

## Can knowledge of the state of the stratosphere be used to improve statistical forecasts of the troposphere?

By A. J. CHARLTON<sup>1</sup>\*, A. O'NEILL<sup>2</sup>, D. B. STEPHENSON<sup>1</sup>, W. A. LAHOZ<sup>2</sup> and M. P. BALDWIN<sup>3</sup>

<sup>1</sup>*Department of Meteorology, University of Reading, UK*

<sup>2</sup>*Data Assimilation Research Centre, University of Reading, UK*

<sup>3</sup>*Northwest Research Associates, Bellevue, Washington, USA*

(Received 23 December 2002; revised 30 May 2003)

### SUMMARY

Recent analysis of the Arctic Oscillation (AO) in the stratosphere and troposphere has suggested that predictability of the state of the tropospheric AO may be obtained from the state of the stratospheric AO. However, much of this research has been of a purely qualitative nature. We present a more thorough statistical analysis of a long AO amplitude dataset which seeks to establish the magnitude of such a link.

A relationship between the AO in the lower stratosphere and on the 1000 hPa surface on a 10–45 day time-scale is revealed. The relationship accounts for ~5% of the variance of the 1000 hPa time series at its peak value and is significant at the 5% level. Over a similar time-scale the 1000 hPa time series accounts for ≤1% of itself and is not significant at the 5% level. Further investigation of the relationship reveals that it is only present during the winter season and in particular during February and March. It is also demonstrated that using stratospheric AO amplitude data as a predictor in a simple statistical model results in a gain of skill of ~5% over a troposphere-only statistical model. This gain in skill is not repeated if an unrelated time series is included as a predictor in the model.

KEYWORDS: Arctic Oscillation Multiple regression

### 1. INTRODUCTION

Much interest in the atmospheric science community has been generated by the recent papers of Baldwin and Dunkerton (1999, 2001). Their work examines the variability of the stratosphere and troposphere in the context of the first empirical orthogonal function (EOF) of wintertime surface pressure, the so called Arctic Oscillation (Thompson and Wallace 1998). The Arctic Oscillation (AO) as described by Baldwin and Dunkerton represents a 'mode of variability' of the stratosphere and troposphere. Baldwin and Dunkerton present evidence that knowledge of the state of the AO in the stratosphere provides information about the future state of the troposphere on extended-range (10–30 days) and long-range (beyond 30 days) time-scales.

The long-held view of connections between the stratosphere and troposphere has been that the stratosphere responded passively to a tropospheric wavemaker. An early example of this paradigm is the work of Matsuno (1971) who was able to demonstrate that the structure of a stratospheric sudden warming could be simulated in a model with prescribed tropospheric planetary wave forcing. In the intervening 30 years there has been a wealth of other studies which have been able to demonstrate that a large part of the variability in the stratosphere is associated with changes in the tropospheric circulation. Examples of this include Lahoz (2000) which showed that a good representation of stratospheric variability could be made by forcing a general-circulation model with observed sea surface temperatures. A review of these ideas can be found in O'Neill (2003).

While it is true that much of the variability seen in the stratosphere may have a tropospheric origin, the papers of Baldwin and Dunkerton suggest that there may be a relationship between the current state of the stratosphere and the future state of the

\* Corresponding author: Department of Meteorology, University of Reading, PO Box 243, Reading, Berkshire RG6 6BB, UK. e-mail: a.j.charlton@reading.ac.uk

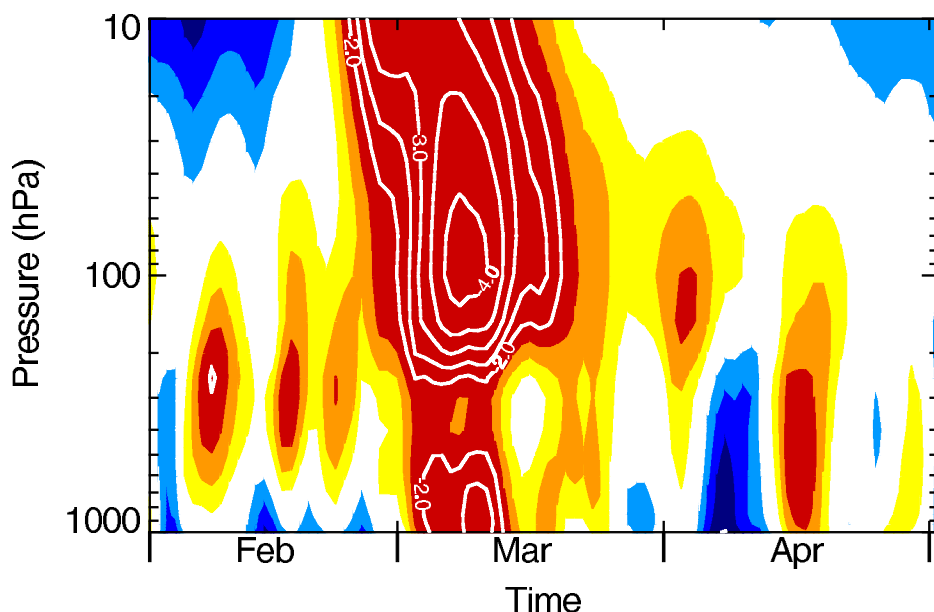


Figure 1. Time–height cross-section of Arctic Oscillation amplitude data for February to April 1999. Red colours indicate negative values of Arctic Oscillation amplitude and blue values indicate positive values. Shading is as Baldwin and Dunkerton (1999).

troposphere. This paper is concerned with examining the link between the stratosphere and troposphere in this sense and its potential use in tropospheric forecasting.

Figure 1 shows a section from a daily amplitude AO time series (Baldwin and Dunkerton 1999). The section centres around a stratospheric sudden warming event in late February 1999. Large negative anomalies (shaded in red colours) appear throughout the troposphere and stratosphere. In the stratosphere this represents a weakening of the strong wintertime polar vortex. In the troposphere the changes are harder to interpret, but in general negative anomalies represent a weakening of the jet structures in both the Pacific and the Atlantic. After the peak of the event, negative values of the AO persist in the lower stratosphere and upper troposphere (50–250 hPa) much longer than in the middle stratosphere (10–50 hPa). There is also some suggestion that these persistent negative amplitudes in the lower stratosphere and upper troposphere are linked to negative values in the middle and lower troposphere. It has been proposed that it would be possible to extract predictability of the troposphere from such events in the stratosphere.

Baldwin and Dunkerton (2001) composite such large-amplitude stratospheric AO events based on thresholds of +1.5 and –3.0 non-dimensional AO amplitude at 10 hPa. The resulting composite AO events show the mean behaviour of the AO throughout the atmosphere after such an event. In both the strong positive (+1.5) and strong negative (–3.0) composites the amplitude of the AO in the lower stratosphere and upper troposphere following the large AO amplitude in the middle stratosphere is on average of the same sign. This anomaly in the lower stratosphere and upper troposphere persists for up to 60 days after the start of the composite, much longer than the anomaly in the middle stratosphere.

Using a similar thresholding technique for 10 hPa zonal-mean zonal wind, Thompson *et al.* (2002) determine a large difference in surface temperature between

TABLE 1. DATASETS

Name	Description	Levels	Time range	Source
AO	Daily amplitude of Arctic Oscillation	17	1958–2000	Baldwin and Dunkerton (1999)
$\bar{u}$	Zonal-mean zonal wind at 60°N	17	1979–2000	ERA-15 and ECMWF operational analysis
$\Phi'$	Geopotential-height r.m.s. error from zonal mean at 60°N	17	1979–2000	ERA-15 and ECMWF operational analysis
$\Phi'$ (filtered)	As above but using geopotential-height dataset truncated at zonal wave number 2	17	1979–2000	ERA-15 and ECMWF operational analysis

opposite phases of the stratospheric AO, up to 60 days after the peak of the stratospheric event.

Much of the previous work in this area has used descriptive statistical techniques, such as the thresholding analysis of Baldwin and Dunkerton (2001), to highlight the relationship between the stratosphere and troposphere. The purpose of this work is to evaluate quantitatively the relationship between the stratosphere and troposphere and its statistical robustness. Our approach to this problem is to examine the predictive capability of the stratosphere to forecast the troposphere in terms of the AO pattern. To do this we use multiple linear regression techniques. This should be seen as the next logical step in the level of complexity of statistical techniques applied to AO datasets.

The method differs from the thresholding methods used previously in a number of important ways. First it uses all of the data available, rather than pre-selecting only large events. Second, it also allows us to quantify the size of any potential relationship. Third it allows comparison between the size of relationships between the troposphere and itself and the stratosphere and troposphere.

