty first century. For precipitation, there is an anthropogenic signal for enhanced precipitation over northern Europe and a slight drying over southern Europe and the Mediterranean region. Changes in the NAO predictor (Figure 4), the volcanic predictor (Figure 5), the QBO (Figure 6) and ENSO (Figure 7) all lead to a dipolar signal in European temperature with northern Europe and southern Europe showing opposite signed signals. Similar dipolar signals are also found in precipitation, albeit shifted slightly northwards. Very similar signals have been identified in numerous modelling and observational studies of European climate variability and relate strongly to the North Atlantic and Arctic Oscillations. This result further indicates that our predictors are largely indicative of the spatial pattern of influence of the NAO rather than those of other circulation patterns that may be important for north European temperature.

Taking the whole set of European winters from 1954 to 2008, we first assess the degree to which our predictors can explain winter interannual variability. Figure 8 shows that if the values of the predictors for the coming winter were precisely known (this includes knowing the 500 hPa NAO value precisely) then correlations of northern European temperature of around 0.7 would be achievable. Of course, while our predictors are well estimated several months ahead, they cannot be precisely known and, in particular, the NAO shows only limited predictability from the North Atlantic SST index used here (RF2002). A more realistic estimate which

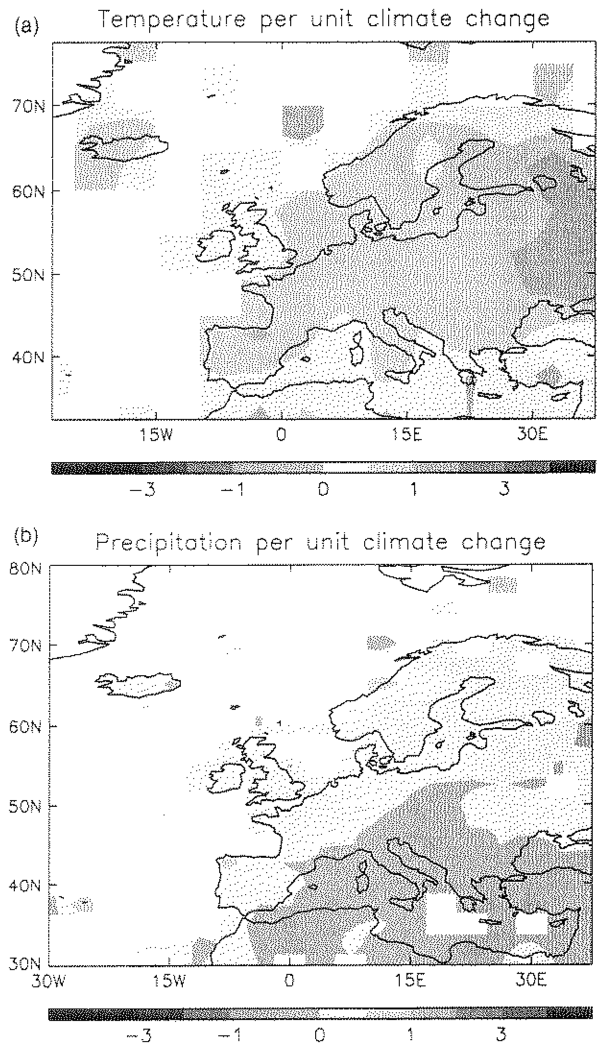

Figure 3. Spatial patterns over Europe for predicted winter temperature and rainfall resulting from systematic climate change. The values correspond to a modelled one standard deviation of the increase in north European mean temperature for 1861–2039 as in shown in Figure 1, together with the corresponding one standard deviation of the rainfall change over 1861–2009. Temperature anomalies in °C and precipitation anomalies in mm/day.

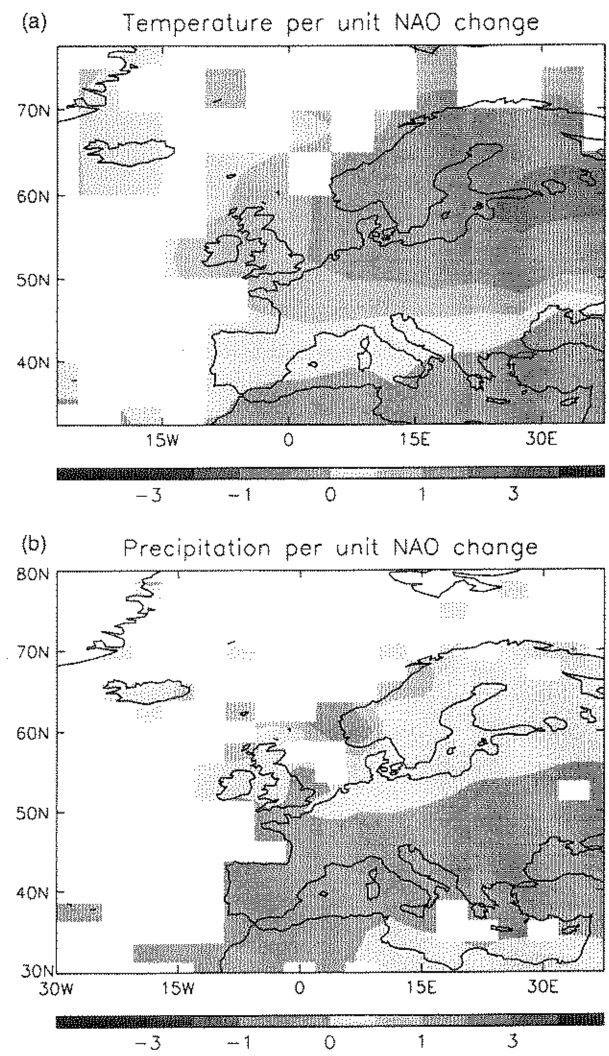

Figure 4. Spatial patterns over Europe of temperature and rainfall for a plus one standard deviation value of the observed NAO index. Units as in Figure 3.

uses the forecast NAO value available in the previous June (Figure 8, lower) indicates that cross-validated correlations of around 0.4 are achievable, assuming other factors such as the winter average QBO and ENSO are known fairly accurately before the winter starts. This is a good assumption as we indicated earlier. As mentioned in Section 2, the QBO can usually be forecast quite accurately from current and recent QBO data about 3 or even more months in advance of the start of winter using the method of Gray *et al.* (1992). The state of ENSO in winter is also usually well predictable 1–3 months in advance of the start of winter using the Glosea 3 coupled dynamical seasonal forecasting model (Graham *et al.*, 2005) or the new Glosea 4 model (Arribas *et al.*, 2009, 2011). The typical value of $R = 0.4$ is similar to the cross-validation result for area mean north European temperature described earlier. Results for precipitation (Figure 9) and for January–February are broadly similar but indicate reduced correlation skill, presumably due to the higher proportion of internal and unpredictable variability in both precipitation and in the shorter 2-month means for temperature discussed earlier.

