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

By A.J.Charlton[1]*, A.O'Neill[2], D.B.Stephenson[1],W.A.Lahoz[2],M.P.Baldwin[3]

[1] *Department of Meteorology, University of Reading, UK*

[2] *Data Assimilation Research Centre,Department of Meteorology, University of Reading,UK*

[3] *Northwest Research Associates, Bellevue,Washington*

(Received ; revised)

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 1000hPa surface on a 10-60 day time scale is revealed. The relationship accounts for $\sim 5\%$ of the variance of the 1000hPa time series at its peak value and is significant at a the 5 % level .Over a similar time scale the 1000hPa time series accounts for $\leqslant 1$ % of itself and is not significant at a 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 $\sim 5\%$ 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.

Previously the work of Matsuno (1970) had considered the dynamics of the polar winter stratosphere to be intimately connected to planetary wave activity originating in the troposphere. Although some limited observational studies (Quiroz,1986) suggested a link between stratospheric sudden warmings and tropospheric blocking, a consistent connection between the two phenomena was difficult to determine.

Figure 1 shows a section from a daily amplitude AO time series (Baldwin,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

* Corresponding author: Department of Meteorology, University of Reading, PO Box 243, Reading, Berkshire RG6 2BB, UK

(50hPa-250hPa) much longer than in the middle stratosphere (10hPa-50hPa). 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.

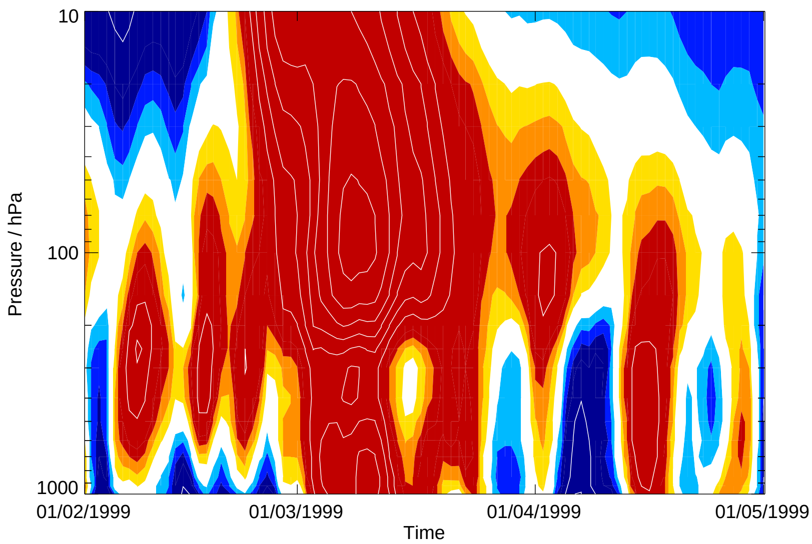

Figure 1. Time Height cross section of AO Amplitude data for February to April 1999. Red colours indicate negative values of AO Amplitude and blue values indicate positive values. Shading is as Baldwin (1999).

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 10hPa. The resulting composite AO events show the typical 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 10hPa zonal mean zonal wind Thompson et al. (2001) determine a large difference in surface temperature between 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.

Statistical modelling attempts to determine a relationship between two or more datasets. This fit to the data should have residuals with an unbiased, normal distribution which are not serially correlated. This definition does not take into account the size of the residual terms, which may be very large even if the model meets our definition of a good fit.

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.

We also test the model derived using our method against one which includes tropospheric information only. This allows us to measure the gain in skill which including stratospheric data in simple statistical forecasts provides.

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 70hPa and 1000hPa 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 summarised 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 Re-Analysis 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 data sets 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.

These extra diagnostics were extracted from ECMWF Re-Analysis (ERA-15, Gibson,1997) datasets held at the British Atmospheric Data Centre (BADC). Using a different re-analysis dataset to the one used to construct the AO dataset lends robustness to conclusions common to both datasets.

TABLE 1. Datasets

Name	Description	Levels	Time Range	Source
AO	Daily Amplitude of AO	17	1958-2000	Baldwin 1999
$\overline{u}$	Zonal Mean Zonal Wind at 60N	17	1979-2000	ERA-15
$\Phi\prime$	Geopotential Height RMS error from zonal mean at 60N	17	1979-2000	ERA-15
Filtered $\Phi\prime$	As above but using Geopotential Height dataset truncated at zonal wavenumber 2	17	1979-2000	ERA-15

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 standardised 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. (2001).

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:

$$y_z^{t+\tau} = \beta_0 y_z^t + \beta_1 x_z^t + \epsilon^t \tag{1}$$

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 and τ is the lag. The parameters β_0 and β_1 can be estimated using least squares regression and ϵ^t is a residual error. 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 β_0 and β_1 of the model using ordinary least squares. Repeating this for a range of values of the lag parameter τ produces a set of model parameter estimates as a function of lag.

Junge and Stephenson (2002) show that when using a multiple regression model with two predictors and standardised 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* (β_0) and an *indirect relationship* ($\rho(y_z^t, x_z^t)\ \beta_1$).

$$\rho(y_z^{t+lag}, y_z^t) = \beta_0 + \rho(y_z^t, x_z^t)\ \beta_1 \tag{2}$$

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 (β_0). The "path" from the stratosphere to the troposphere, taking into account the mutual correlation between the stratosphere and the troposphere, represents the indirect relationship ($\rho(y_z^t, x_z^t)\ \beta_1$).

