J3.12 SEASONAL FORECASTING OF STRONG WINDS OVER EUROPE

J. P. Palutikof, T. Holt, and T.J. Osborn
Climatic Research Unit, University of East Anglia, Norwich, UK

1. INTRODUCTION

Seasonal forecasting for tropical regions, based on statistical and dynamical modelling approaches, is well-established (Goddard et al., 2001). Prediction models have been developed for phenomena as diverse as Atlantic hurricanes (Lehmiller et al., 1997), the Indian Monsoon (Krishna Kumar et al., 1995) and Sahelian rainfall (Ward, 1998). By comparison, the potential for seasonal forecasting in temperate latitudes is recognized to be lower (Lloyd-Hughes and Saunders, 2001). Over Europe, recent work has emphasised the prediction of rainfall (Rimbu et al., 2001; Lloyd-Hughes and Saunders, 2001) and the role of the North Atlantic Oscillation (Rodwell et al., 1999).

As financial losses due to weather extremes escalate, there is growing interest from end-users, for example from forestry and insurance, in the development of seasonal forecasting models for weather extremes in temperate latitudes. Here, we present results from a pilot project to explore the potential for seasonal forecasting of wind storm over Europe.

2. THE WIND SPEED PREDICTAND

The study area covers northwestern Europe between 45°N to 65°N and 15°W to 25°E. Over this region, the wind climatology varies substantially. Mean speeds increase from south-east to north-west. Superimposed on this pattern, speeds are highest in coastal areas and at high-altitude locations. Against this background, we sought an approach which would generate useful predictions at the regional scale, whilst permitting spatial intercomparison of the results. To achieve this goal, the predictand variable was based not on absolute wind speeds, but on the number of exceedances of percentile thresholds. By using a relative rather than an absolute measure, useful and comparable results are generated.

The predictand variable time series were taken from six-hourly gridded (at a resolution of 2.5° latitude by 2.5° longitude) scalar wind speeds from the National Centers for Environmental Prediction (NCEP) reanalyses. These data are now available from 1948 to present. This provided a readily-available homogeneous data source entirely appropriate for the exploratory purposes of a pilot project. As a first step, the daily maximum wind speed was selected from the four available values. The resulting time series was used to compute from the 1961-90 normal period, for each grid point, the 90th, 95th and 99th percentile values. Figure 1 shows the value of the 90th percentile wind speed for the 153 grid boxes in the domain. A time series of the annual number of exceedances of this 90th percentile threshold is shown in Figure 2, averaged across all grid boxes. These NCEP extremes suggest an overall increase in wind speed extremes over Europe in the last 40 years, accompanied by an increase in year-to-year variability. The predictand variable was found by calculating the number of exceedances of each threshold during each long winter season, October-March. Damaging high winds may occur in any month in this period.

FIGURE 1 90th percentile wind speed (ms^{-1}).

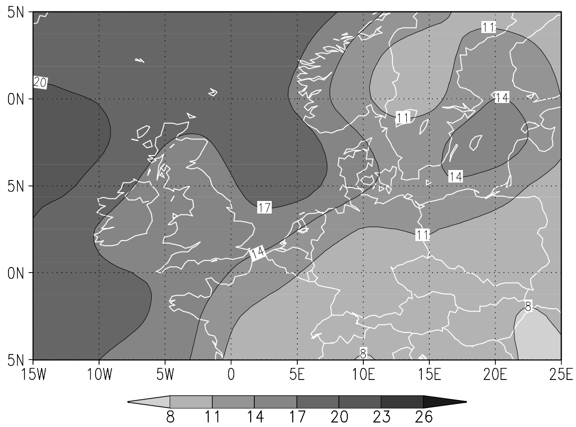
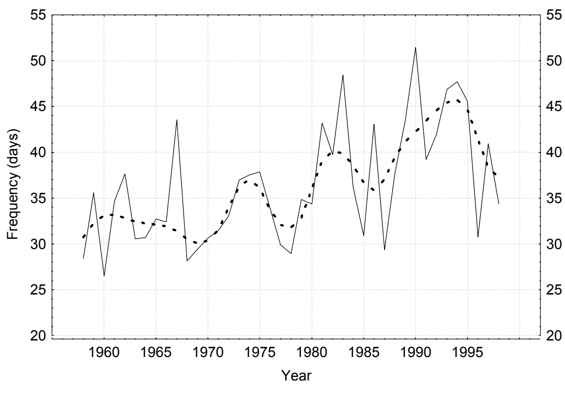


FIGURE 2 Time series of the number of exceedances of the 90th percentile, averaged across all grid boxes in the domain.



Corresponding author address: Jean Palutikof, Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK; e-mail: j.palutikof@uea.ac.uk.