In conclusion, the cross-validated correlation of about 0.4 represents a significant increase beyond the skill of current dynamical seasonal forecast systems for winter over Europe (Junge and Stephenson, 2003; Rodwell and Doblas-Reyes, 2006). Thus, over the period 1989–2002, a set of hindcasts for the winter NAO from the state-of-the-art Glosea4 model (Arribas *et al.*, 2011) only had a correlation of about 0.2 with the observed winter NAO. This is not statistically significant and represents considerably less variance explained than implied by this paper. Thus, the May NAO SST factor used in this paper alone provides a cross-validated correlation of 0.45 between its hindcast NAO and the observed NAO over winters 1949–1998 (RF 2002).

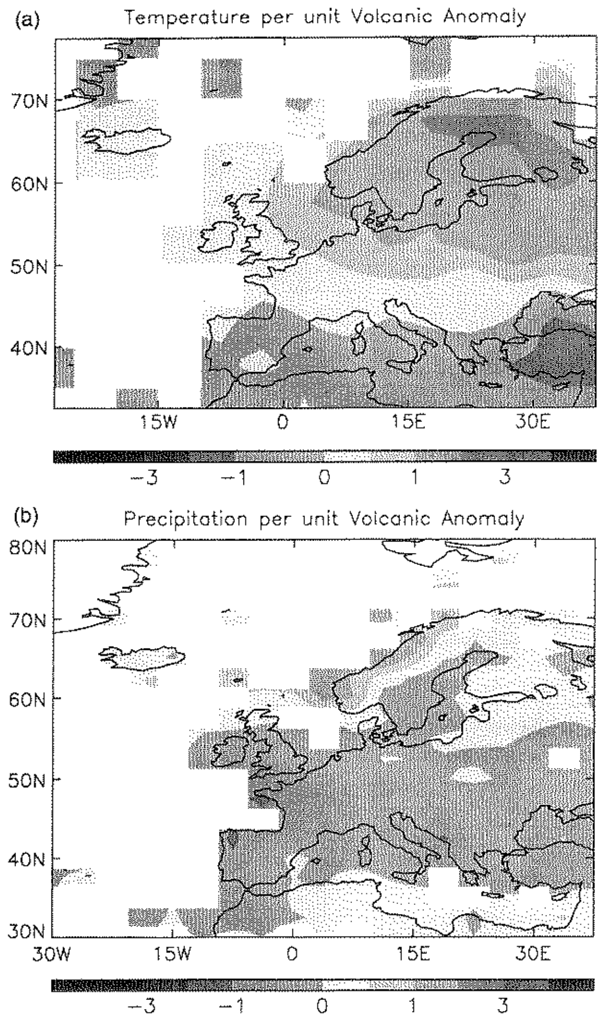

Figure 5. Spatial patterns of predicted (a) temperature, and (b) rainfall anomalies over Europe for the unit value of the volcanic index for temperature and rainfall. Units as in Figure 3. These are fixed anomalies for a volcanic winter.

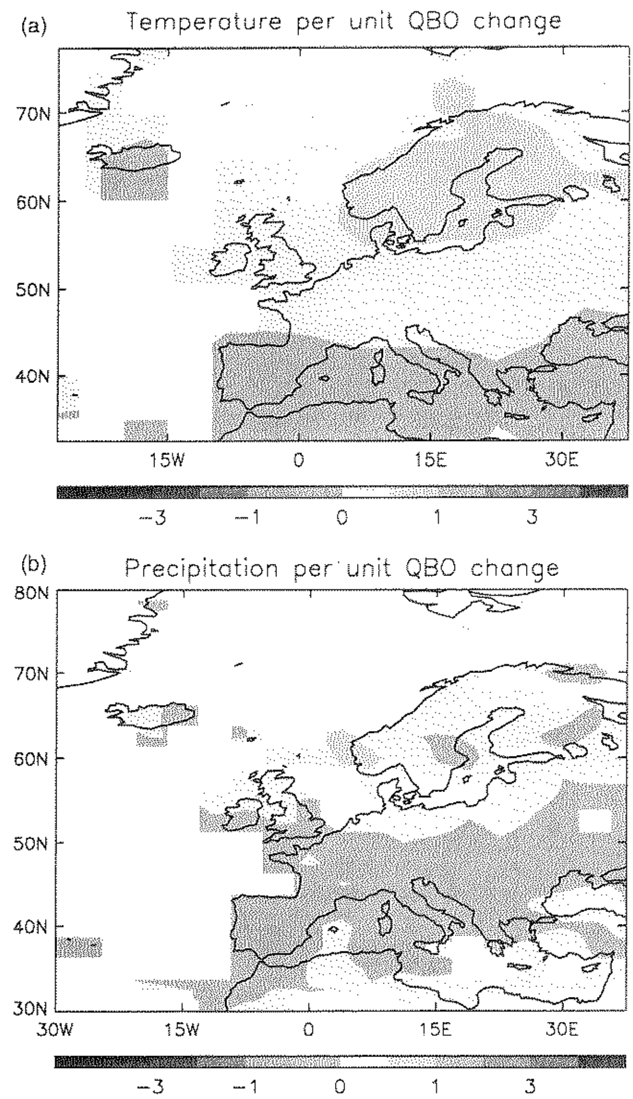

Figure 6. Spatial patterns over Europe of predicted temperature and rainfall for plus one standard deviation value of the of QBO index. Units as in Figure 3.