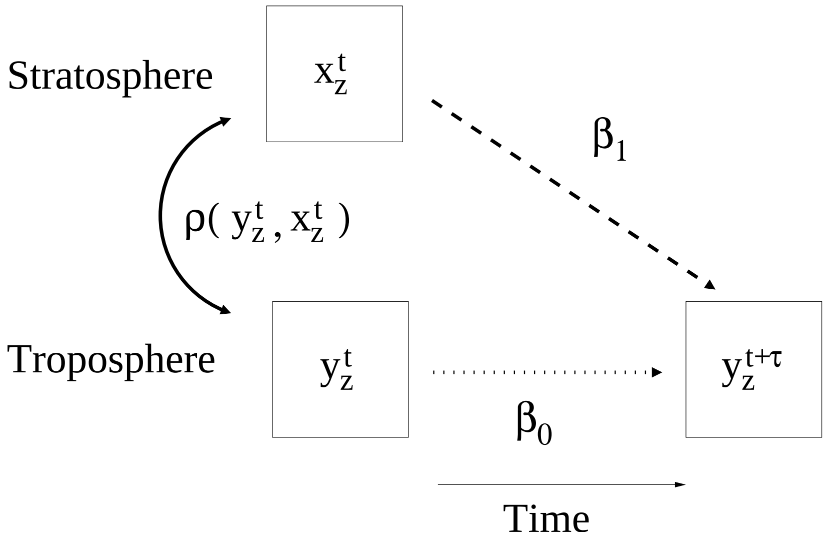

Figure 2. Idealised view of regression relationships. Squares represent state of atmosphere at some time. Dotted arrow indicates direct influence of troposphere on itself, dashed arrow indicates influence of stratosphere on troposphere, solid, curved arrow indicates instantaneous correlation between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

The parameters β_0 and β_1 represent correlations between datasets. While correlations give no information about causality, a statistically significant correlation between a value at some time t and a value at time t+τ 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 non-linearity of relationship between variables*

An assumption implicit in the thresholding techniques employed by both Baldwin and Dunkerton (1999) and Thompson et al. (2001) is that there is a different relationship between the stratosphere and troposphere when the amplitude of the AO in the stratosphere is large compared to when the amplitude of the AO in the stratosphere is small. This suggests that in order to examine the quantitative nature of this relationship a non-linear statistical model is required.

An inspection of the data shows that this is not the case. Figure 3 shows a scatter plot of the AO amplitude at 70hPa plotted against the AO amplitude at 1000hPa, for both 0 and 20 days lag. We choose 70hPa and 1000hPa to illustrate points which apply to the relationship between a number of time series in the lower and middle stratosphere and 1000hPa. Both plots show a general ellipsoidal shape. If the relationship between the variables were non-linear 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 further, more quantitative, check of the non-linear hypothesis is to perform a simple 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 the points at 70hPa which have magnitude greater than 1,2 and 3 non-dimensional AO amplitude. A linear fit to these sub-sampled data is also shown in dashed (threshold 1), dotted (threshold 2) and dot-dashed (threshold 3) lines in Fig. 3. If there were a non-linear effect present in the data it would be expected that

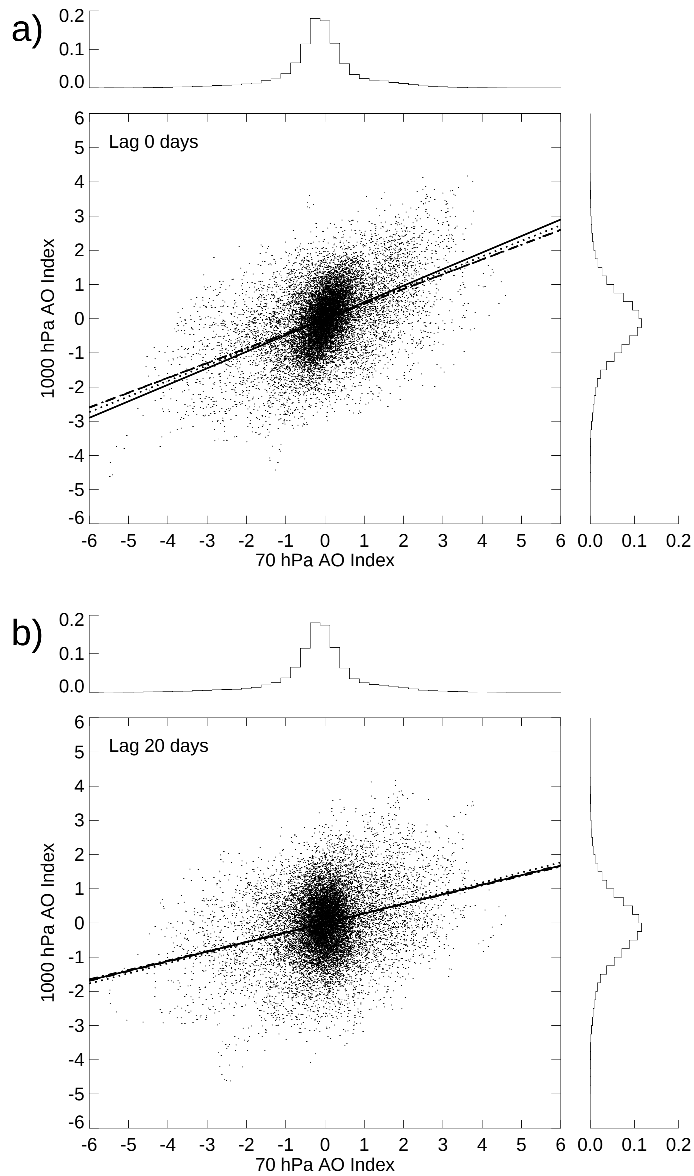

Figure 3. Scatter plot showing 70hPa AO Index plotted against 1000hPa AO index for a) lag=0 days and b) lag=20 days. Solid line shows fit with all of the data. Dashed line shows fit for data with a magnitude of 1 or greater. Dotted line shows fit for data with a magnitude of 2 or greater. Dot dashed line shows fit for data with a magnitude of 3 or greater. In each plot all fit lines lie approximately on top of one another.

the slope of the lines would change substantially when only data past a given threshold was included. In Fig. 3 the slope of the lines is almost identical for the fit with all data included and the fits with the sub-sampled datasets.