3. THE PREDICTORS

The candidate predictors are regional indices of sea surface temperatures over the Atlantic and eastern Pacific, together with a range of indices of the state of the atmospheric circulation. Sea surface temperatures, largely because of their long time scales of variability compared to the atmosphere, are widely-recognized as offering potential predictability with lead times of a few months to a year (Goddard et al., 2001). In addition, a number of authors report improvements in seasonal forecasting skill when atmospheric predictors are included (e.g., Francis and Renwick, 1998).

3.1 Sea surface temperatures

The data set used was the global GISST 2.3b monthly estimates of sea surface temperature (SST). This is provided on a 1^o by 1^o grid from 1871 to present. The predictor variables were identified using Principal Components Analysis (PCA), which requires that the number of variables (grid squares in this case) in the data set must be less than the number of cases (months). To achieve this whilst considering a domain large enough to properly identify all patterns influencing European high wind occurrence, the SST data were interpolated onto a 5^o by 5^o grid over the Atlantic and Eastern Pacific from 100^oW to 20^oE and 40^oS to 70^oN, giving 621 grid boxes.

The SSTs were normalized (mean of 0, standard deviation of 1) and deseasonalized before performing the PCA. Rotation was not found to be necessary. In order to identify significant modes, a number of tests have been proposed. For example, the Kaiser criterion argues that all modes with eigenvalues greater than 1 should be considered. Here, this would require consideration of 57 modes, too many to reasonably incorporate as predictor variables. There are 21 modes which individually explain at least 1% of the variance. Together, these 21 explain nearly 70% of the variance, as shown in Table 1. We take this as a reasonable cutoff in terms of meaningful predictors.

TABLE 1 Eigenvalues from the SST PCA

Mode	Eigenvalue	% variance	Cum. % variance
1	49.64	12.14	12.14
2	34.46	8.43	20.56
3	26.70	6.53	27.09
4	20.98	5.13	32.22
5	18.92	4.63	36.85
…	…	…	…
21	4.32	1.06	69.17

3.2 *Atmospheric indices*

A number of indices based on atmospheric pressure have been identified as candidate predictors in the seasonal forecasting models. Their association with European weather ranges from well-established (in the case of the North Atlantic Oscillation) to being tentative and the subject of research (Southern Oscillation Index). They were obtained from the web pages of institutions named below and are:

- Arctic Oscillation (AO) from JISAO/University of Washington (JISAO/UW)[1]
- East Atlantic (EA) from Climate Prediction Center (CPC)[2]
- East Atlantic Jet (EAJ) from CPC[2]
- East Atlantic/Western Russia (EAWR) from CPC[2]
- North Atlantic Oscillation (NAO) from Climatic Research Unit (CRU)[3]
- Quasi-Biennial Oscillation (QBO) from JISAO/UW[1]
- Southern Oscillation Index (SOI) from CPC[2]

They encompass a wide range of potential influences on Europe, as shown in Table 2.

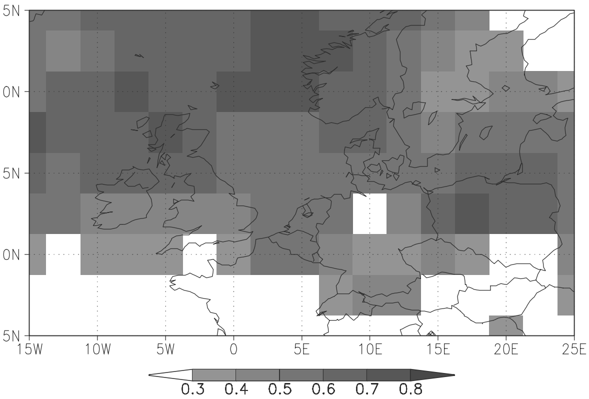
As a first step, the time series of these indices were correlated with the wind speed exceedance time series for each grid box. With one exception, this was carried out for simultaneous winter data, and also with the index time series leading the exceedance time series by one winter. The exception is the EAJ, which is a summer-active series, and for which only relationships with the summer preceding the winter exceedance values was explored. The results for the NAO are shown in Figure 3.