The datasets and modelling procedure used are outlined in section 2. Section 3 evaluates the suitability of the statistical model. Section 4 presents the results examining the relationship between the AO at 70 and 1000 hPa and section 5 examines the relationship at other levels. Section 6 determines the stationarity of the statistical relationship and section 7 presents the forecasting skill of the model. Section 8 presents conclusions.

## 2. DATASETS AND METHODOLOGY

### (a) Datasets

The datasets used in the study are summarized in Table 1. The daily AO amplitude time series used is described in Baldwin and Dunkerton (1999). It contains the amplitude of the AO on 17 pressure levels extracted from NCAR/NCEP\* reanalysis geopotential-height data between 1958 and 2000. For technical details see Baldwin and Dunkerton (1999).

We also examine connections between the stratosphere and troposphere in other datasets. This provides a test of relationships found in the AO dataset which could be a product of the AO diagnostic. These datasets are zonal-mean diagnostics traditionally used in stratospheric analysis. They only consider the variability around one latitude circle and may be less representative of the variability over the northern hemisphere as a whole. If any relationship between the variables can be found in non-AO diagnostics it would suggest the relationship is robust and not a product of the AO diagnostic.

\* National Center for Atmospheric Research/National Centers for Environmental Prediction.

These extra diagnostics were extracted from European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA-15) (Gibson *et al.* 1997) and ECMWF Operational Analysis datasets held at the British Atmospheric Data Centre.

Before any analysis is performed the mean annual cycle is removed from the ERA-15 datasets. This prevents the annual cycle from contaminating the results. All datasets are then standardized to have a mean of zero and standard deviation of one. This allows us to simplify some of the equations describing the relationships between variables in the statistical model.

In this study we focus on the use of daily data. It has been suggested that the signal-to-noise ratio could be reduced by smoothing data in some way. In this paper we focus on daily data as this is the simplest way of addressing the problem.

### (b) Methodology

To investigate the relationship between the stratospheric and tropospheric parts of the AO we construct a linear statistical model. This model attempts to quantify the effect of relationships between the stratospheric AO and the tropospheric AO and the tropospheric AO and itself. This is the next logical step from the work of Baldwin and Dunkerton. It attempts to quantitatively test ideas that are implicit in the compositing techniques employed by both Baldwin and Dunkerton (2001) and Thompson *et al.* (2002).

By fitting the model for a variety of lags between different time series we are able to examine the time-scales on which each of these relationships is important and how large the relationships are.

The statistical model is given by

$$y_z(t + \tau) = \beta_0(\tau)y_z(t) + \beta_1(\tau)x_z(t) + \epsilon(t), \quad (1)$$

where  $y_z(t)$  is the AO index on a pressure surface  $z$  at time  $t$  (in days),  $x_z(t)$  is the AO index on a different pressure surface at time  $t$ ,  $\tau$  is the lag and  $\epsilon(t)$  is a residual error. The parameters  $\beta_0(\tau)$  and  $\beta_1(\tau)$  are both functions of lag and can be estimated using least-squares regression. The statistical properties of the error determine the suitability of the model to the dataset. If the model is a good fit, that is to say it is a good representation of the dataset, the residual time series should be serially independent and normally distributed. Our criteria for a good model fit do not depend on the size of the error, that is to say a model may be a 'good fit' to the data even if the error term is very large.

Fitting the model involves estimating parameters  $\beta_0(\tau)$  and  $\beta_1(\tau)$  of the model using ordinary least squares. Repeating this for a range of values of the lag parameter  $\tau$  produces a set of model parameter estimates as a function of lag.

When using a multiple regression model with two predictors and standardized data, the autocorrelation  $\rho(y_z(t + \tau), y_z(t))$ , in the tropospheric time series can be decomposed into the sum of a *direct relationship*  $\beta_0(\tau)$  and an *indirect relationship*  $\rho(y_z(t), x_z(t))\beta_1(\tau)$ . Further details of this approach are given in Junge and Stephenson (2002).

$$\rho(y_z(t + \tau), y_z(t)) = \beta_0(\tau) + \rho(y_z(t), x_z(t))\beta_1(\tau). \quad (2)$$

The series of relationships represented by the model is shown in Fig. 2. The 'path' from the troposphere at some given time to the troposphere at some later time represents the direct relationship  $\beta_0(\tau)$ . The 'path' from the stratosphere to the troposphere, taking into account the mutual correlation between the stratosphere and the troposphere, represents the indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$ .

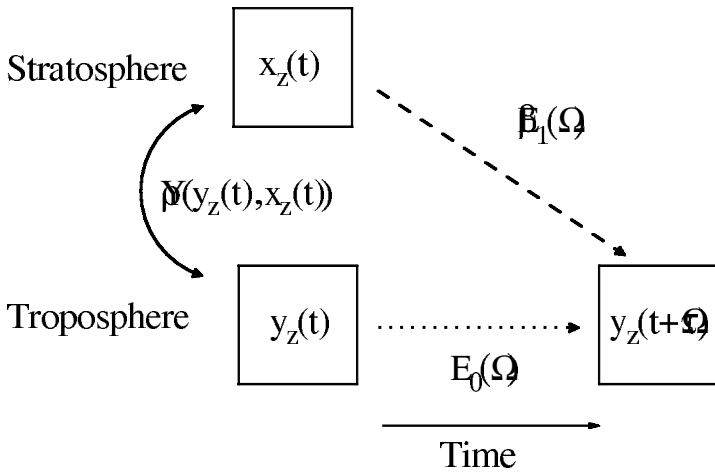


Figure 2. Idealized view of regression relationships. Squares represent state of atmosphere at some time. Dotted arrow indicates direct influence of troposphere on itself, dashed arrow indicates influence of stratosphere on troposphere, solid curved arrow indicates instantaneous correlation between troposphere and stratosphere.

The parameters  $\beta_0(\tau)$  and  $\beta_1(\tau)$  represent correlations between the time series. While correlations give no information about causality, a statistically significant correlation between a value at some time  $t$  and a value at time  $t + \tau$  can be exploited for predictive purposes.

We do not suggest that this is the best method of understanding the links in the AO, since the statistical method relies on linear statistical relationships between variables. The question we are asking is: Can we apply a statistical model to AO variables to gain useful predictive skill?

### 3. VALIDITY OF MODEL

#### (a) Evidence for nonlinearity of relationship between variables

Baldwin and Dunkerton use thresholding techniques to determine the relationship between the stratosphere and troposphere in the case of large amplitude stratospheric AO events. This technique uses only the end points of the AO dataset. An issue that arises from this analysis is whether the statistical relationship between the stratosphere and troposphere is the same for points in the middle of the distribution as it is for the end points of the distribution. In other words, is there a nonlinear relationship between the stratosphere and the troposphere present in the data? We try to answer this question by examining scatter plots of the AO dataset.

In Fig. 3(a) the AO amplitude at 1000 hPa is plotted against the AO amplitude at 70 hPa. This level is chosen for illustration; the conclusions in this section are true for other levels in the upper troposphere and lower stratosphere (50–250 hPa).

In these plots we use only winter, November to March (NDJFM), data. The variance of the 70 hPa AO in the summer months is much less than in the winter. There is also a much weaker correlation between the 70 and 1000 hPa AO during the summer months (this is discussed in section 4(c)). To determine if the relationship between the 70 and 1000 hPa AO is linear we have to consider data which are not affected by this seasonal change in correlation. For this reason we only consider winter data in the following analysis.