The thresholding analyses conducted by Baldwin and Dunkerton (2001) and Thompson et al. (2001) do not show the presence of a non-linear relationship between the AO on different levels. In their analysis they only consider data on the extremes of the ellipse in Fig. 3. The thresholding technique highlights the presence of a relationship between the stratospheric and tropospheric time series which is irrespective of the value of the stratospheric AO. This suggests that a linear statistical model is sufficient to investigate the relationship between the stratospheric and tropospheric parts of the AO.

(*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.

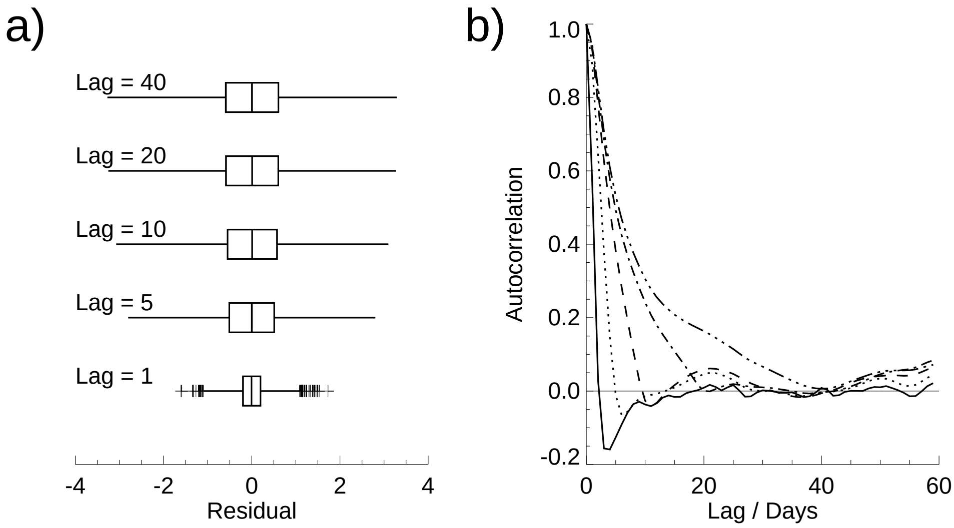

Figure 4. Figures showing distribution of residuals of model fit. Panel a) shows box plots of the distribution of residuals from the same model fit. Central line of box shows median residual, outer lines of box shows upper and lower quartile. Whiskers are plotted at 1.5 times the inter-quartile range.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

Figure 4 shows some diagnostics of the residuals for a fit of the model where the predictand series ($y_z^{t+\tau}$) is the 1000hPa AO time series and the predictor series (x_z^t) is the 70hPa AO time series. Figure 4 (a) shows box plots for a number of different model lags between one and forty 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 one and forty 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 (solid,dotted and dashed lines) this is the case; but for larger model lags (dot-dash and triple-dot dash lines) 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 overconfidence 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 10hPa. 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 data on the 70hPa pressure surface is referred to as the stratospheric data series and data on the 1000hPa pressure surface is referred to as the tropospheric data series. We chose 70hPa to illustrate points which are generalised to include a range of levels in the Upper Troposphere Lower Stratosphere (UTLS) region (which we define here as between 50hPa and 250hPa) in Section 5.

Statistical testing of many of the results is conducted. This testing uses a student 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 1000hPa were dependent only upon itself then β_1 would be equal to 0 and Eq. 1 could be re-written as:

$$y_z^{t+\tau} = \beta_0 y_z^t + \epsilon^t. \tag{3}$$

Equation 3 is an example of an autoregressive (AR(1)) or 'red noise' time series (Chatfield,1995). In Eq. 3 the autocorrelation would be equal to the value of β_0 at any lag.

The value of y at lag one day is equal to $y_z^{t+1} = \beta_0 y_z^t + \epsilon^t$

and so the value of y at lag two days is then given by $y_z^{t+2} = \beta_0 y_z^{t+1} + \epsilon^{t+1} = \beta_0(\beta_0 y_z^t + \epsilon^t) + \epsilon^{t+1}$.

In general,

$$y_z^{t+\tau} = (\beta_0)^\tau y_z^t + \sum_{l=0}^{\tau} (\beta_0)^{\tau-l} \epsilon^t \tag{4}$$

The maximum value of β_0 is one for stationary time series, therefore Eq. 4 shows exponential decay of β_0 with increasing lag. This means that autocorrelation of y would also decrease exponentially with lag. An exponential decay of the autocorrelation of a time series with lag suggests that it could be modelled as an autoregressive process.

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 of a dataset indicates that over these time scales the tropospheric AO time series could be modelled as an autoregressive process. The direct relationship (β_0,dotted line) is much larger than the indirect relationship ($\rho(y_z^t, x_z^t)$ β_1,dashed line). This suggests that only the direct relationship (β_0) is important on 1-10 day timescales.