4.2. Reconstructed and hindcast temperature maps for individual years, December–February

Examples of very cold (1963) and very mild (1990) years are given in Figure 10, along with an example of a cold year which followed a run of mild years (1996). Notwithstanding the potential role of unpredictable atmospheric variability, in each of these successfully forecast cases, the fact that the reconstruction explains the observed anomaly suggests that our estimates of the different influences are quantitatively large enough to explain a considerable proportion of the observed anomalies in these years, though some of the apparent signal might in reality be noise.

Other years in Figure 10 were poorly forecast by our simple method. In the case of 2008, the reconstruction is able to explain the observed warm anomaly very well but the hindcast failed. The NAO SST index prediction in this year was particularly poor and much cooler than observed and this mainly gave rise to the poor hindcast. This failure reflects the likelihood that although there is a clear potential for significantly improved predictions of the NAO and European winter temperature, NAO predictability may remain relatively limited, at least partly due to substantial internal extratropical winter climate variability (Stephenson *et al.*, 2000). Substantial internal variability, and therefore limited predictability, affects not just the NAO, as underlined by the discussion in Kushnir *et al.* (2006), but other north Atlantic atmospheric circulation patterns as well. However, offset against this to some extent are the intriguing insights of Keeley *et al.* (2009) who indicate that a substantial fraction of seasonal NAO variability may be externally forced. However, this is not a definite indication of a similar level of predictability as the forcings may be unpredictable. Finally, the last row in Figure 10 shows a case (1971) where neither the hindcast nor the reconstructions were able to reproduce the observed anomaly. This is one of many broadly similar cases as indicated by Figure 2(d), as it is by no means the worst case. In this particular case, the limited set of predictors in our simple statistical model was unable to explain the observed anomaly.

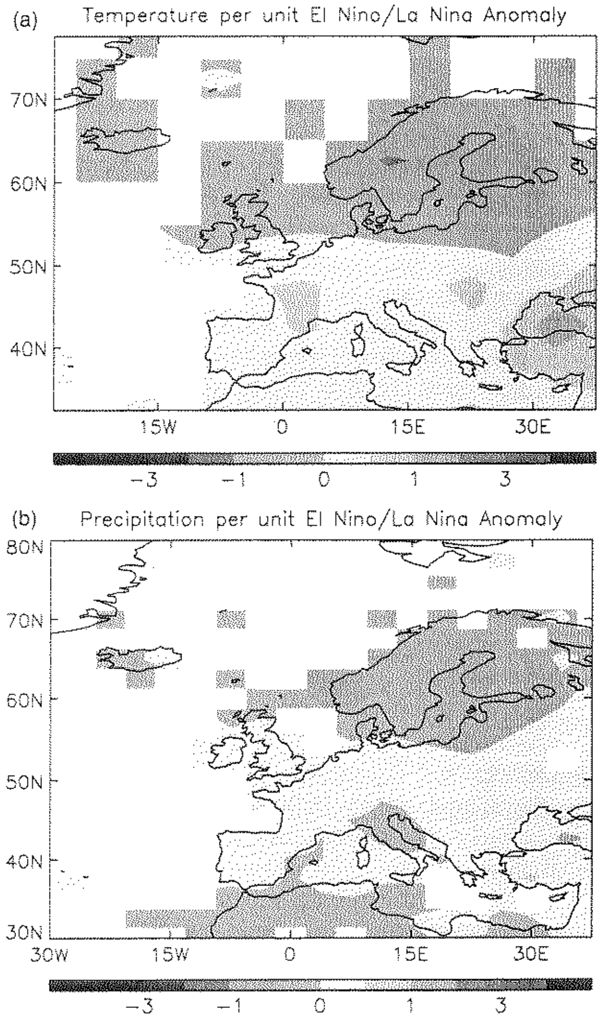

Figure 7. Spatial patterns of predicted (a) temperature, and (b) rainfall over Europe of a positive unit index for an El Nino for temperature and rainfall. This is a fixed anomaly for a (moderate) El Nino with this index value.

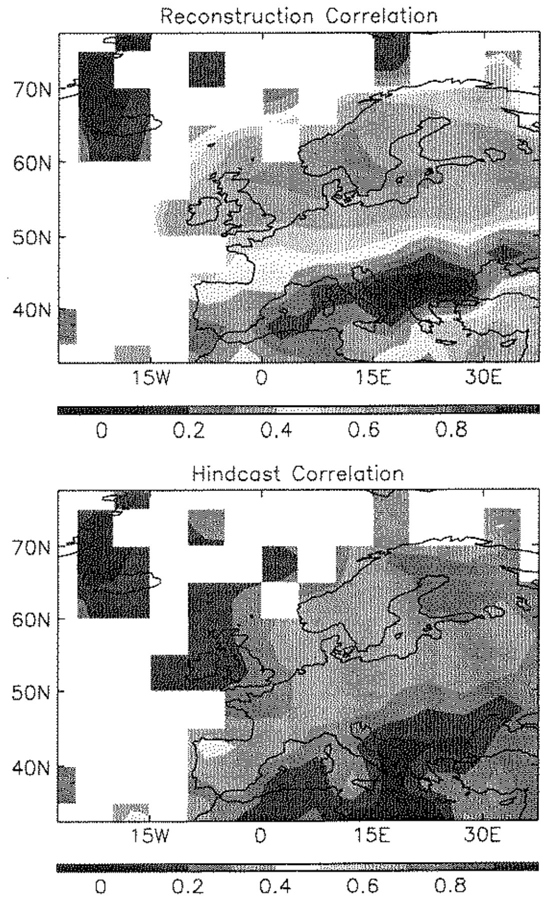

Figure 8. Correlation skill for near-surface temperature reconstructions given perfect predictor data (upper) and hindcasts (lower). Correlations were assessed over 1954–2008. Reconstructed correlation skill uses the observed values of the predictors to assess the maximum amount of variability that can be explained by our predictors. Hindcast skill uses predicted values of the NAO from May SST patterns (RF2002) and other predictors as described in the text.