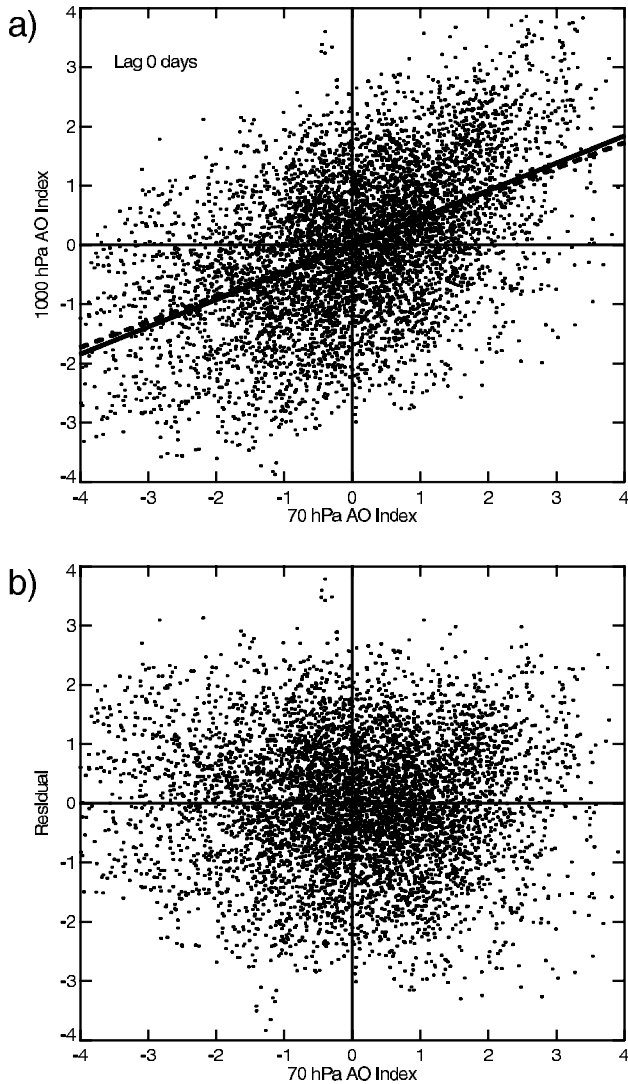


Figure 3. (a) Scatter plot showing 1000 hPa Arctic Oscillation (AO) index against 70 hPa AO index for lag = 0 days. Solid line shows fit with all of the data. Dashed line shows fit for data with an AO amplitude of magnitude 1 or less at 70 hPa. (b) Scatter plot showing residuals from fit with all data plotted against 70 hPa AO index.

Figure 3(a) shows a general ellipsoidal shape. If the relationship between the variables were nonlinear and dependent upon the value of the 70 hPa AO, a scatter plot of the two variables would show a general random cloud of points in the centre of the diagram and an ellipsoidal shape at one or both ends of the distribution.

A simple test of the linearity of the relationship between the stratosphere and troposphere can be performed by making a linear fit to different parts of the data. A linear fit to all the data is shown as a solid line in Fig. 3. Data is then sub-sampled to include only points at 70 hPa which have a magnitude less than 1 non-dimensional AO amplitude. This is shown in the dashed line. In Fig. 3 the slope of both of the lines is very similar. This shows that the correlation between the 70 and 1000 hPa AO amplitude

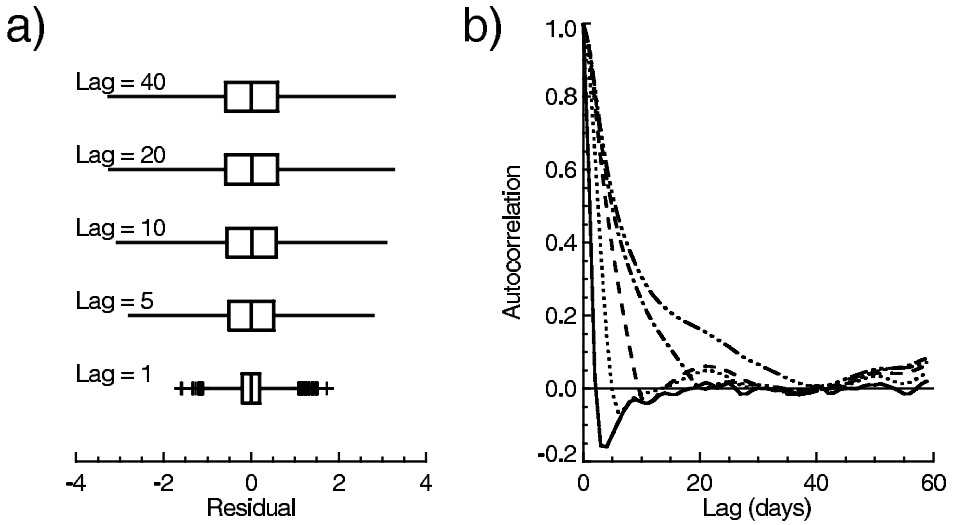


Figure 4. Distribution of residuals of model fit. (a) Box plots of the distribution of residuals from the same model fit. Central line of box shows median residual, outer lines of box shows upper and lower quartile. Whiskers are plotted at 1.5 times the inter-quartile range. Crosses show data points outside 1.5 times the inter-quartile range. (b) Autocorrelation function of residuals when model is fitted using 1000 hPa as the predictand series and 70 hPa as the predictor series. Autocorrelation is shown for model lag of 1 day (solid line), 5 days (dotted line), 10 days (dashed line), 20 days (dot-dash) and 40 days (triple-dot dash).

for small values of 70 hPa AO is very similar to the correlation when using all of the data. This suggests that the relationship between the AO at 70 and 1000 hPa does not depend on the amplitude of the AO at 70 hPa.

An alternative technique to test for nonlinearity in the relationship between the 70 and 1000 hPa AO is to use lowess smoothing (Chambers *et al.* 1983). A fit to the data using this technique (not shown) shows that there is some weak nonlinearity for extreme values of the 70 hPa AO index, but that a linear fit to the data provides a good representation of the variability.

Figure 3(b) shows a scatter plot of the residuals about the linear fit to all of the data plotted against the 70 hPa AO amplitude. The residuals show little dependence on the value of the 70 hPa AO. If there were a nonlinear relationship between the 70 and 1000 hPa AO then we would expect to see a dependency of these residuals on the AO amplitude at 70 hPa.

### (b) Residual diagnostics

It is important to establish the suitability of the statistical model to the datasets investigated. The criteria we use to judge if the model is a good fit to the dataset is that the residuals should be serially uncorrelated and normally distributed.

Figure 4 shows some diagnostics of the residuals for a fit of the model where the predictand series  $y_z(t + \tau)$  is the 1000 hPa AO time series and the predictor series  $x_z(t)$  is the 70 hPa AO time series. Figure 4(a) shows box plots for a number of different model lags between 1 and 40 days. We define a good model to have normally distributed residuals. The box plots show that the residuals have a median value close to zero and are symmetrically distributed about this median. This indicates that there is no bias in the model and the residuals left over from the model fit are approximately Gaussian noise.

Figure 4(b) shows the autocorrelation of residuals for a number of different model lags between 1 and 40 days. A good model fit is defined as one in which the residuals are independent. In this case, the autocorrelation of residuals should decay rapidly with increasing lag. At small model lags (1–10 days) this is the case; but for larger model lags (20–40 days) the residual autocorrelation remains large beyond 10 days. This is common in atmospheric data (Wilks 1995, section 5.2.3) and is a product of time-dependence in the data used to construct the model.

Ignoring serial correlation in the data can lead to an underestimate of the variance of the sampling distribution and hence to over confidence in the significance of a hypothesis test. In order to account for the time dependence of the data when calculating the significance of the model correlations, we reduce the degrees of freedom in our hypothesis test by a factor proportional to the typical time between uncorrelated points in the input dataset (Wilks 1995, section 5.2.3).

The largest autocorrelation in the AO time series is found at 10 hPa. This time series has a decorrelation time of approximately 10 days. We reduce the number of degrees of freedom in all our significance testing calculations by a factor of 10 in line with this result. Although this technique is not ideal it provides a good indication of the significance of the model correlations.

We conclude that the model is a good fit to the dataset as defined by our criteria.

#### 4. EXAMINING CONNECTIONS BETWEEN THE LOWER STRATOSPHERE AND LOWER TROPOSPHERE

In the following section the model is fitted using the 1000 hPa time series as the  $y$  time series in our model (see Eq. (1)) and the 70 hPa time series as the  $x$  time series in our model. We chose 70 hPa to illustrate points which are generalized to include a range of levels in the upper-troposphere/lower-stratosphere (UTLS) region (which we define here as between 50 and 250 hPa) in section 5.

Statistical testing of many of the results is conducted. This testing uses a Student's  $t$ -test. Results referred to as 'significant at the 5% level' refer to the test being conducted at 95% confidence. That is to say, there will be a 5% chance of a false-positive result.

##### (a) Whole-year behaviour

The model described by Eq. (1) was fitted to the AO dataset (Table 1) for a range of lags. The parameters of the fit are shown in Fig. 5(a).