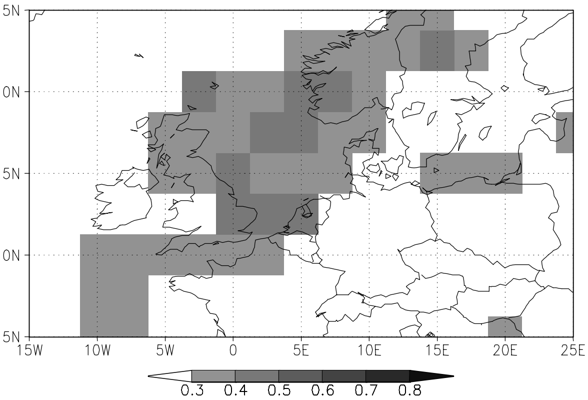
TABLE 2 Atmospheric indices tested as predictors

Index	At. circulation	Active period	Link to Europe
AO	Hemispheric westerly windiness	Mostly winter	Winter winds
EA	Westerly windiness south of UK	Mostly winter	Winter SW winds
EAJ	Westerly windiness south of UK	Summer	Summer SW winds
EA/WR	Blocking over western Europe	Winter, early spring	Winter winds over mainland
NAO	North Atlantic westerly windiness	Mostly winter	Winter winds
QBO	Stratospheric winds	Approx. 2-year cycle	Uncertain
SOI	Pressure difference Tahiti/Darwin	3 – 7 year timescale	Uncertain

[1] http://tao.atmos.washington.edu/main.html
[2] http://www.cpc.ncep.noaa.gov/index.html
[3] http://www.cru.uea.ac.uk/

FIGURE 3 Correlations between October-March NAO and wind speed exceedances of the 90th percentile. Upper map: simultaneous data. Lower map: NAO leads wind speed by one year.





Maps of lagged relationships for each of the atmospheric indices in Table 2 were inspected and, on the basis of this inspection, we were able to reduce the number of potential predictors in Table 2 to just four:

- Arctic Oscillation
- North Atlantic Oscillation
- East Atlantic pattern
- East Atlantic/Western Russia pattern

The issue of multicollinearity in these indices is touched on in Section 4. Together with the 21 components from the PCA of SST, this gave a total of 25 potential predictors.

4. THE MODELS

Three regression methods were tested as the basis for the seasonal forecasting models. These were:

a. Multiple linear regression (MLR). This 'standard' regression method was rejected because of inherent assumptions and constraints A particular problem here would be that the predictor variables should not themselves be correlated (multicollinearity).

b. Partial least squares regression (PLSR). This technique (Merola and Abraham, 2001) considers the variance structure of predictand(s) and predictors jointly. It has a number of advantages for the purposes of this study, including that it:
   - can handle multiple predictands,
   - can accept more variables than cases and
   - uses eigenvector methods, such that multicollinearity is not a problem.

c. Principal component regression (PCR). This technique is very similar to PLSR, but the variance structure of the predictand and predictor variables are considered separately. It requires fewer variables than cases.

In addition, we considered canonical correlation analysis, which was found to be similar to a combination of PCR and MLR, but lacking the flexibility of PLSR.

The y-variables to be predicted are exceedances of the 90th percentile of wind speed at 153 grid points over Europe, accumulated into annual winter series from October to March. The predictor (x) variables are 4 atmospheric and 21 SST indices. Thus, the MLR uses 25 predictors per grid square and the PCR optimises 25 predictors per grid square. The PLSR is used both on a grid square basis, optimising 25 predictors per grid square, and over the whole domain optimising 25 x 153 variables in one operation. The whole-domain use of PLSR is about 200 times faster than the analysis by grid square. A range of lags were tested, starting at 4 months and repeating the analysis with progressively longer lags at one-month increments. This method gave an optimal lag at 10 months, based on a cross-validated minimum PRESS (Prediction REsidual Sum of Squares). Essentially, therefore, the exceedances are being predicted by a set of predictors drawn from the previous winter.

In order to compare model performance, the root mean square error (RMSE) was computed between observed and predicted exceedances across all grid squares. For the models with a lag of 10 months, the RMSE is:

- 2231 for MLR
- 1498 for PCR
- 1421 for PLSR over the whole domain
- 1365 for PLSR by grid square

The ranked order of these values remains the same at all lags. Thus, MLR is shown to be much inferior to PCR and PLSR. Also, the PLSR optimising both x and y variables provides better forecasts than the PCR, which took about 200 times longer.

Figure 4 shows predicted exceedances for the 1999/2000 winter using PLSR on a grid square basis. The errors (observed minus predicted number of exceedances) are shown in Figure 5. Although some grid points have very large errors, there are many regions where the forecast error is quite small. Generally, the model tends to underpredict in areas of high wind speed and overpredict where wind speeds

FIGURE 4 Predicted exceedances of the 90th percentile for the 1999/2000 winter using PLSR in grid-square mode.