So we must assume that either another factor or simply internal unpredictable variability in the European circulation was responsible for the relatively mild winter of 1971 over much of Europe.

5. Discussion

The residuals of our one-stage regression models for 1954–2008 in DJF and JF show a strong spectral peak (not shown) with a broad maximum around 7–8 years (slightly sharper for JF), with no other spectral peak of comparable size. This is similar to the period (6–10 years) of a known spectral peak in the winter NAO (Hurrell and Van Loon, 1997; Venegas and Mysak, 1997) and hints at a possible missing influencing factor on European winter climate, though this spectral peak may really be rather ill-defined. Figures 1(b) and (d) also indicate that, in some winters, factors other than those considered here may be particularly important or alternatively that internal climate variability may be dominant in some winters. Further research would be especially worthwhile to investigate these winters more closely, e.g. as identified by Figure 2(d) for northern Europe.

The factors identified here as being important for the winter circulation over Europe include the influence of the stratosphere. The stratosphere was shown to be a fundamental component of longer time scale interactions between the tropospheric NAO and the tropospheric AO by Ambaum and Hoskins (2002). Recently Maycock *et al.* (2011) have shown that climate models without a fully resolved stratosphere are unlikely to pick up the interactions between sudden stratospheric warmings and the troposphere correctly. To further check this, we have calculated the cross-validated correlation of the winter (DJF) hindcasts with observed north European temperature omitting the QBO and also the moderate El Nino index as capturing this component of ENSO influence on European winter climate has been shown to be crucially dependent on stratospheric interactions by Ineson and Scaife (2009). The value of $R = 0.31$, about 10% of the variance, can be contrasted with the 13% of the variance explained by all factors discussed above. So the physical studies and our results

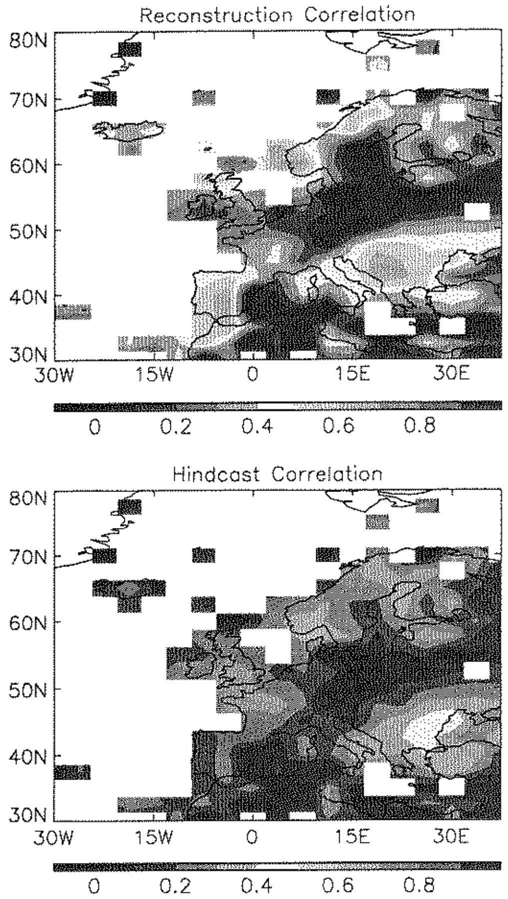

Figure 9. Correlation skill for winter rainfall for reconstructions (upper) and hindcasts (lower). Correlations were assessed over 1954–2008. Reconstructed correlation skill uses the observed values of the predictors to assess the maximum amount of variability that can be explained by our predictors. Hindcast skill uses predicted values of the NAO from May SST patterns and other predictors as described in the text.

imply that a model with a realistic representation of stratospheric variability is required to fully capture the level of European winter predictability described by this paper.

Although data are too few to investigate this purely statistically, there may be a need to modify the Atlantic SST predictor when El Nino and possibly La Nina is active. It is well known that El Nino tends to favour a strong winter storm track and enhanced rainfall over the southern USA extending into the western subtropical Atlantic (Ropelewski and Halpert, 1986, 1996; Smith and Ropelewski, 1997) commonly seen to cause some cooling of SST in that region. At the same time, anomalous atmospheric descent is favoured over the western tropical Atlantic; that area of the Atlantic warms as well as the tropical Atlantic further east due to reductions in the trade winds associated with El Nino (Curtis and Hastenrath, 1995), though maximum tropical Atlantic warming often occurs after boreal winter. These effects might sometimes disrupt the re-emergence of the May SST field in El Nino winters. So the RF2002 method may be less skilful when El Nino is moderate or strong in winter, and possibly when La Nina is strong when broadly opposite effects on SST anomalies might be expected. Progress requires appropriate observational and dynamical modelling studies to clarify how the previous May north Atlantic subsurface temperatures are modified. Such studies will be helped by the recent development of several improved subsurface ocean datasets (Domingues *et al.*, 2008; Ishii *et al.*, 2009; Levitus *et al.*, 2009). So far, observed data have not been sufficient to carry out such studies and they would need to take account of the variable effects of El Nino on winter North Atlantic climate found by Toniazzo and Scaife (2006) and Ineson and Scaife (2009). In the meantime, it will be obvious before winter starts whether the re-emergent SST pattern has been disrupted. Such knowledge can be used inform a revised forecast.

A possible predictor which we do not include is pre-winter season Northern Hemisphere snow cover anomalies. Some skill was reported by Cohen and Fletcher (2007) for empirical forecasts of winter land surface temperatures over the Northern Hemisphere, including Europe on its own, using a combination of October snow cover and sea level pressure anomalies. This effect has since been looked at in a climate model (e.g. Fletcher *et al.*, 2009) using a Siberian snow cover anomaly in October and November. This was shown to force the stratosphere which in turn forced the AO in a manner quite like the process described by Ambaum and Hoskins (2002). The main forcing effect appears to act through the surface albedo change associated with snow cover anomalies. Thus, October snow anomalies may have at least as big an effect as the November ones (Fletcher *et al.*, 2009). So there may well be some predictability of the Northern Hemisphere winter atmosphere circulation from preceding autumn snow cover. This may also be a further example of the need for accurate stratospheric–tropospheric interactions.

