Volume variability diagnostic for 4D datasets

Natalia Andronova[1] and Steven Boland[1]

Received 18 February 2011; revised 23 March 2011; accepted 30 March 2011; published 19 May 2011.

[1] Recent advances in the compilation of the climate science databases, derived from observations and model simulations, help in estimating trends and variability of current and past atmospheric conditions. However, an exponential increase in the amount of information limits its comprehension. In this paper we present a simple diagnostic, named as "volume variability", that allows for analysis of multidimensional fields of climate variables. The simplicity inherent in the new diagnostic, calculated as a product of variances along a set of directions, and its transparent interpretation thereof will enable its use as a metric for the quick and easy comparison of the 3D time series of models and observations. Using this diagnostic, we explore the evolution of the temperature and geopotential height fields over the Southern Hemisphere Cap region by way of four reanalysis systems. **Citation:** Andronova, N., and S. Boland (2011), Volume variability diagnostic for 4D datasets, *Geophys. Res. Lett.*, *38*, L10805, doi:10.1029/2011GL047193.

1. Introduction

[2] The detection of climate change has become one of the most prominent scientific topics, and estimating climate variability and trends has become increasingly important for defining potential socioeconomic development paths. Thus, the key advances have been made in the compilation of climate databases from observations and model simulations. Concurrently, over the course of the last decade, the volume of climate data has been increasing exponentially due to the increasing number of models and reanalysis systems and their respective grid resolutions. Therefore, there is still a need for a methodology that will allow for the simple assessment of large datasets and their differences.

[3] In this paper we present a new diagnostic for use with large, multidimensional datasets. We advocate simplicity and transparency of the new diagnostic. The methodology is presented in Section 2. We test the diagnostic using the 4D fields from four reanalysis systems: MERRA, NCEP, ERA40 and JRA25, which are presented in Section 3. In Sections 4 and 5 we present the results of the data analyses, based on temperature and geopotential fields over SH cap region; and in Section 6 we presents the conclusions.

2. Methodology

[4] The new diagnostic introduced in this paper is based on a calculation of variances of a quantity Q along a set of directions. Because in this paper we used three geometric directions aligned with (x, y, z), we named the diagnostic as "volume variability," VVar. Namely, for a 4D quantity Q(x,y,z,t) the VVar is defined as:

$$\mathrm{VVar} = \mathrm{VarX} \cdot \mathrm{VarY} \cdot \mathrm{VarZ} \tag{1}$$

where VarX, VarY, VarZ are time series of the Q's respective variances, calculated along one of the directions while keeping the quantity averaged over other two complimentary directions. For example, VarX would mean the variance along x-direction of the variable Q, beforehand averaged over y- and z-directions:

$$VarX = \frac{1}{n}\sum\left(\overline{Q}_i - \mu\right)^2;\ \mu = \frac{1}{n}\sum\overline{Q}_i;\ \overline{Q}_i = \frac{1}{mp}\sum_{j=1,m}\sum_{k=1,p} Q_{jk} \tag{2}$$

where n, m and p are the number of points in x, y and z directions. Each of three components of VVar represents a 1D time series, and their product makes an analogy of evolution of the spatial volume. Taking the natural logarithm to accommodate a few orders of VVar converts the metric into a non-dimensional index. This index can be generalized on more than three dimensions.

[5] The VVar diagnostic sheds new light upon the underlying physical processes with regards to a particular quantity through the incorporation of the threefold product of the one-dimensional variances, each calculated with regards to their respective directions. Variance measures the spread of data about its mean. An increase in variance is a consequence of the existence of a gradient or a trend in a given quantity or of an increase in the amplitude of a quantity's variability. Conversely, a decrease in variance is the result of an absence of significant changes in a quantity's magnitude along a given direction. The product of three one-dimensional variances comprises the volume variability (VVar) diagnostic for a particular quantity. A large VVar value indicates that a given quantity's capacity for the temporal volumetric fluctuation is correspondingly sizable, and vice-versa; the VVar diagnostic incorporates the three directional variances simultaneously, lending it a statistical value for the verification and validation of numerical models. Among other advantages of the VVar are its simplicity and ability to handle large arrays of data.

[6] We envision this diagnostic as the first step in tackling sets of large 4D arrays (for example, the model simulations for the forthcoming AR5 report). The process is twofold: we search, first, for similar tendencies in the VVar's amplitude and phase between datasets, and, second, for similar tendencies in its respective single-direction components. There is, therefore, no predefined threshold for VVar; rather, the degree of similarity of datasets is determined by the correlation of the VVar (and its components') time series.