If the time series at 1000 hPa were dependent only upon itself then we could model the AO time series at 1000 hPa as an AR(1) or red-noise process (Chatfield 1996):

$$y(t+1) = \alpha y(t) + \epsilon(t+1). \quad (3)$$

The autocorrelation of the time series at lag 1 is equal to  $\alpha$  and is less than one for a stationary time series. Substituting shows that

$$y(t+2) = \alpha(\alpha y(t) + \epsilon(t)) + \epsilon(t+1),$$

and so, in general,

$$y(t+\tau) = \alpha^\tau y(t) + \sum_{l=0}^{\tau-1} \alpha^{\tau-l} \epsilon(t+l). \quad (4)$$

There is an exponential decay of the autocorrelation with lag:  $\rho(\tau) = e^{\tau \ln \alpha}$ .



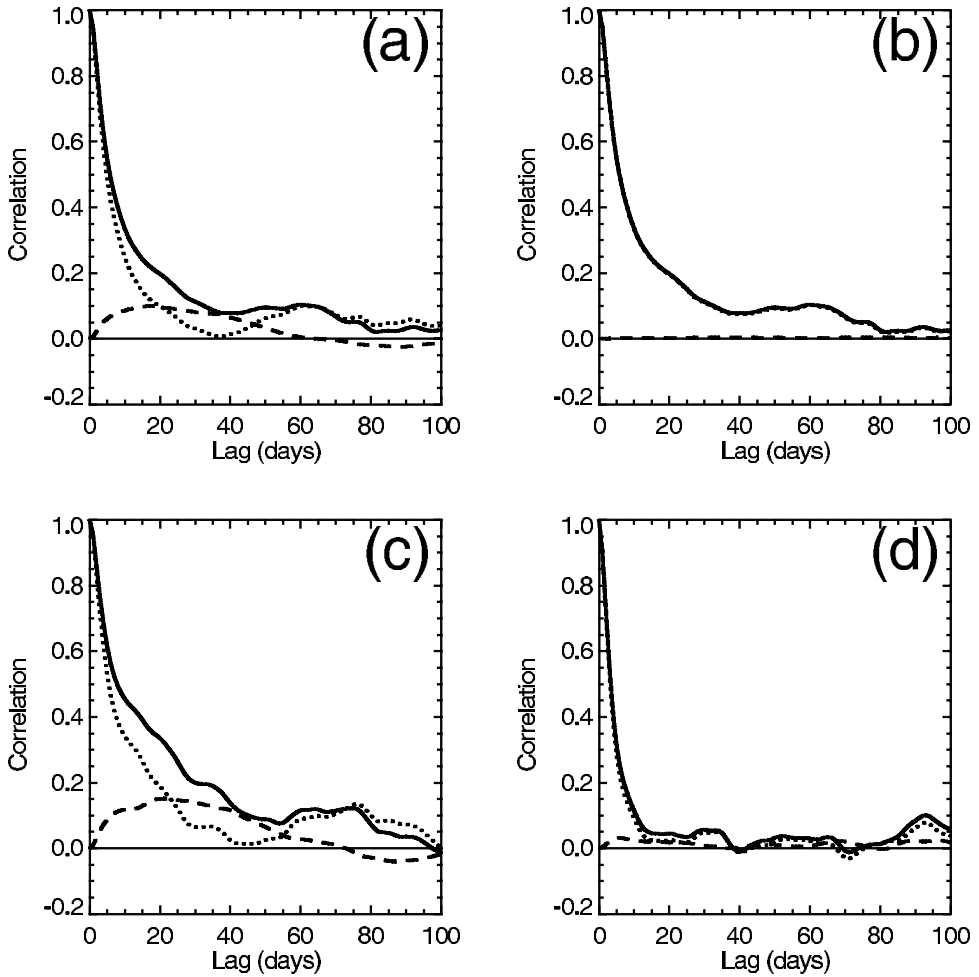


Figure 5. Decomposition of autocorrelation of 1000 hPa AO using the model in Eq. (1), where 70 hPa is the  $x_z$  series and 1000 hPa the  $y_z$  series. Autocorrelation of 1000 hPa series  $\rho(y_{1000}(t + \tau), y_{1000}(t))$  is shown in solid line,  $\beta_0(\tau)$  in dotted line and the product  $\beta_1(\tau)\rho(y_{1000}(t), x_{70}(t))$  in dashed: (a) using all of the data, (b) with  $y_{1000}$  time series reversed in time, (c) for December–February data only, and (d) for June–August data only.

Parameters from the fit of the model show that over the medium-range time-scale (1–10 days lag) the decay of the autocorrelation function is near to exponential. Exponential decay of the autocorrelation with increasing lag over these time-scales means the 1000 hPa AO time series could be modelled as an autoregressive process. The direct relationship  $\beta_0(\tau)$  is much larger than the indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$ . This suggests that only the direct relationship  $\beta_0(\tau)$  is important on 1–10 day time-scales.

On extended range (10–30 days lag) and slightly longer (30–45 days lag) time-scales, the decay of the autocorrelation function is less than exponential. The direct relationship  $\beta_0(\tau)$  is much smaller than the autocorrelation and is not significant at the 5% level. The indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$  increases in magnitude and is significant at the 5% level. On 10–45 day time-scales the direct relationship accounts for  $\leq 1\%$  of the variance of the 1000 hPa time series. In contrast, the indirect relationship accounts for  $\sim 5\%$  of the variance of the 1000 hPa time series. Although both the direct

relationship and the indirect relationship account for very small amounts of the variance of the 1000 hPa time series, the indirect relationship accounts for a larger proportion of the variance than the direct relationship. It can be inferred from these results that a significant though small statistical relationship between the AO at 70 and 1000 hPa is seen on time-scales of 10–45 days.

On much longer time-scales (45–100 days lag) the autocorrelation of the 1000 hPa time series becomes smaller. The indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$  is much reduced and is not significant at the 5% level. The direct relationship  $\beta_0(\tau)$  accounts for most of the autocorrelation of the 1000 hPa dataset.

These results suggest a small statistically significant relationship between the 70 and 1000 hPa AO exists on 10–45 day time-scales. The autocorrelation of the 1000 hPa dataset on these time-scales is accounted for mainly by the indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$ .

Ambaum and Hoskins (2002) examined the autocorrelation of a smoothed North Atlantic Oscillation (NAO) index. They comment that the decrease in the rapid decay of the autocorrelation of the NAO index starting at 10 days (which they refer to as ‘shouldering’) may be due to a relationship with the stratosphere. Our result shows that the reduction in the decay of the autocorrelation between 10–45 days is related to a connection between the lower stratosphere and the troposphere. This result also agrees with a model study by Norton (2003), who found a significant difference in the autocorrelation of the tropospheric AO on 10–25 day time-scales when changes were made to the stratospheric variability.

### (b) *Time-order dependence*

There is a large difference in the statistical properties of the AO amplitude at 70 and 1000 hPa. In particular the autocorrelation of the AO at 70 hPa is substantially larger than the autocorrelation of the AO at 1000 hPa for the same lag. It could be suggested that the statistical relationship between the 70 and 1000 hPa AO highlighted in section 4(a) is due to the difference in autocorrelation of the 70 and 1000 hPa time series.

A simple way to test this hypothesis is to fit the model with the same 70 hPa time series and a time-reversed copy of the 1000 hPa time series. The autocorrelation of the new reversed 1000 hPa time series is identical to the normal 1000 hPa time series. If the statistical relationship highlighted in section 4(a) is due to the difference in autocorrelation of the 70 and 1000 hPa time series, then a fit with the 70 hPa and reversed 1000 hPa time series will show correlations identical to the fit with the 70 hPa and normal 1000 hPa time series.

The parameters of the model fit with the 70 hPa AO time series and the time reversed 1000 hPa AO time series are shown in Fig. 5(b). There is no evidence of a similar increase in the value of the indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$  on 10–45 day time scales as is seen in Fig. 5(a). Therefore it can be inferred that the small, statistical relationship between the 70 and 1000 hPa AO amplitude on 10–45 days is a product of the particular time orientation of the 1000 hPa AO time series.

### (c) *Winter and summer behaviour*

Baldwin and Dunkerton (1999) found that connections between the stratospheric and tropospheric parts of the AO only occur during the winter season. To quantitatively investigate this seasonal dependence the model was fitted to subsets of the AO dataset which only included winter (DJF) and summer (JJA) data. In order to keep a constant

data size between fits at different lags, the data for the predictor  $x_{70}(t)$  and  $y_{1000}(t)$  series included all of that particular season (e.g. DJF) and the predictand  $y_{1000}(t + \tau)$  series is taken to be a slice of data of the same size displaced by the lag in question. For example the data for the DJF fit at 31 days lag would be DJF for the predictor (70 hPa AO) series and January to March (JFM) for the predictand (1000 hPa AO) series. The model parameters are shown in Figs. 5(c) and (d).