On extended range (10-30 days lag) and slightly longer (30-60 days lag) time scales, the decay of the autocorrelation function (solid line) is less than exponential. The direct relationship (β_0) is much smaller than the autocorrelation and is not significant at the 5 % level. The indirect relationship($\rho(y_z^t, x_z^t)$ β_1) increases in magnitude and is significant at the 5 % level. On 10-60 day timescales the direct relationship accounts for $\leqslant 1$ % of the variance of the 1000hPa time series. In contrast the indirect relationship accounts for $\sim 5\%$ of the variance of the 1000hPa time series. Although both the direct relationship and the indirect relationship account for very small amounts of the variance of the 1000hPa 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 in the lower stratosphere and troposphere is seen on time scales of 10-60 days.

On much longer time scales (60-100 days lag) the autocorrelation of the tropospheric time series becomes smaller. The indirect relationship ($\rho(y_z^t, x_z^t)$ β_1) is much reduced and is not significant at the 5 % level. The direct relationship (β_0) accounts for most of the autocorrelation of the 1000hPa dataset.

These results suggest a small statistically significant relationship between the 70hPa AO and 1000hPa AO exists on 10-60 day time scales. The autocorrelation of the 1000hPa dataset on these time scales is accounted for mainly by the indirect relationship ($\rho(y_z^t, x_z^t)$ β_1).

Ambaum and Hoskins (2002) examined the autocorrelation of a smoothed 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 10-60 days is related to a connection between the lower stratosphere and the troposphere. This result also agrees with a model study by Norton (2002), who found a significant difference in the autocorrelation of the tropospheric AO on 10-25 day time scales when changes were made to stratospheric vortex.

(*b*) *Time Order Dependence*

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

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

The parameters of the model fit with the 70hPa AO time series and the time reversed 1000hPa AO time series are shown in Fig. 5(b). There is no evidence of a similar increase in the value of the indirect relationship ($\rho(y_z^t, x_z^t)\ \beta_1$) on 10-60 day time scales as is seen in Fig. 5(a). Therefore it can be inferred that the small, statistical relationship between the 70hPa and 1000hPa AO amplitude on 10-60 days is a product of the particular time orientation of the 1000hPa 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 (eg 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 (70hPa AO) series and JFM for the predictand (1000hPa AO) series. The model parameters are shown in Fig. 5(c) (DJF) and Fig. 5(d) (JJA) .

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 (($\rho(y_z^t, x_z^t)\ \beta_1$,dotted line) over the 10-60 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 70hPa and 1000hPa is in the winter season.In contrast no such structures are seen in the JJA fit. The stratospheric parameter 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.