Curiously, however, Saunders *et al.* (2003) had already indicated that Northern Hemisphere summer (June–July) snow cover anomalies might have a larger predictive power for the winter NAO than snow cover in preceding autumn. According to their observational analysis, predictive power peaks in the preceding summer (for the NAO) and reduces in the following autumn, a counter-intuitive result. Albedo effects would of course maximize in summer, but it is not easy to see how such effects could persist through to winter. Further research is clearly needed to explain and quantify the predictive influence of snow cover, especially given Keeley *et al.*'s (2009) estimate that more interannual forcing of the NAO may exist than explained in this paper.

Finally, although our results are presented in terms of the skill of 'best estimate' forecasts, the methodology can readily be extended to investigate the skill of probability forecasts. This is the form in which dynamical seasonal forecasts are usually made, based on a fairly large ensemble of forecasts starting from a range of different initial conditions.

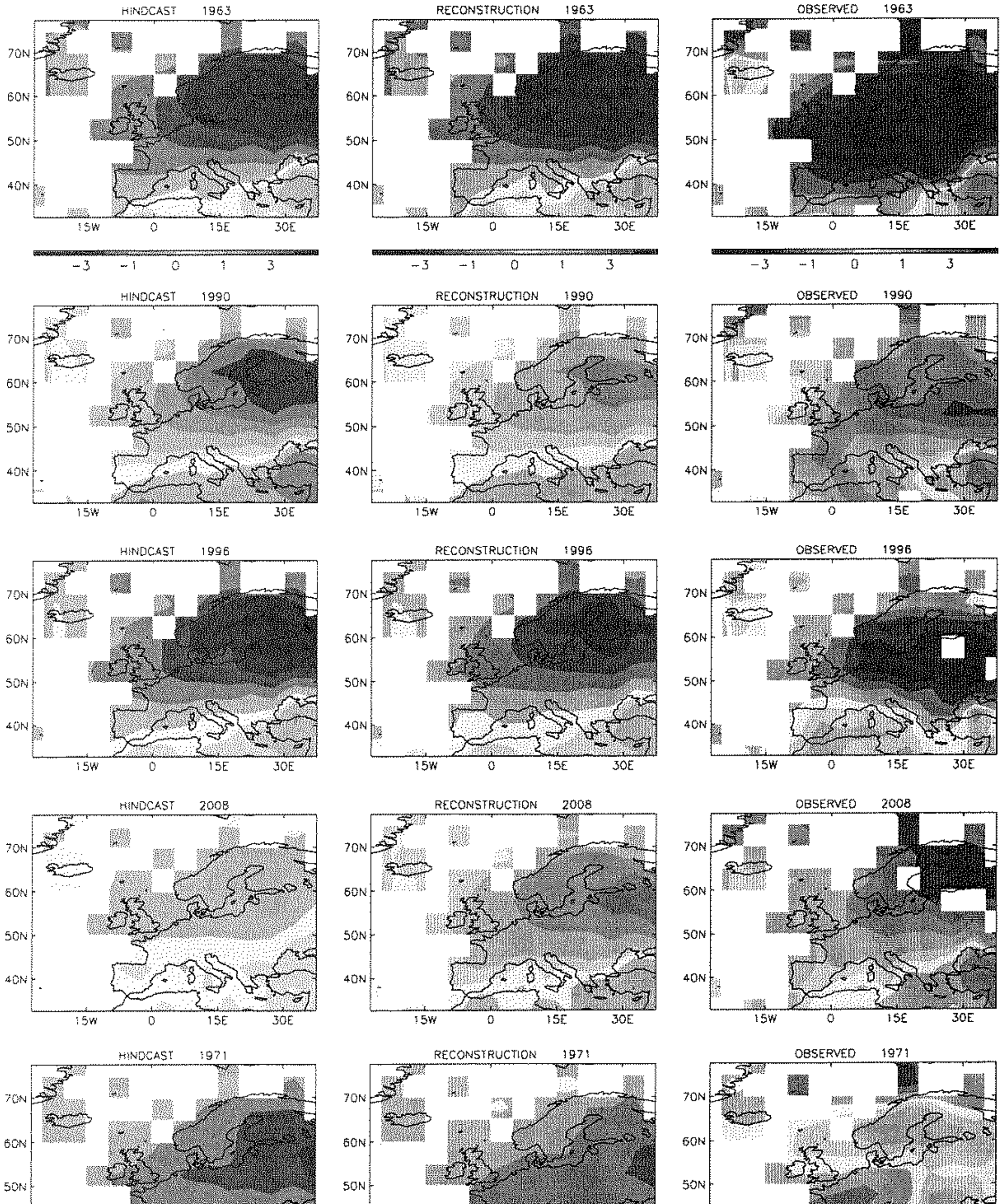

Figure 10. Examples of hindcast, reconstructed and observed winters. Examples of very cold (1963), very mild (1990) and a large observed departure from the previous year (1996) are shown in the first three rows. Both the hindcast and the reconstruction reproduced the observed anomaly in these years quite well. The penultimate row shows 2008 when the reconstruction explained the observed anomaly quite well but the forecast was poor. The last row shows 1971 when both the reconstruction and the hindcast failed to reproduce the observed anomaly pattern, though small parts of northern Europe had the correct sign of the anomaly. This is a fairly typical error, averaged over northern Europe, as indicated in Figure 2(d).

6. Summary and conclusion

Practically useful potential predictability of winter climate over Europe a season ahead has been identified using a small set of predictors based mainly on dynamical studies together with supporting observational data. Predictability identified here exceeds the levels in current dynamical forecast systems, e.g. as indicated above by Arribas *et al.* (2011). Our study indicates some sources of predictability that are usually not well represented in current dynamical forecast systems. These include much of the skill inherent in the May Atlantic SST predictor, despite being available in the initial conditions, and much of the influence of the QBO. In addition, current dynamical models are poor at picking up the signal from moderate El Ninos, though our statistical model may also underestimate this source of skill, especially in January and February.