In DJF (Fig. 5(c)) the correlation structure of the model is very similar to the model fit with all of the data included (Fig. 5(a)). The DJF fit shows a peak in the indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$  over the 10–45 day range. The magnitude of the indirect relationship is larger than in the fit with all the data, suggesting that the main contribution to the relationship between the AO at 70 and 1000 hPa is in the winter season. In contrast no such structures are seen in the JJA fit. The indirect relationship remains very small at all lags and is not significant at the 5% level.

This confirms the suggestion that any connection between the stratosphere and troposphere is only likely to occur during the winter season. Baldwin and Dunkerton (1999) suggested that connections between the stratospheric and tropospheric parts of the AO were linked to stratospheric sudden warming events in the stratosphere. These events occur between December and March and are not present in JJA.

#### (d) *Month-by-month behaviour*

A further examination of the seasonality of the relationship is shown in Fig. 6. In this analysis we fit the model for subsets of the AO dataset which include data from each calendar month. As in the seasonal analysis, care is taken to preserve the data size for each regression.

Figure 6(a) shows the autocorrelation for each calendar month plotted against lag. Figure 6(b) shows the value of the direct, tropospheric correlation for each calendar month. Figure 6(c) shows the indirect, stratospheric correlation for each calendar month. It is important to remember that, although the plots are shown with contours, they represent 12 independent sets of 100 model fits and values between the marked months are artificial. Contouring is used as it makes the plots easier to read and interpret.

Figure 6(a) shows the autocorrelation of the 1000 hPa AO. In general this autocorrelation increases during the winter months. During January, February and March the autocorrelation decays slowly with lag, having values larger than 0.1 beyond 30 days lag (see B in Fig. 6(a)). The increase in the autocorrelation of the 1000 hPa AO in January (B) is attributable to the increase in the direct relationship  $\beta_0(\tau)$  seen in January (D). A similar increase in the direct relationship is not seen in February and March. The increase in autocorrelation in February and March (E) is due to an increase in the indirect relationship  $\rho(y_{1000}(t), x_{70}(t))\beta_1(\tau)$ .

The dynamics of the stratosphere in February and March are dominated by the break-up of the weakening stratospheric vortex. There is large variability in the timing of the break-up of the vortex (O'Neill 1995). In some years the vortex breaks down in late February with an early final warming. It is plausible that the larger values of the indirect relationship (E) in February and March are associated with the timing of the final warming. A final warming involves a reversal of the jet from winter westerly values to summer easterly values. Such a wind reversal is a major dynamical event in the stratosphere and as such might have a significant effect on the position of the tropopause and consequently the evolution of the troposphere.

There is also evidence of a relationship between the 70 hPa AO and the 1000 hPa AO during December and January (F) but the magnitude of the correlation is much

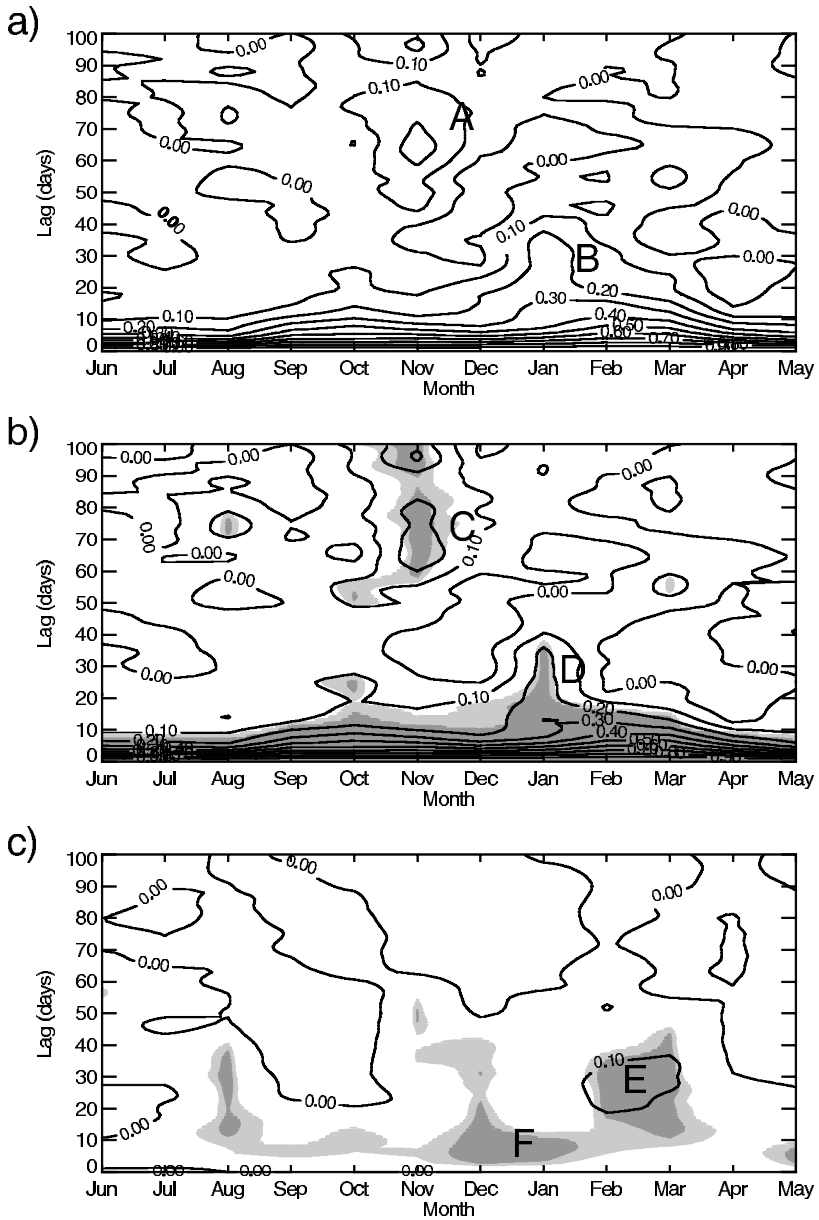


Figure 6. Contour plots of decomposition of 1000 hPa AO autocorrelation of 1000 hPa. (a) Autocorrelation of 1000 hPa AO as a function of month and lag, (b) direct effect  $\beta_0(\tau)$  as a function of month and lag, and (c) indirect effect  $\rho(y_{1000}(t), x_{70}(t))\beta_1(\tau)$  as a function of month and lag. Contour interval is 0.1. Dark shading shows correlation is significant at the 5% level, light shading shows correlation is significant at the 10% level. A–F mark salient features. See text for further details.

smaller and on shorter (5–10 day) time-scales. On these time-scales the direct effect is much larger.

Figure 6(a) also shows large autocorrelation at a lag of 60 days and greater during November (A). Feature A is accounted for by the large direct relationship in November (C). This suggests that the state of the tropospheric AO in early autumn has some influence on the evolution of the AO throughout the winter.

Fitting the model to monthly subsets of the AO datasets shows that the relationship between 70 hPa and 1000 hPa identified in section 4(a) is confined to February and March. This might suggest that the relationship between 70 and 1000 hPa might be linked to the timing of the final warming of the stratospheric vortex.

(e) *Relationship in simple diagnostics*

There are many questions about the suitability of the AO to fully represent the variability of the northern hemisphere. As a partial check of the robustness of the relationships between 70 and 1000 hPa established in section 4(a) using the AO dataset, we repeat the analysis using three other simple zonal-mean diagnostics. A relationship between 70 and 1000 hPa in these datasets would suggest the relationship found in the AO dataset is not a product of the AO diagnostic.

The datasets used are outlined in Table 1. Figure 7(a) shows correlations from a zonal-mean zonal-wind dataset. Figure 7(b) shows correlations from a filtered  $\Phi'$  dataset. This quantity is defined as follows.

$$\Phi' = \sqrt{(\Phi - \overline{\Phi})^2}, \quad (5)$$

where  $\Phi$  represents geopotential height and the overbar represents a zonal mean. Before calculating this diagnostic we filter the geopotential-height analysis to only include zonal wave numbers up to and including zonal wave number two. The stratosphere exhibits primarily low wave-number variability and it is reasonable to expect that any relationship between the stratosphere and troposphere is likely to occur through these wave numbers. The  $\Phi'$  diagnostic would include higher zonal wave-number variability in the troposphere which may confuse any relationship between the stratosphere and troposphere. A fit is also made with a  $\Phi'$  diagnostic extracted from the full, unfiltered geopotential-height analysis (not shown).