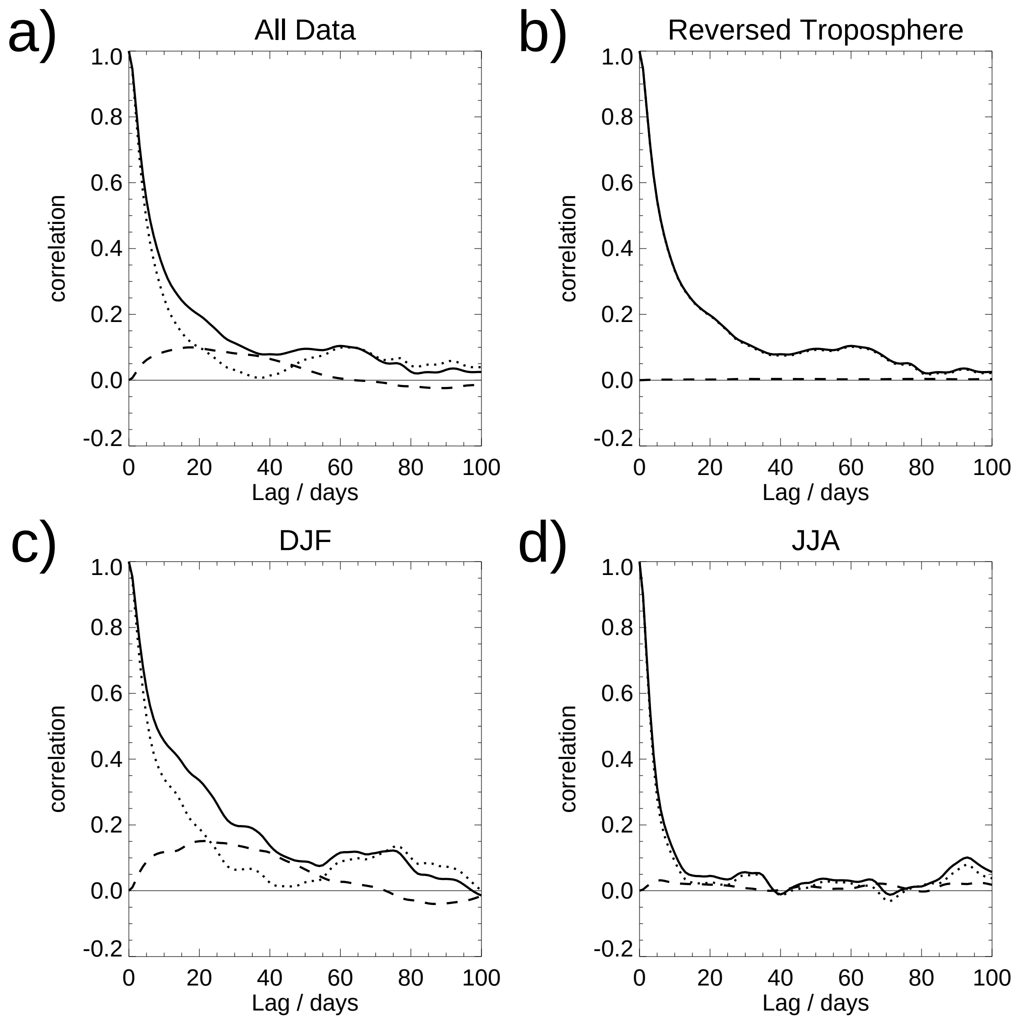

Figure 5. Decomposition of autocorrelation of 1000hPa AO using the model in Eq. 1, where 70hPa is the x_z series and 1000hPa the y_z series. Autocorrelation of 1000hPa series $r(y_{1000}(t+\tau), y_{1000}(t))$ shown in solid line, β_0 shown in dotted line and the product $\beta_1 r(y_{1000}(t), x_{70}(t))$ shown in dashed. Panel (a) shows results using all of the data, panel (b) shows results when the y_{1000} time series is reversed in time, panel (c) shows results for DJF data only and panel (d) shows results for JJA data only.

(*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. Shading in Fig. 6 (b) and (c) shows significance at the 10 % (light shading) and 5 % levels. It is important to remember that although the plots are shown with contours they represent 12 independent sets of 100 model

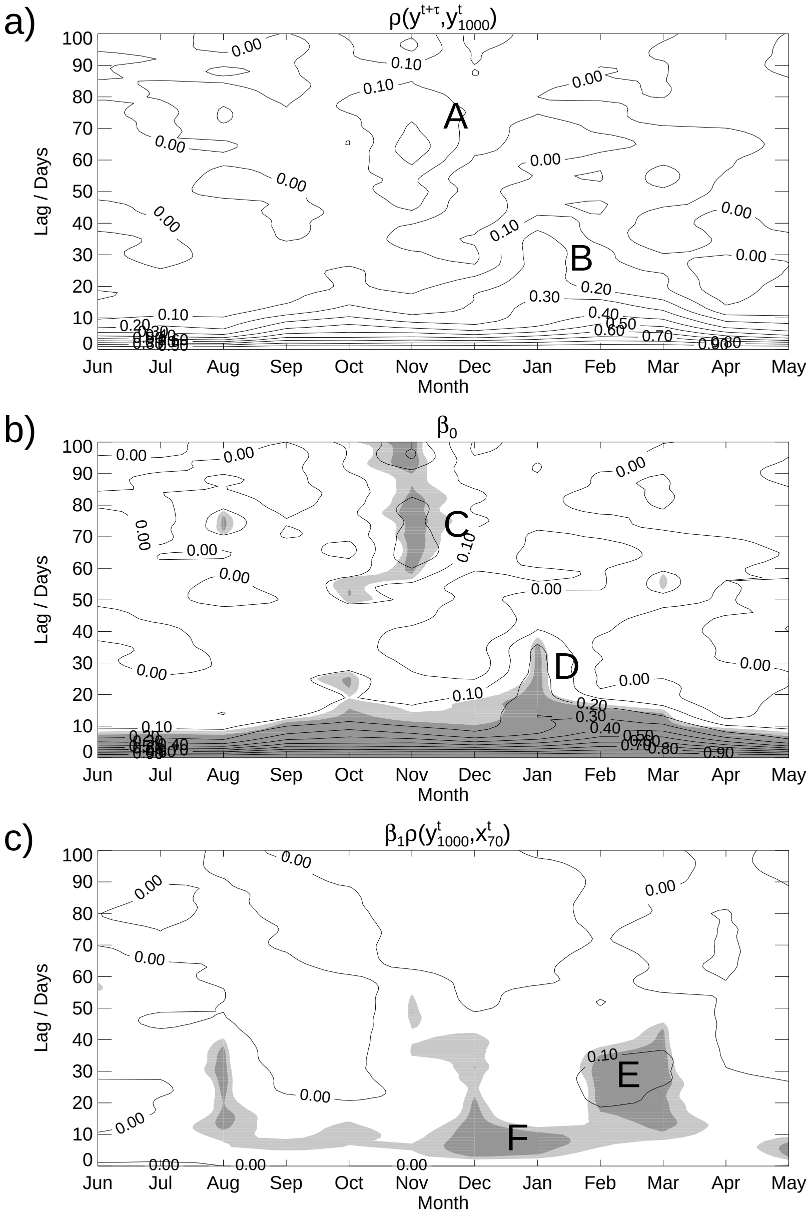

Figure 6. Contour plots of decomposition of 1000hPa AO autocorrelation of 1000ha. a) shows autocorrelation of 1000hPa AO as a function of month and lag. b) shows direct effect (β_0) as a function of month and lag. c) shows indirect effect ($\rho(y^t_{1000}, x^t_{70})\ \beta_1$) as a function of month 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 5 % level of 0.10. A-F mark salient features see text for details.

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 1000hPa 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 (B).The increase in the autocorrelation of the 1000hPa AO in January (B) is attributable to the increase in the direct relationship (β_0) 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 is due to an increase in the indirect relationship ($\rho(y^t_{1000}, x^t_{70})\ \beta_1$) E.

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 breakup 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 70hPa AO and the 1000hPa AO during December and January (F) but the magnitude of the correlation is much 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 sub-sets of the AO datasets shows that the relationship between 70hPa and 1000hPa identified in section 4(a) is confined to February and March. This might suggest that the relationship between 70hPa and 1000hPa might be linked to the timing of the final warming of the stratospheric vortex.