In future, it may be advantageous to use the new ENSEMBLES high resolution climate data set (Haylock *et al.*, 2008). Averaged over the north European region, the HadCRUT3v and ENSEMBLES temperature data are highly correlated with $R = 0.95$ for DJF, and $R = 0.96$ for JF. So the main conclusions of our paper would be little affected by using the better ENSEMBLES dataset and the data are also limited to a start date of 1950. They could, though, add significant value to the patterns in predictability maps.

We do not advocate our simple methods as an alternative to dynamical seasonal forecasts because dynamical methods are potentially capable of describing nonlinear interactions between the predictable components of the climate system. Our simple linear regression methods cannot do this. However, they are likely to be useful until dynamical forecast systems improve considerably. Improvements to dynamical methods are becoming more urgent as winter climate over Europe is now clearly changing under the influence of global climate warming: forecasting under conditions of non-stationarity may lead to different responses which can be modelled in a dynamical framework but cannot be included in simple linear statistical methods. Thus, although our potential predictability model does include the non-stationarity of the warming signal, it cannot include future changes in the responses of the climate to factors already identified, or possible new forcing factors like a permanent large reduction in winter Arctic sea ice extent.

In conclusion, this study points toward the phenomena to which dynamical models need to respond skilfully, and where appropriate process-based model developments are needed. This is likely to require the extension of current dynamical seasonal forecasting models up to the mesosphere, at the same time ensuring an accurate climatology in the stratosphere and the troposphere. It is also likely to require substantially improved horizontal resolution as current low-resolution models are not likely to respond correctly to North Atlantic SST patterns and strong SST gradients (Minobe *et al.*, 2008). Thus, even the most recent Met Office model, Glosea4, does not appear to pick up much of the winter signal of the NAO SST tripole (A. Arribas, personal communication), and we have identified this as one of our most important winter predictors. Finally, there is the prospect of better worldwide initialisation of dynamical seasonal forecasts now that the ARGO array (Wilson, 2000; Gould and Turton, 2006), which measures the temperature and salinity of the top 2000 m of the ocean is now essentially complete, worldwide.

Acknowledgements

This research was, firstly, supported by the Joint DECC and Defra Integrated Climate Programme – DECC/Defra (GA01101). Janet Lindesay contributed to this work mainly as a Visiting Scientist at the Met Office Hadley Centre for Climate Change. David B. Stephenson thanks the Met Office for joint funding of his Chair at Exeter University which made this collaboration possible. We wish to thank Pascal Mailier and two anonymous reviewers for helpful comments. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://eca.knmi.nl).

References

Afifi AA, Azen SP. 1979. *Statistical analysis – a computer oriented approach*. Academic Press: London; pp. 442.

Ambaum MHP, Hoskins BJ. 2002. The NAO tropospheric-stratospheric connection. *Journal of Climate* **15**: 1969–1978.

Arribas A, Glover M, Maidens A, Peterson K, Gordon M, MacLachlan C. 2009. Towards a new generation of seasonal forecasting systems. *Revista Fisica Tierra* **21**: 1–8.

Arribas A, Glover M, Maidens A, Peterson K, Gordon M, MacLachlan C, Graham R, Fereday D, Camp J, Scaife AA, Xavier P, McLean P, Colman A, Cusack S. 2011. The GloSea4 ensemble prediction system for seasonal forecasting. *Monthly Weather Review* (in press).

Baldwin MP, Dunkerton TJ. 1999. Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *Journal of Geophysical Research [Atmospheres]* **104**: 30937–30946.

Baldwin MP, Gray LG, Dunkerton TJ, Hamilton K, Haynes PH, Randel WJ, Holton JR, Alexander MJ, Hirota I, Horinouchi T, Jones DBA, Kinnersley JS, Marquardt C, Sao K, Takahasi M. 2001. The quasi-biennial oscillation. *Reviews in Geophysics* **39**: 179–229.

Boer GJ. 2009. Climate trends in a seasonal forecasting system. *Atmosphere-Ocean* **47**: 123–138.

Boer GJ, Hamilton K. 2008. QBO influence on extratropical predictive skill. *Climate Dynamics* **31**: DOI:10.1007/s00382-008-0379-5.

Böhm R, Jones PD, Hiebl J, Frank D, Brunetti M, Maugeri M. 2010. The early instrumental warm-bias: a solution for long central European temperature series 1760–2007. *Climatic Change* **101**: 41–67.

Brohan P, Kennedy JJ, Harris I, Tett SFB, Jones PD. 2006. Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *Journal of Geophysical Research [Atmospheres]* **111**: D12106, DOI:10.1029/2005JD006548.

Brönnimann S. 2007. ENSO influence on Europe during the last centuries *Clim. Dynam.* **28**: DOI:10.1007/s00382-006-0175-z.

Cagnazzo C, Manzini E. 2009. Impact of the stratosphere on the winter tropospheric teleconnections between ENSO and the North Atlantic and European region. *Journal of Climate* **22**: 1223–1238.

Cohen J, Fletcher C. 2007. Improved skill of Northern Hemisphere winter surface temperature prediction based on land-atmosphere fall anomalies. *Journal of Climate* **20**: 4118–4132.

Curtis S, Hastenrath S. 1995. Forcing of anomalous sea surface temperature evolution in the tropical Atlantic during Pacific warm events. *Journal of Geophysical Research [Oceans]* **100**: 15835–15847.

Davies JR, Rowell DP, Folland CK. 1997. North Atlantic and European seasonal predictability using an ensemble of multidecadal AGCM simulations. *International Journal of Climatology* **12**: 1263–1284.

Domingues CM, Church JA, White NJ, Gleckler PJ, Wijffels SE, Barker PM, Dunn JR. 2008. Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature* **453**: 1090–1093, DOI:10.1038/nature07080.