Parameters from the model fit using the two ERA-15 datasets are shown in Fig. 7. Both the  $\overline{u}$  and filtered  $\Phi'$  datasets have qualitatively similar correlation series to the AO data set. The indirect relationship  $\rho(y_{1000}(t), x_{70}(t))\beta_1(\tau)$  has larger values over the 10–45 day lag as in the AO dataset. Over a similar time-scale there is also a reduction in the size of the direct relationship  $\beta_0(\tau)$  as in the AO dataset. This indicates that the effect observed in the AO data is robust.

An interesting comparison can be made between the results for the filtered and unfiltered  $\Phi'$  datasets. The model parameters in the unfiltered  $\Phi'$  dataset (not shown) do not show correlation similar to that in the other three datasets. There is no evidence of a relationship between the UTLS region and the surface. The stratosphere is dominated by variability at the low planetary wave numbers. Any relationship between the stratosphere and troposphere is likely to occur on this large scale. This is confirmed by the differences in the filtered and unfiltered fits.

By fitting the same statistical model to  $\overline{u}$  and filtered  $\Phi'$  datasets it is possible to determine a similar connection between the lower stratosphere and troposphere without using an AO diagnostic. While the relationship in other diagnostics is smaller, its presence suggests that the relationship is robust and not a product of the AO diagnostic.

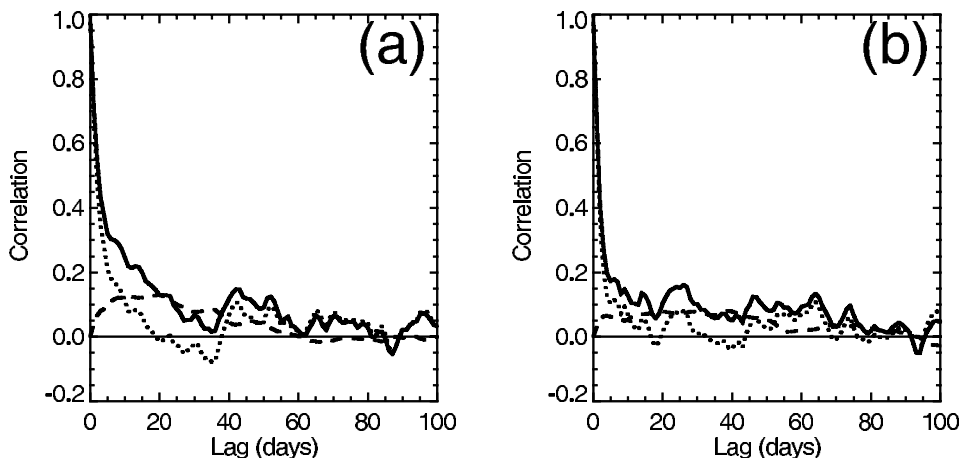


Figure 7. Decomposition of autocorrelation of 1000 hPa dataset as Fig. 5 but for (a) zonal-mean zonal-wind time series at 60°N and (b) filtered  $\Phi'$  time series at 60°N (see text). For more details of datasets see Table 1.

Even if the AO provides the best way of revealing a link between the stratosphere and troposphere, it is not certain that this link exists exclusively through a large-scale hemispheric change to the flow. Examination of geopotential-height anomaly maps at various times in the evolution of ‘downward propagating events’ shows a much more highly convoluted anomaly pattern than a simple hemisphere-scale exchange of mass between the polar cap and sub-polar latitudes.

##### 5. EXTENDING THE MODEL TO OTHER LEVELS

Fitting the model with 70 hPa as one of the predictors suggested that a relationship between the stratosphere and troposphere may exist. An extension of this approach to other pressure levels is necessary to fully understand the nature of the relationship. This is done by fitting the model with the stratospheric predictor  $x_z(t)$  replaced by each of the other levels in the dataset. The fit parameters for different levels are shown in Fig. 8. The parameters for each model are plotted at the corresponding pressure. For example, a cut across Fig. 8(a) at 70 hPa would produce the dotted line in Fig. 5(a) and a cut across Fig. 8(b) at 70 hPa would produce the dashed line in Fig. 5(a).

Figure 8(b) shows the large increase in the value of the indirect relationship  $\rho(y_{1000}(t), x_z(t))\beta_1(\tau)$  which can be seen on 10–60 day time-scales at the 70 hPa level (B). There are similar effects on surrounding levels (50–250 hPa), but this increase is smaller at levels in the middle stratosphere (50–10 hPa) and the middle and lower troposphere (250–925 hPa).

The large increase in the indirect relationship is accompanied by a similar decrease in the direct relationship  $\beta_0(\tau)$  (Fig. 8(a)). This reduction is largest in the same region between 50 and 250 hPa (A), but there is a general reduction in the significance for levels into the middle stratosphere. The indirect relationship has largest magnitude on the 150 hPa surface. This level is in the troposphere at most latitudes. It is therefore suggested that while some predictability of the 1000 hPa AO may be obtained from the UTLS region, the relationship with levels in the middle stratosphere is very weak. The relationship is strongest for levels near the tropopause. It is important to note that this region is influenced both by tropospheric and stratospheric dynamics.

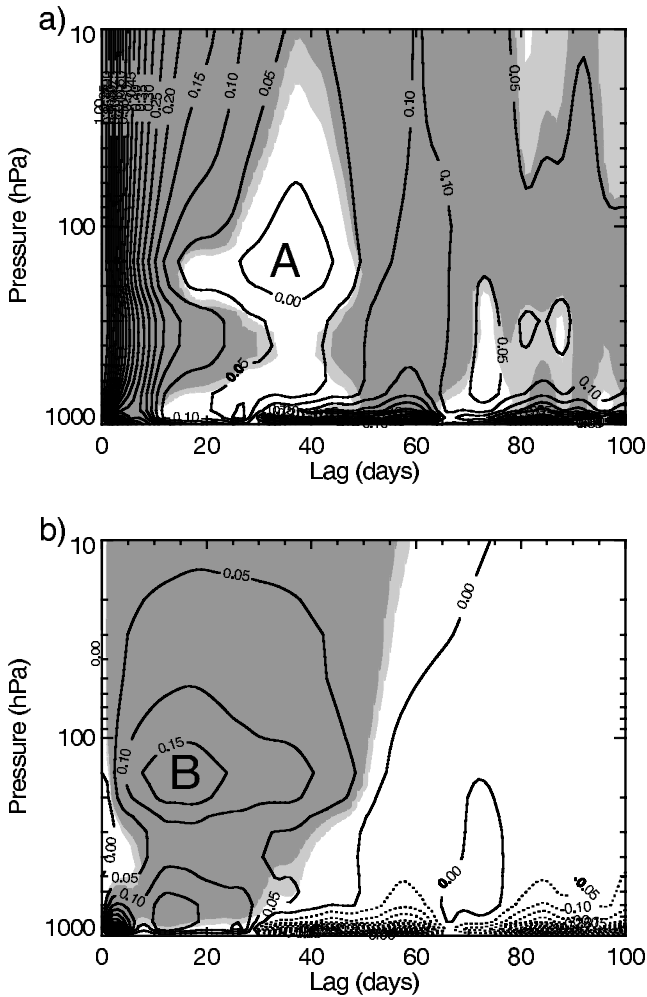


Figure 8. Model parameters (solid contours, negative values dotted) for various predictor levels. (a)  $\beta_0(\tau)$  for fits with various levels, and (b) product of  $\beta_1(\tau)$  and instantaneous correlation for various predictor levels. Regions where the parameters are significantly different from zero at the 5% level are shaded in dark grey, and at the 10% level are shaded in light grey. A and B mark salient features (see text for details).

It might be expected that the state of the AO near the tropopause has an impact on the surface AO; but the longer time-scale (10–45 days) of this link is unexpected. The long time-scale of this relationship requires further investigation in a dynamical context.