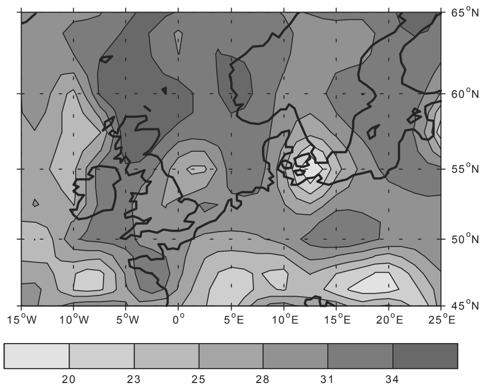


FIGURE 5 Errors in the prediction of Figure 4, expressed as the number of exceedances, for the PLSR in grid square model.

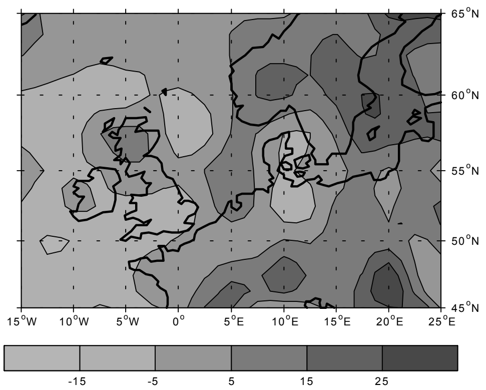
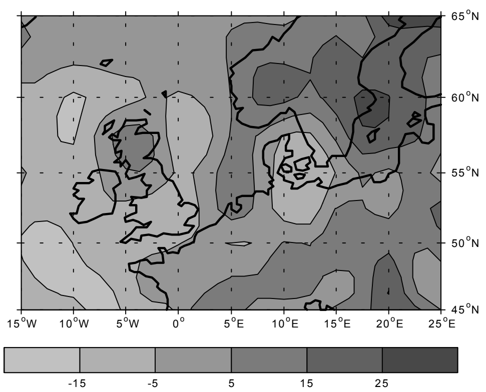


FIGURE 6 Errors for the PLSR model for the whole domain for the winter of 1999/2000, expressed as the observed minus predicted number of exceedances.



are lowest. Comparison of Figures 5 and 6 (the latter showing errors for PLSR applied to the whole domain) suggests that the penalty for optimising predictor and predictand at the same time is relatively small. The analysis by grid square has less error in some areas but the patterns are very similar

5. CONCLUSIONS

The comparison of models in Section 4 suggests that the PLSR models generate the most accurate results in this application. Although the PLSR model in grid square mode is superior overall, where computation time is a consideration it may well be that PLSR over the whole domain is to be preferred.

The method of optimising used by PLSR over the whole domain offers further potential advantages. As used here, it reduces the x and y variance to a single "latent variable" giving a minimum PRESS statistic. However, this is minimizing the error over all grid squares. Experimenting with more latent variables could increase the error in some areas and reduce it in others. This approach could be used to minimise the forecast error for particular areas of interest at the expense of others.

ACKNOWLEDGEMENTS

This work was funded by the TSUNAMI initiative under Contract A2P06 'Investigation of the risk from climate change variability over Northern Europe'.

REFERENCES

Francis, R.I.C.C., and J.A. Renwick, 1998: A regression-based assessment of the predictability of New Zealand climate anomalies. *Theor. Appl. Climatol.*, **60**, 21-36.

Goddard, L., S.J. Mason, S.E. Zebiak and others, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol.*, **21**, 111-1152.

Krishna Kumar, K., M.K. Soman, and K. Rupa Kumar, 1995: Seasonal forecasting of Indian summer monsoon rainfall: a review. *Weather*, **50**, 449-466.

Lehmiller, G.S., T.B. Kimberlain, and J.B. Elsner, 1997: Seasonal prediction models for North Atlantic basin hurricane location. *Mon. Wea. Rev.*, **125**, 1780-1791.

Lloyd-Hughes, B., and M.A. Saunders, 2001: Seasonal prediction of European spring precipitation from ENSO and local sea surface temperatures. Accepted for publication in *Int. J. Climatol*.

Merola, G. M., and B. Abraham, 2001: Dimensionality reduction approach to multivariate prediction. *Canadian Journal of Statistics*, **29**, 191-200.

Rimbu, N., H. Le Treut, S. Janicot, C. Boroneant, and C. Laurent, 2001: Decadal precipitation variability over Europe and its relation with surface atmospheric circulation and sea surface temperature. *Quart. J. Roy. Meteorol. Soc.,* **127**, 315-329.

Rodwell, M.J., and C.K. Folland, 2001: Atlantic air-sea interaction and seasonal predictability. *Q. J. R. Meteor. Soc.*, **127**, 1-31.

Ward, N.M., 1998: Diagnosis and short-lead prediction of summer rainfall in tropical North Africa at interannual and multidecadal timescales. *J. Climate*, **12**, 3167-3191.