Draper NR, Smith H. 1998. *Applied Regression Analysis*. Wiley-Interscience; pp. 706.

Ebdon RA. 1975. The quasi-biennial oscillation and its association with tropospheric circulation patterns. *Meteorological Magazine* **104**: 282–297.

Ebdon RA, Veryard RG. 1961. Fluctuations in equatorial stratospheric winds. *Nature* **189**: 791–793.

Fereday DR, Knight JR, Scaife AA, Folland CK, Philipp A. 2008. Cluster analysis of North Atlantic/European circulation types and links with tropical Pacific sea surface temperatures. *Journal of Climate* **21**: 3687–3703, DOI:10.1175/2007JCLI1875.1.

Fischer EM, Luterbacher J, Zorita E, Tett SFB, Casty C, Wanner H. 2007. European climate response to tropical volcanic eruptions over the last half millennium. *Geophysical Research Letters* **34**: Lo5707, DOI:10.1029/2006GL027992.

Fletcher CG, Saunders MA. 2006. Winter North Atlantic Oscillation hindcast skill: 1900–2001. *Journal of Climate* **19**: 5762–5776.

Fletcher CG, Hardiman SC, Kushner PJ, Cohen J. 2009. The dynamical response to snow cover perturbations in a large ensemble of atmospheric GCM integrations. *Journal of Climate* **22**: 1208–1222.

Folland CK. 2005. Assessing bias corrections in historical sea surface temperature using a climate model. *International Journal of Climatology*. **25**: 895–911.

Folland CK, Karl TR. 2001. Observed Climate Variability and Change. In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds). Cambridge University Press: 99–181.

Folland CK, Parker DE, Scaife AA, Kennedy J, Colman A, Brookshaw A, Cusack S, Huddleston MR. 2006. The 2005–6 winter in Europe and the United Kingdom: Part 2 prediction techniques and their assessment against observations. *Weather* **61**: 337–346.

Fraedrich K. 1992. Climate anomalies in Europe associated with ENSO extremes. *International Journal of Climatology* **12**: 25–31.

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA. 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics* **16**: 147–168.

Gould WJ, Turton J. 2006. Argo – sounding the oceans. *Weather* **61**: 17–21.

Graf H-F, Perlwitz J, Kirchner I. 1994. Northern Hemisphere tropospheric mid-latitude circulation after violent volcanic eruptions. *Contributions to Atmospheric Physics* **67**: 3–13.

Graham RJ, Gordon M, Mclean PJ, Ineson S, Huddleston R, Davey MK, Brookshaw A, Barnes TH. 2005. A performance comparison of coupled and uncoupled versions of the Met Office seasonal prediction general circulation model. *Tellus* **57A**: 320–339.

Gray WM, Landsea CW, Mielke PW, Berry KJ. 1992. Predicting Atlantic seasonal hurricane activity 6–11 months in advance. *Weather and Forecasting* **7**: 440–455.

Hamilton K. 1998. Effects of an imposed quasi-biennial oscillation in a comprehensive troposphere-stratosphere-mesosphere general circulation model. *Journal of Atmospheric Science* **55**: 2393–2418.

Haylock MR, Hofstra N,, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded dataset of surface temperature and precipitation. *Journal of Geophysical Research* [Atmospheres] **113**: D20119, DOI:10.1029/2008JD10201.

Hegerl GC, Zwiers FW. 2007. Understanding and Attributing Climate Change. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor MMB, Miller HL, Chen Z. (eds). Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA; 663–745.

Holton JR, Tan H-C. 1980. The influence of the equatorial Quasi-Biennial Oscillation on the global circulation at 50 mb. *Journal of Atmospheric Science* **37**: 2200–2208.

Hurrell, JW. 1995. Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science* **269**: 676–679.

Hurrell JW, van Loon H. 1997. Decadal variations in climate associated with the North Atlantic Oscillation. *Climate Change* **36**: 301–326.

Ineson S, Scaife AA. 2009. The role of the stratosphere in the European climate response to El Nino. *Nature Geoscience* **2**: 32–36, DOI: 10.1038/ngeo381.

IPCC 2007. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor MMB, Miller HL, Chen Z. (eds). Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA; 235–336.

Ishii M, Kimoto M. 2009. Re-evaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections. *Journal of Oceanography* **65**: 287–299.

Jin EK, Kinter III JL, Wang B, Park C-K, Kang LS, Kirtman BP, Kug J-S, Kumar A, Luo J-J, Schemm J, Shukla J, Yamagata T. 2008. Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Climate Dynamics* **31**: 647–664.

Jones PD, New M, Parker DE, Martin S, Rigor IG. 1999. Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics* **37**: 173–199.

Junge MM, Stephenson DB. 2003. Mediated and direct effects of the North Atlantic ocean on winter temperatures in northwest Europe. *International Journal of Climate* **23**: 245–261.

Keeley SPE, Sutton RT, Shaffrey LC. 2009. Does the North Atlantic Oscillation show unusual persistence on intraseasonal timescales? *Geophysical Research Letters* **36**: L22706, DOI:10.1029/2009GL040367.

Kushnir Y, Robinson WA, Bladé I, Hall NMJ, Peng S, Sutton R. 2002. Atmospheric GCM response to extratropical SST anomalies: synthesis and evaluation. *Journal of Climate* **15**: 2233–2256.

Kushnir Y, Robinson WA, Chang P, Robertson AW. 2006. The physical basis for predicting Atlantic sector seasonal-to-interannual climate variability. *Journal of Climate* **19**: 5949–5970.

Levitus SJL, Antonov TP, Boyer RA, Locarnini HE, Garcia, Mishonov AV. 2009. Global ocean heat content 1955–2007 in light of recently revealed instrumentation problems. *Geophysical Research Letters* **36**: L07608, DOI:10.1029/2008GL037155.

Liniger MA, Mathis H, Appenzeller C, Doblas-Reyes FJ. 2007. Realistic greenhouse gas forcing and seasonal forecasts. *Geophysical Research Letters* **34**: DOI 10.1029/2006GLO28335.