(*e*) *Relationship in simple diagnostics*

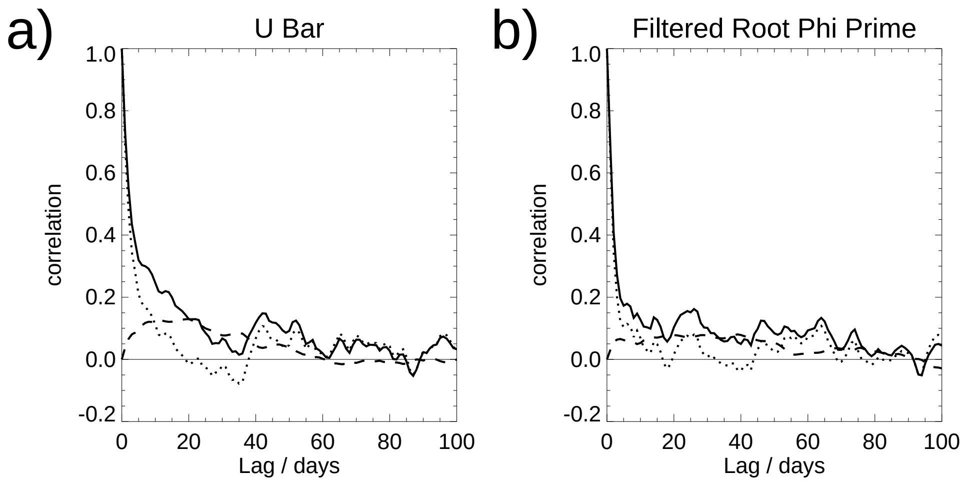

Figure 7. Decomposition of autocorrelation of 1000hPa dataset as figure 5 but for (a) $\overline{u}$ time series and (b) filtered $\Phi\prime$ time series (right column).

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 70hPa and 1000hPa established in section 4(a) using the AO dataset we repeat the analysis using three other simple zonal mean diagnostics. A relationship between 70hPa and 1000hPa 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 Φ' dataset. This quantity is defined as follows.

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

Where Φ represents geopotential height and the overbar represents a zonal mean. Before calculating this diagnostic we filter the geopotential height analysis to only include zonal wavenumbers up to and including zonal wavenumber two. The stratosphere exhibits primarily low wavenumber variability and it is reasonable to expect that any relationship between the stratosphere and troposphere is likely to occur through these wavenumbers. The Φ' diagnostic would include higher zonal wavenumber variability in the troposphere which may confuse any relationship between the stratosphere and troposphere. A fit is also made with a Φ' 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 Φ' datasets have qualitatively similar correlation series to the AO data set. The indirect relationship ($\rho(y^t_{1000}, x^t_{70})\ \beta_1$) has larger values over the 10-60 day lag as in the AO dataset. Over a similar time scale there is also a reduction in the size of the direct relationship (β_0) as in the AO data set. 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 Φ' datasets. The model parameters in the unfiltered Φ' dataset (not shown) do not show similar correlation as 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 Φ' 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.

5. Extending the Model to Other Levels

Fitting the model with 70hPa 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^t_z) 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

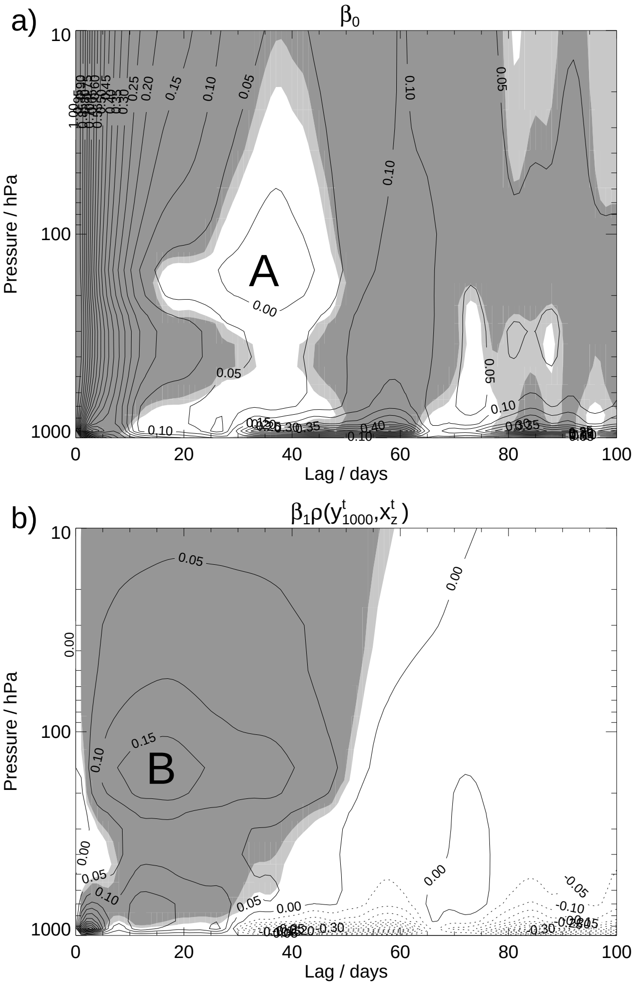

Figure 8. Model parameters (solid contours) for various predictor levels. Panel a) shows β_0 for fits with various levels, panel b) shows product of $\beta 1$ and instantaneous correlation for various predictor levels. Regions where the parameters are significantly different from zero at the 5% level 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.

are plotted on the panels at the corresponding pressure. For example, a cut across Fig. 8 (a) at 70hPa would produce the dotted line in Fig. 5 (a) and a a cut across Fig. 8 (b) at 70hPa would produce the dashed line in Fig.5 (a).