## 6. STABILITY OF RELATIONSHIP

In order to assess the stability of the relationship between upper levels and the surface AO, it is necessary to investigate the relationship for different sub-periods within the data record. To do this the data were split into a series of ten-year blocks and the model fitting procedure applied to each block. The model fit is made for the 1000 and 70 hPa levels in the dataset as in section 4. The lag is fixed at 30 days as the largest indirect correlation is seen at this lag. Other lags were investigated and it was

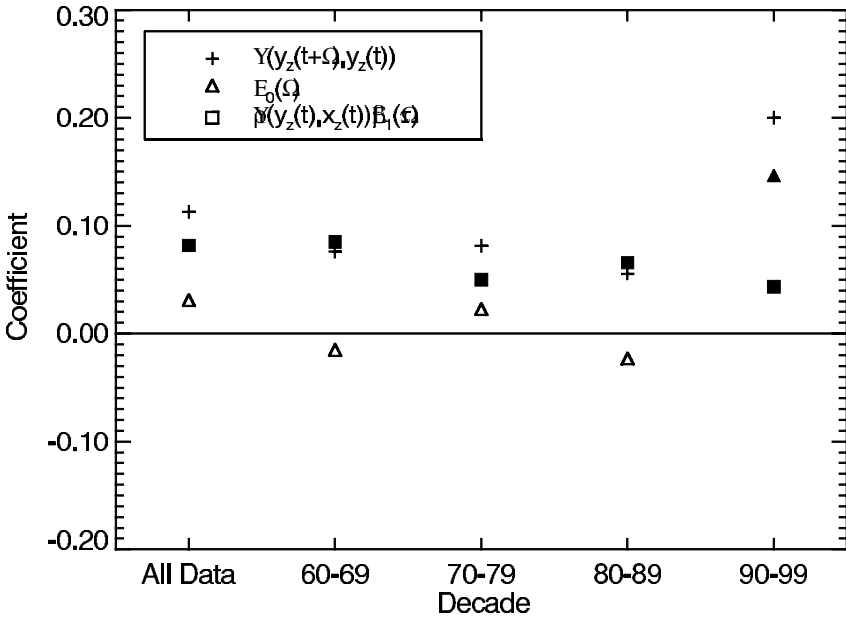


Figure 9. Model parameters for various subsections of data at lag of 30 days. The autocorrelation of the dataset is plotted with crosses,  $\beta_0(\tau)$  with triangles and  $\rho(y_z(t), x_z(t))\beta_1(\tau)$  is plotted with squares. Symbols for model parameters are filled in black if the parameter is significantly different from zero at the 5% level. See text for explanation.

found that the results were robust within the region of increased indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$  (10–45 days).

Figure 9 shows the autocorrelation  $\rho(y_{1000}(t + \tau), y_{1000}(t))$ , direct relationship  $\beta_0(\tau)$  and indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$  at 30-days lag for each decade of the data. The size of the indirect correlation is relatively constant between each decade and is of similar magnitude to the indirect relationship for the entire record. This suggests that the indirect relationship is stable throughout the data. It is also interesting that the relationship between 70 and 1000 hPa is relatively similar between decades with significantly different variability in the stratosphere. In particular the 1990s had relatively few stratospheric sudden warmings but the relationship is still statistically significant.

At greater lags (not shown) the size of the indirect relationship is reduced during the 1990s and not statistically significant. This suggests that on longer time-scales the relationship between the stratosphere and troposphere may be influenced more strongly by the number of large-amplitude stratospheric warming events.

The magnitude of the direct relationship  $\beta_0(\tau)$  (and therefore the autocorrelation, see Eq. (2)) is extremely variable between different decades. In particular, during the 1990s the direct correlation is very large at this lag. An examination at other lags (not shown) reveals that this is part of a large increase in the direct relationship between 20 and 60 days.

7. OUT-OF-SAMPLE LINEAR PREDICTIVE SKILL

The ultimate application of the relationships suggested by the Baldwin and Dunkerton dataset is to improve forecasting of the tropospheric AO and hence surface parameters. A simple experiment was constructed to test the forecasting capability



of this dataset. In order to test the fitted model it should be tested against an independent dataset. As no other dataset is available we divide the data in half and then fit the model for one half of the dataset and test it using the other half.

To assess the benefit of using stratospheric data to forecast the 1000 hPa AO we fit two different models to the dataset. The first one is structured as in Eq. (1). The second control model is shown in Eq. (6):

$$y_z(t + \tau) = \gamma_0(\tau)y_z(t) + \epsilon(t). \quad (6)$$

This model only has only one predictor, the state of the 1000 hPa AO at a previous time. It is not expected to be a good model of the future state of the 1000 hPa AO.

We measure the skill of each of the models by comparison with an AO climatology using the skill score (SS)

$$SS = 1 - \frac{MSE_{\text{forecast}}}{MSE_{\text{climatology}}}, \quad (7)$$

where MSE represents the mean square error of the forecast. The difference in SS between the two models gives a measure of the gain in skill obtained by including extra information in the model on each level.

Figures 10(a) and (b) show the difference in SS between the 1000 hPa only control regression model and the two-predictor model as in Eq. (1). Positive values indicate that including data at a particular pressure level and lag adds skill to forecasts of the 1000 hPa AO (compared with a 1000 hPa AO only model), and negative values indicate that including data at a particular lag and pressure level reduces skill to forecasts of the 1000 hPa AO (compared with a 1000 hPa AO only model). The skill is plotted for different lags and different pressure levels. The two columns show results when different halves of the dataset are used to train the model.

Figures 10(a) and (b) show that the SS of the two predictor model is greater than the 1000 hPa only control model in the lower and middle stratosphere (250–10 hPa) on time-scales between 10 and 60 days. This is the region highlighted in the model fit as the significant region for the indirect relationship  $\rho(y_z(t), x_z(t))\beta_1(\tau)$ . The magnitude of the increase is small at  $\sim 5\%$ .

In contrast, for levels in the middle and lower troposphere the SS of the two-predictor model and the 1000 hPa only control model is approximately comparable. The addition of extra information from the middle and lower troposphere into a statistical model of the 1000 hPa AO provides little extra forecast skill.

As was suggested in the introduction to this section, it was not expected that the troposphere-only model would provide useful skill on longer time-scales. Figures 10(c) and (d) show the actual skill for the model with only a 1000 hPa predictor and both a 1000 hPa predictor and a 70 hPa predictor. The region of increased skill highlighted above can clearly be seen between 10 and 60 days for both training periods. On 10–60 day time-scales the 1000 hPa only model has less than 5% skill. The inclusion of extra lower-stratospheric information results in a large increase in this skill. For example, at 20-days lag the forecast skill is increased from 5% to 10%. The inclusion of extra information in the two-level model does give a significant increase in the skill of 1000 hPa forecasts. Nevertheless, the actual forecast skill derivable from such a model is still small.

It could be suggested that the increase in skill is simply due to the addition of an extra predictor in the two-level model. This hypothesis can be tested by repeating the analysis with the 70 hPa time series reversed in time. In this case the 1000 hPa only control model will have identical skill and the additional predictor dataset will have