Marshall AG, Scaife AA. 2009. Impact of the QBO on surface winter climate. *Journal of Geophysical Research [Atmospheres]* **114**: D18110, DOI:10.1029/2009JD011737.

Marshall AG, Scaife AA, Ineson S. 2009. Enhanced seasonal prediction of European winter warming following volcanic eruptions. *Journal of Climate* **22**: 6168–6180.

Maycock AC, Keeley SPE, Charlton-Perez AJ, Doblas-Reyes FJ. 2011. Stratospheric circulation in seasonal forecasting models: implications for seasonal prediction. *Climate Dynamics* **36**: 309–321, DOI:10.1007/s00382-009-0665-x.

Minobe S, Kuwano-Yoshida A, Komori N, Xie S-P, Small RJ. 2008. Influence of the Gulf Stream on the troposphere. *Nature* **452**: 206–209. DOI:10.1038/nature06690.

Moron V, Gouirand I. 2004. Seasonal modulation of the El Nino-Southern Oscillation relationship with sea level pressure anomalies over the North Atlantic in October-March 1873–1996. *International Journal of Climatology* **23**: 143–155.

Naujokat B. 1986. An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics. *Journal of Atmospheric Science* **43**: 1873–1877.

New M, Hulme M, Jones P. 2000. Representing twentieth-century space–time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *Journal of Climate* **13**: 2217–2238.

Parker DE, Folland CK, Jackson M. 1995. Marine surface temperature: observed variations and data requirements. *Climate Change* **31**: 559–600.

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A. 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research [Atmospheres]* **108**(D14): 4407, DOI:10.1209/2002JD002670, 2003.

Robock A. 2000. Volcanic eruptions and climate. *Reviews of Geophysics* **38**: 191–219.

Robock A, Mao, J. 1995. The volcanic signal in surface temperature observations. *Journal of Climate* **8**: 1086–1103.

Rodwell M, Doblas-Reyes FJ. 2006. Predictability and prediction of European monthly to seasonal climate anomalies. *Journal of Climate* **19**: 6025–6046.

Rodwell MJ, Folland CK. 2002. Atlantic air-sea interaction and seasonal predictability. *Quarterly Journal of the Royal Meteorological Society* **128**: 1413–1443.

Rodwell MJ, Rowell DP, Folland CK. 1999. Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature* **398**: 320–323.

Ropelewski CF, Halpert, MS. 1986. North American precipitation and temperature patterns associated with the El Nino/Southern Oscillation (ENSO). *Monthly Weather Review* **114**: 2352–2362.

Ropelewski CF, Halpert, MS. 1996. Quantify Southern Oscillation Precipitation relationships. *Journal of Climate* **9**: 1043–1059.

Saunders MA, Qian B, Lloyd-Hughes B. 2003. Summer snow extent heralding of the winter North Atlantic Oscillation. *Geophysical Research Letters* **30**: DOI: 10.1029/2002GL016832.

Scaife, AA, Folland CK, Alexander LV, Moberg A, Brown S, Knight JR. 2008. European climate extremes and the North Atlantic Oscillation. *Journal of Climate* **21**: 72–83.

Scaife AA, Knight J, Vallis G, Folland CK. 2005. A stratospheric influence on the winter NAO and North Atlantic surface climate. *Geophysical Research Letters* **32**: L18715, DOI:10.1029/2005GL023226.

Smith TH, Ropelewski CF. 1997. Quantifying Southern Oscillation-precipitation relationships from an atmospheric GCM. *Journal of Climate* **10**: 2277–2284.

Stenchikov G, Hamilton K, Stouffer S, Robock R, Ramaswamy V, Santer B, Graf H-F. 2006. Arctic Oscillation response to volcanic eruptions in the IPCC AR4 climate models. *Journal of Geophysical Research (Atmospheres)* **111**: DOI:10.1029/2005JD006286.

Stenchikov GL, Kirchner I, Robock A, Graf H-F, Antuña JC, Grainger RG, Lambert A, Thomason L. 1998. Radiative forcing from the 1991 Mount Pinatubo volcanic eruption. *Journal of Geophysical Research (Atmospheres)* **103**: 13837–13857.

Stephenson DB, Pavan V, Bojariu R. 2000. Is the North Atlantic Oscillation a random walk? *International Journal of Climatology* **20**: 1–18.

Stott PA, Tett SFB, Jones GS, Allen MR, Mitchell JFB, Jenkins GJ. 2000. External control of 20th century temperature by natural and anthropogenic forcings. *Science* **290**: 2133–3137.

Thompson DWJ, Baldwin MP, Wallace JM. 2002. Stratospheric connection to Northern Hemisphere wintertime weather: implications for prediction. *Journal of Climate* **15**: 1421–1428.

Thompson DWJ, Wallace JM. 2001. Regional climate impacts of the northern hemisphere annular mode. *Science* **293**: 85–89.

Toniazzo T, Scaife AA. 2006. The influence of ENSO on winter North Atlantic climate. *Geophysical Research Letters* **33**: L24704, DOI:10.1029/2006GL027881.

Trenberth KE, Jones PD. 2007. *Chapter 3 of Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor MMB, Miller HL, Chen Z. (eds). Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA; 235–336.

van Oldenborgh GJ, Balmaseda MA, Ferranti L, Stockdale TN, Anderson DLT. 2005. Evaluation of atmospheric fields from the ECMWF seasonal forecasts over a 15 year period. *Journal of Climate* **18**: 5188–5198.

van Oldenborgh GJ, Drijfhout SS, van Ulden A, Haarsma A, Sterl A, Severijns C, Hazeleger W, Dijkstra H. 2009. Western Europe is warming much faster than expected. *Climate of the Past* **5**: 1–12.

Venegas, SA, Mysak, LA. 2000. Is there a dominant timescale of natural climate variability in the Arctic? *Journal of Climate* **13**: 3412–3434.

von Storch H, Zwiers FW. 1999. *Statistical analysis in climate research*. Cambridge University Press; pp. 484.

Wilson S. 2000. Launching the Argo armada. *Oceanus* **42**: 17–19.