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

The large increase in the indirect relationship is accompanied by a similar decrease in the direct relationship (β_0) (Fig. 8 (a)). This reduction is largest in the same region between 50hPa and 250hPa (A), but there is a general reduction in the significance for levels into the middle stratosphere. The indirect relationship has largest magnitude on the 150hPa surface. 150hPa is in the troposphere at most latitudes. It is therefore suggested that while some predictability of the 1000hPa 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.

Fitting the statistical model at all levels in the dataset shows a similar relationship to that revealed at 70hPa exists for a number of levels in the upper-troposphere lower-stratosphere (UTLS) region (50hPa-250hPa). A similar relationship between the middle stratosphere and the 1000hPa AO exists but the magnitude of the correlation here is much smaller than in the UTLS region. 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 (10-60 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 was split into a series of ten year blocks and the model fitting procedure applied to each block. The model fit is made for the 1000hPa and 70hPa 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 found that the results were robust within the region of increased indirect relationship ($\rho(y_z^t, x_z^t)\ \beta_1$) (10-60 days).

Figure 9 shows the autocorrelation ($\rho(y_{1000}^{t+\tau}, y_{1000}^t)$), direct relationship ($\beta_0$) and indirect relationship ($\rho(y_z^t, x_z^t)\ \beta_1$) 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 70hPa and 1000hPa is relatively similar between decades with significantly different variability in the stratosphere. In particular the 1990 had relatively few stratospheric sudden warmings but the relationship is still statistically significant.

In contrast the magnitude of the direct relationship (β_0) (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.

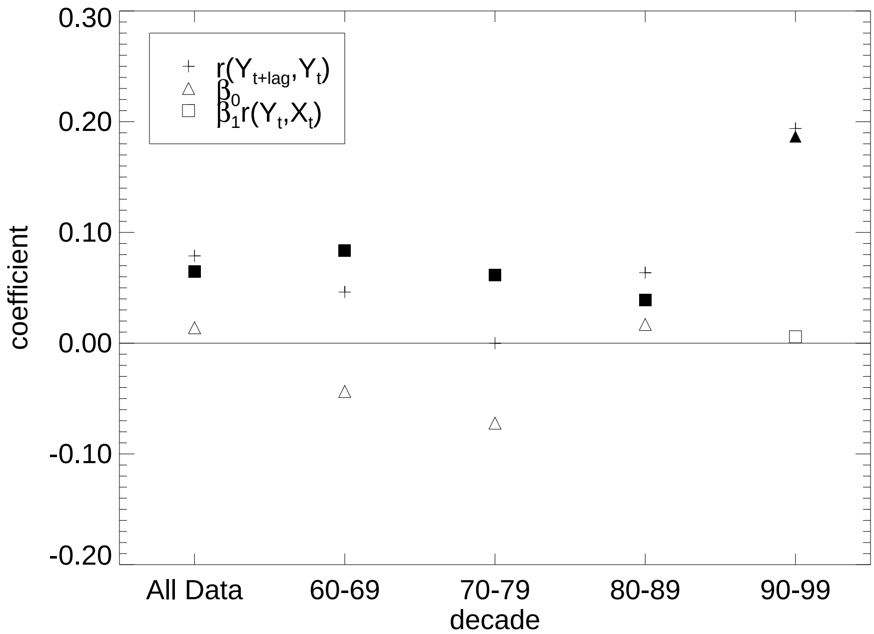

Figure 9. Model parameters (solid contours) for various subsections of data at lag of 30 days. The autocorrelation of the dataset is plotted with crosses, β_0 is plotted with triangles and $\beta_1 r(Y_t, X_t)$ is plotted with squares. Symbols for model parameters are filled in black if the parameter is significantly different from zero at the 5 % level.

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

$$y_z^{t+\tau} = \beta_0 y_z^t + \epsilon^t \tag{6}$$

This model only takes into account the previous value of the 1000hPa AO and is not expected to be a good model of the future state of the 1000hPa AO.

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

$$SS = 1.0 - \frac{MSE_{forecast}}{MSE_{climatology}} \tag{7}$$

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

Figure 10 shows the difference in skill score between the 1000hPa only control regression model and the two predictor model as in Eq. 1. Positive values indicate including data at a particular pressure level and lag adds skill to forecasts of the 1000hPa AO (compared with a 1000hPa AO only model) and negative values indicate including data at a particular lag and pressure level reduces skill to forecasts of the 1000hPa AO (compared with a 1000hPa AO only model).

The skill is plotted for different lags and different pressure levels. The two panels show results when different halves of the data set are used to train the model.