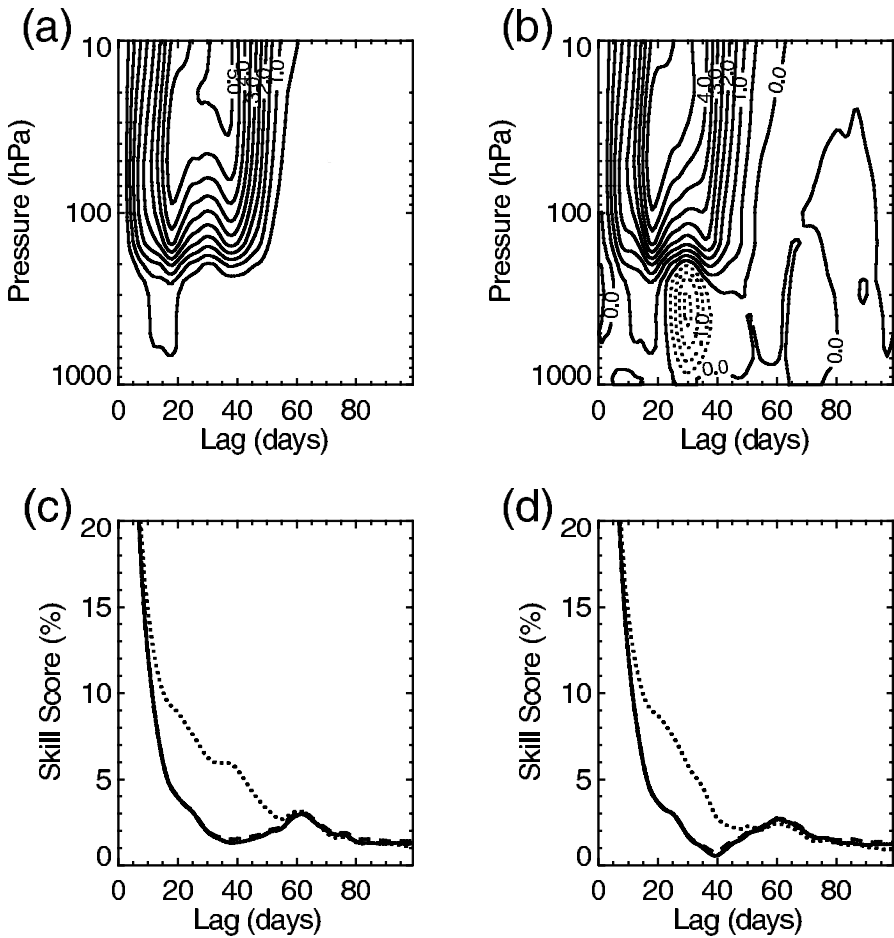


Figure 10. Skill score (SS) diagnostics: (a) difference in SS of multiple-predictor model and 1000 hPa predictor only model versus climatology in percentage gain in skill (1958–78 used as training period, predicting 1979–2000). Contour interval is 0.5%. Solid contours show the multiple predictor model has larger SS than the 1000 hPa predictor only model. See text for details of models. (b) As (a) but 1979–2000 used as a training period, predicting 1958–78. Dotted contours show the multiple-predictor model has smaller SS than the 1000 hPa predictor only model. (c) Actual percentage skill against lag for 1000 hPa only model (solid line), model with 1000 hPa and 70 hPa as model predictors (dotted line) and model with 1000 hPa and a time reversed 70 hPa time series as model predictors (dashed line); 1958–78 is used as training period, predicting 1979–2000. (d) As (c) but for 1979–2000 used as training period, predicting 1958–78.

identical statistical properties as in the normal fit. Any predictive relationship between the two datasets is destroyed. Therefore, if the gain in skill in this test is comparable to the gain in skill in the normal case, then this is likely to be due to the addition of an extra predictor.

The SS of a model which has a 1000 hPa AO predictor and a time-reversed 70 hPa predictor is shown in Figs. 10(c) and (d). It is hard to distinguish this from the skill of a 1000 hPa AO only model. This indicates that including extra, unrelated information with the same statistical properties as the 70 hPa time series results in a very small increase in skill. The gain in skill introduced by including extra lower-stratospheric information in a statistical model of the 1000 hPa AO represents a real increase in the forecasting skill of such a model.

The lack of increase in the skill for tropospheric levels is somewhat surprising. It might be expected that including information in the troposphere, which could have a direct impact on the development of individual weather systems in the middle troposphere, may lead to better forecasts of the AO. However, it seems that in terms of the AO the lower and middle troposphere contains very little information not contained in the 1000 hPa AO. Examination of AO time series, such as Fig. 1, suggests that the middle- and lower-tropospheric AO often have very similar variability in time. The suitability of the AO diagnostic in a forecasting context is therefore somewhat limited, as we do not suggest that a tropospheric forecasting model should not include information in the lower and middle troposphere. The gain in forecast skill presented here is for a forecast of the hemispheric-scale AO structure. A standard tropospheric forecast would be dominated on daily time-scales by more localized variability.

## 8. CONCLUSIONS

The papers of Baldwin and Dunkerton (1999, 2001) have suggested a relationship between the AO in the stratosphere and troposphere. Much of the previous analysis of this relationship has been qualitative. In particular, Baldwin and Dunkerton (2001) and Thompson *et al.* (2002) show that there is a large change in the mean tropospheric amplitude of the AO after the amplitude of the stratospheric AO crosses a particular threshold.

We have presented a quantitative examination of the relationship between the lower stratosphere and surface using a simple statistical model. The model relates the amplitude of the tropospheric AO at some time to the previous amplitude of the AO in the troposphere and the previous amplitude of the AO in the stratosphere.

A relationship between the amplitude of the AO in the lower stratosphere and 1000 hPa has been identified. Typical correlations between the lower stratosphere and 1000 hPa are small ( $\sim 0.2$ ), but significant (at the 5% level) over extended range time-scales (10–45 days).

The character of this relationship has been determined by further analysis.

(i) The relationship is most prominent in the upper troposphere/lower-stratosphere region (50–250 hPa). This region spans different parts of the atmosphere at different latitudes, but can broadly be thought of as the location of the tropopause.

(ii) The relationship is strongest during the winter season, in particular during February and March. This is the time in which the polar vortex undergoes major dynamical changes in the final warming phase.

(iii) The relationship is present in all periods of the data, and shows remarkable consistency throughout the time series. In contrast, the relationship between the 1000 hPa AO and itself over extended range time-scales is extremely variable between different 10-year slices of the data  $-0.02 \leq \beta_0(\tau) \leq 0.15$ .

Including stratospheric information in a simple statistical forecasting model of the 1000 hPa AO provides an increase in SS of  $\sim 5\%$  over a statistical forecasting model which only includes 1000 hPa AO information. This increase is not due to the inclusion of an extra predictor in the model.

This analysis appears contradictory to the findings of Baldwin and Dunkerton (2001) and Thompson *et al.* (2002) that composites of large AO events in the stratosphere show a large change in the tropospheric AO some time after the event. We show that, while the relationship between the stratosphere and troposphere is real in a statistical sense, the quantitative size of the relationship is small. It is important to remember

that their analysis is based on a mean picture. Consideration of Fig. 3 shows that there is large spread around the regression line fitted to the dataset. This means the predictive skill of the relationship between the stratosphere and troposphere is small.

The procedures applied in this paper may not be the optimum method of obtaining predictive skill from the relationship between the lower stratosphere and troposphere. In particular, it may be possible to reduce the noise in the data by applying a smoothing filter. A simple, quantitative approach to the problem was adopted to avoid model-specific and filter-specific effects contaminating the results.

While the analysis presented here is of a purely statistical nature, it does nonetheless raise several questions about the dynamical nature of links between the stratosphere and troposphere. We are currently investigating these dynamical questions using a high-resolution general-circulation model. The results will be published in a companion paper.

#### ACKNOWLEDGEMENTS

Thanks to Brian Hoskins, Ellie Highwood and John Austin for useful discussions. AJC is funded by a Natural Environment Research Council Studentship and by a CASE award from the Met Office. AO and WL are part of the Data Assimilation Research Centre which is funded by the Natural Environment Research Council. We would also like to thank two anonymous reviewers whose comments on style and scientific content greatly improved this paper.

#### REFERENCES

- Ambaum, M. H. P. and Hoskins, B. J. 2002 The NAO troposphere–stratosphere connection. *J. Climate*, **15**, 1969–1978
- Baldwin, M. P. and Dunkerton, T. J. 1999 Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *J. Geophys. Res.*, **104**, 30937–30946
- 2001 Stratospheric harbingers of anomalous weather regimes. *Science*, **294**, 581–584
- Chambers, J. M., Cleveland, W. S., Kleiner, B. and Tukey, P. A. 1983 *Graphical methods for data analysis*. Wadsworth, Belmont, CA, USA
- Chatfield, C. 1996 *The analysis of time series*. Chapman and Hall, London
- Gibson, J. K., Kallberg, P. K., Uppala, S., Nourara, A., Hernandez, A. and Serrano, E. 1997 ‘ERA description’. ECMWF reanalysis project, Report No. 1. Available at <http://www.ecmwf.int>
- Junge, M. M. and Stephenson, D. B. 2002 Mediated and direct effects of the North Atlantic Ocean on winter temperatures in north-west Europe. *Int. J. Climatol.*, **23**, 245–261
- Lahoz, W. A. 2000 Northern hemisphere winter stratospheric variability in the Met Office Unified Model. *Q. J. R. Meteorol. Soc.*, **126**, 2605–2630
- Matsuno, T. 1971 A dynamical model of stratospheric warmings. *J. Atmos. Sci.*, **27**, 871–883
- Norton, W. A. 2003 Sensitivity of northern hemisphere surface climate to simulation of the stratospheric polar vortex. *Geophys. Res. Lett.*, **30**(12), 1627, doi:10.1029/2003GL016958
- O’Neill, A. 1995 ‘Observations of dynamical processes’. Pp. 59–82 in *The stratosphere and its role in the climate system*. Ed. G. Brasseur. Springer, Berlin, Germany
- 2003 ‘Stratospheric sudden warmings’. Pp. 1342–1353 in *Encyclopedia of atmospheric sciences*. Eds. J. R. Holton, J. A. Curry and J. A. Pyle. Academic Press, London
- Thompson, D. W. J. and Wallace, J. M. 1998 The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, **25**, 1297–1300
- Thompson, D. W. J., Baldwin, M. P. and Wallace, J. M. 2002 Stratospheric connection to northern hemisphere wintertime weather: Implications for prediction. *J. Climate*, **15**, 1421–1428
- Wilks, D. S. 1995 *Statistical methods in the atmospheric sciences*. Academic Press, San Diego, USA