[1]Atmospheric, Oceanic and Space Science, University of Michigan, Ann Arbor, Michigan, USA.

0094-8276/11/2011GL047193

Table 1. Assimilating Systems, Used in the Paper

Assimilated System	Temporal Range	Horizontal Resolution	Vertical Resolution
MERRA	Jan 1979 to Jan 2007; 29 years, 348 months	X288 × Y144	L42 1000 hPa to 0.1 hPa
JRA25	Jan 1979 to Jan 2007; 29 years, 348 months	X288 × Y145	L23 1000 hPa to 0.4 hPa
ERA40	Dec 1957 to Dec 2002; 45 years, 540 months	X145 × Y73	L23 1000 hPa to 1 hPa
NCEP	Jan 1948 to Apr 2010; 62 years, 748 months	X144 × Y73	L17 1000 hPa to 10 hPa

[7] In the next section we demonstrate the application of the diagnostic to four reanalysis systems.

3. Data

[8] During the last ten years numerous efforts have been made in developing consistent atmospheric reanalysis products (so called the assimilated observations) that provide estimates of variables from the surface to the top of the atmosphere. The most widely used reanalysis are those developed by NCEP (USA Climate Prediction Center [*Kalnay et al.*, 1996], often referred to as NNR) and by the European Centre for Medium-Range Weather Forecasts, ECMWF (often referred to as ERA-40 [*Kistler et al.*, 2001; *Uppala et al.*, 2005]), which extend back to 1948 and 1958 respectively. Other reanalysis data are JRA-25 (Japan [*Onogi et al.*, 2007]) and MERRA (USA, NASA [*Bosilovich et al.*, 2008; M. M. Rienecker et al., MERRA—NASA's Modern-Era Retrospective Analysis for Research and Applications, submitted to *Journal of Climate*, 2011]), both of which cover a much shorter period (1979-present). The MERRA new reanalysis is a data assimilation system that grew out of the NASA Goddard Earth Observing System of the Global Modeling and Assimilation Office. Brief characteristics of MERRA, JRA25, ERA40 and NCEP are presented in Table 1.

[9] We should mention that the quality of the reanalysis systems is significantly reduced when the amount of input observations assimilated by the climate model is reduced. This is true, for instance, for the SH polar cap region, where there are still not enough comprehensive atmospheric observations, made especially during polar nights. We have chosen this region for analysis. This choice allowed us to focus on comparing the reanalysis systems to one another sans passing judgment on correctness. We test our methodology using the month of July, which represents the middle of the polar night and is far from the seasonal transition times, and two variables: the temperature (in Section 4) and geopotential height (GPH) (in Section 5).

[10] Table 1 shows that four reanalysis systems are different in their spatial resolutions. Among the four considered models, MERRA had the highest top, with 42 levels spanning over the surface to 0.1 hPa. In this paper we used the version of MERRA that has the resolution 1.25 by 1.25 degrees, but which does not assimilate some land points up to approximately 500 hPa. The JRA25 and ERA40 models both have 23 vertical levels, but with the tops at 0.4 hPa and 1 hPa respectively. The horizontal resolution of the JRA25 is comparable with MERRA, however the horizontal resolution of ERA40 is coarser by a factor of two than that of the MERRA (and JRA25) resolution in both x- and y-directions. The horizontal resolution of the NCEP model is comparable to ERA40, but it has the lowest top located at 10 hPa. The length of the time series, represented by these four re-analysis systems, is also different: 29 years for MERRA and JRA25; 45 years for ERA40; and 61 years for NCEP.

[11] We carry out the analysis using monthly averages; for different overlapping 4D domains; with all grid averages being made area weighted; and, when needed, with the data re-gridding based on the lowest model's grid.

4. Results

[12] First, for each modeling system, based on the natural logarithm of variances along the overlapping vertical domain (500 to 10 hPa) of the temperature, averaged over the overlapping time domain (1979–2002), we have constructed the July temperature temporal variability maps. We use these maps to define the latitudinal boundaries of the region where we carry out the analysis. We should mention, that in the general case a choice of the domain depends on a particular task and the researchers' expertise. For a particular case of the SH cap, it appears (not shown) that the contour of ln(VVar)=6.4 fits the threshold value the best as for each dataset the results almost perfectly align with the latitude 40°S. Therefore, the first domain for the volume variability calculation will be: 40S–90S in latitudes; 180E–180W in longitudes and 10–500 hPa in vertical direction.