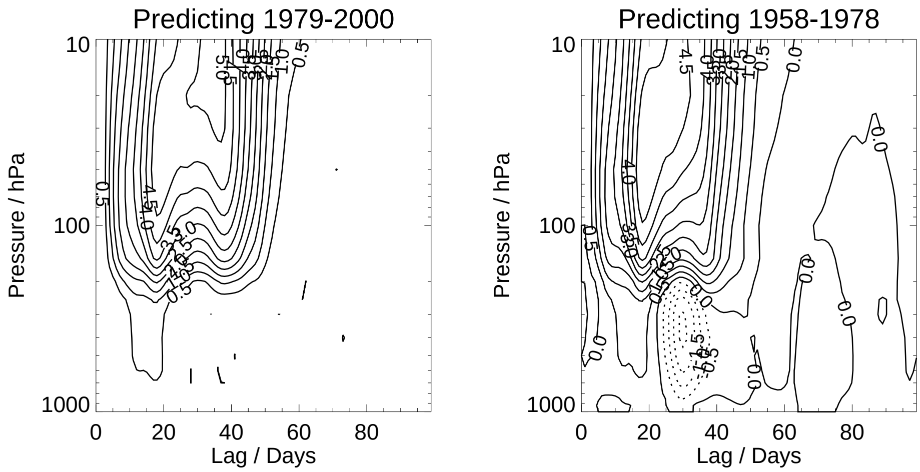

Figure 10. Difference in Skill Score of multiple predictor model and 1000hPa predictor only model versus climatology. The left panel shows the difference when data series from 1958-1978 are used to fit the model which is tested against data series from 1979-2000. Right panel shows the difference when data series from 1979-2000 are used to fit the model which is tested against data series from 1958-1978. Contour interval is 1%. Solid contours show the multiple predictor model has larger Skill Score than the 1000hPa predictor only model. Dotted contours show the multiple predictor model has smaller Skill Score than the 1000hPa predictor only model. See text for details of models.

Figure 10 shows the SS of the two predictor model is greater than the 1000hPa only control model in the lower and middle stratosphere (250hPa - 10hPa) on timescales between 10 and 60 days. This is the region highlighted in the model fit as the significant region for the indirect relationship $(\rho(y_z^t, x_z^t)\ \beta_1)$.The magnitude of the increase is small $\sim 5\%$.

In contrast for levels in the middle and lower troposphere the SS of the two predictor model and the 1000hPa only control model is approximately comparable. The addition of extra information from the middle and lower troposphere into a statistical model of the 1000hPa AO does not provide any benefit.

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 1000hPa time series reversed in time. In this case the 1000hPa only control model will have identical skill and the additional predictor dataset will have 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.

Tests with a time-reversed 1000hPa AO series showed close to zero difference in skill between the control 1000hPa only model and the two predictor model at all levels and all lags. The gain in skill shown in Fig. 10 is a product of real improvements to the forecast and not due to addition of another predictor into the 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 1000hPa AO. Examination of AO time series such as Fig. 1 suggests that the middle and lower tropospheric AO often has 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.

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 focussed on qualitative analysis of such a relationship. In particular, Baldwin and Dunkerton (2001) and Thompson et al. (2001) 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 1000hPa has been identified. Typical correlations between the lower stratosphere and 1000hPa are small (~ 0.2), but significant (at the 5 % level) over extended range time scales (10-60 days).

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

- The relationship is most prominent in the upper troposphere lower stratosphere region (50-250hPa). This region spans different parts of the atmosphere at different latitudes, but can broadly be thought of as the location of the tropopause.
- 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.
- The relationship is present in all periods of the data, and shows remarkable consistency throughout the data series. In contrast the relationship between the 1000hPa AO and itself over extended range time scales is extremely variable between different 10 year slices of the data ($-0.02 \leqslant \beta_0 \leqslant 0.15$).

Including stratospheric information in a simple statistical forecasting model of the 1000hPa AO provides an increase in Skill Score of $\sim 5\%$ over a statistical forecasting model which includes information only includes 1000hPa 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 (2001) and Thompson et al. (2001) that composites of large AO events in the stratosphere show a large change in the tropospheric AO some time after the event. However it is important to remember that their analysis is based on a mean picture. Consideration of Figure 3 shows 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 running mean 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 NERC Studentship and by a CASE award from the UK Met Office. AO and WL are part of the Data Assimilation Research Centre which is funded by NERC.

References

Ambaum, M. H. P. and Hoskins, B. J. 2002 The NAO troposphere-stratosphere connection. *J. Climate*, **15,** 1969–1978

Andrews, D. G. and Holton, J. R. and Leovy, C. B. 1987 *Middle Atmosphere Dynamics*, Academic Press, San Diego **489pp**

Baldwin, M. P. and Dunkerton, T. J. 1999 Propagation of the Arctic Oscillation from the Stratosphere to the Troposphere. *J. Geophys. Res.*, **104,** 30,937–30,946

Baldwin, M. P. and Dunkerton, T. J. 2001 Stratospheric Harbingers of Anomalous Weather Regimes. *Science* **294,** 581–584

Chatfield, C. 1996 *The analysis of Time Series*Champman and Hall, London, UK, **304pp**

Gibson, J. K., Kallberg, P. K., Uppala, S., Noumara, A., Hernandez, A. and Serrano, E. 1997 1997:ERA Description. *ECMWF Re-Analysis Project Report Series,* ECMWF,Reading,UK, **77pp**

Junge, M. M. and Stephenson, D. B. 2002 Quantifying the effects of North Atlantic sea surface temperatures on Central England winter temperature variations *J. Geophys. Res.*, **in press,** –

Matsuno, T. 1970 A dynamical model of stratospheric warmings *J. Atmos. Sci.*, **27,** 871–883

Norton, W. A. 2002 Sensitivity of Northern Hemisphere surface climate to simulation of the stratospheric polar vortex *J. Climate*, **submitted,** –

O'Neill, A. 1995 Observations of Dynamical Processes *The Stratosphere and Its Role in the Climate System. Ed. G. Brasseur,* Springer-Verlag, Berlin, Germany, **368pp**

Quiroz, R. S. 1986 The association of stratospheric warmings with blocking. *J. Geophys. Res.*, **91,** 5277–5285

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

Wilks, D. S. 1995 *Statistical Methods in the Atmospheric Sciences. Academic Press,San Diego,USA* **464pp**