[13] Figure 1 shows the July ln(VVar) for each dataset as a function of time. From Figure 1a it follows that during 1979–2002 all reanalysis systems are generally in agreement with each other, especially before 1990. After 1990, the ERA40 starts to exhibit a slightly different behavior than other reanalysis data. The MERRA is very similar to the NCEP and JRA25 evolution, but exhibits a higher, closer to ERA40, reaching values of ln(VVar) between 11.5 and 12. Before 1992, the maxima of ln(VVar) for all four datasets have their appearance close to ten-year frequency period. Also, the known large El-Nino events and the Pinatubo eruption event can be acknowledged. Figure 1b shows the long-term evolution of ln(VVar) from ERA40 and NCEP, from where it can be seen that ERA40 exhibits a much higher variability level for the SH cap atmosphere than does NCEP. Plus, ERA40 shows the existence of large-scale multi-decadal variability, while NCEP does not. *Bromwich and Fogt* [2004] mentioned that both datasets, NCEP and ERA40, have their shortcomings in high southern latitudes before 1970, the beginning of the satellite era, due to small number of observations. However, Figure 1b shows that both the NCEP and the ERA40 (at least in terms of the level of the temperature volume variability index) are consistent in the representation of the VVar for the whole period, during which the ERA40 exhibits a much larger amplitude of the ln(VVar) variability than the NCEP model. We determined (not shown), that between these two models the main difference in the VVar is due to the vertical component of the VVar, which magnitude is much larger for ERA40. We compare the separate components of the VVar later in this section.

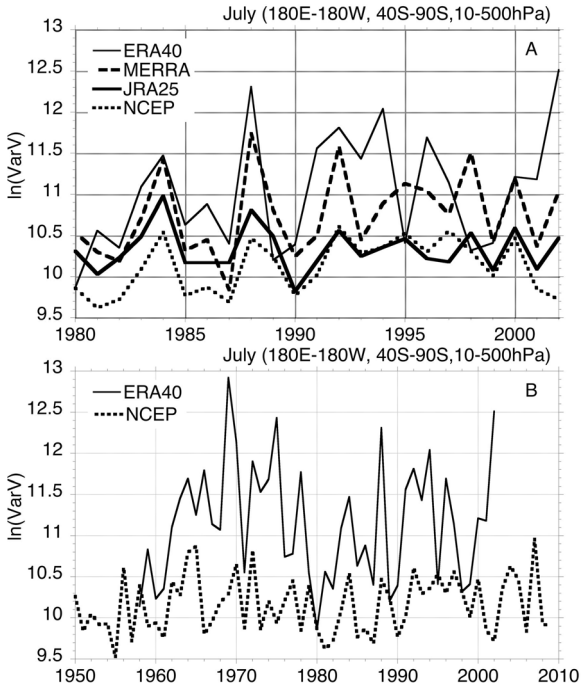

Figure 1. The July variability volume (VVar) over (180E–180W, 40S–90S, 10–500 hPa) domain as a function of time: (a) ERA40, MERRA, JRA25, NCEP data over 1980–2002, (b) ERA40 and NCEP over 1950–2009.

[14] Figure 2 shows the behavior of the July ln(VVar) calculated for two parts of the atmosphere: below and above 10 hPa. In Figure 2 we used only three reanalysis systems: ERA40, MERRA and JRA25, as the NCEP model has the top located at 10hPa. As expected, all three systems show a larger VVar in the upper atmosphere than in the lower atmosphere. Because of its top at 1 hPa (lower than two other), ERA40 shows the highest correlation (0.82) between its two atmospheric parts. The MERRA has a slightly higher connection between its lower and upper parts (correlation is 0.68) than JRA25 (correlation is 0.63) despite MERRA's top at 0.1 hPa being higher than the JRA25's 0.4 hPa top. Also, MERRA exhibits a small upward trend in the variability (statistically insignificant) of the upper atmosphere, while JRA25 does not.

[15] Based on the interpretation of VVar given in the methodology section, a small magnitude of the temperature-based VVar would mean that along a particular atmospheric direction the atmosphere is close to being homogeneous, which might be a demonstration of a well-mixed atmosphere in this direction(s) (e.g., due to small scale convection) or a result of a forced atmospheric state (e.g., due to large scale transport). A large magnitude of VVar would mean that along particular directions the air masses are heterogeneous and/or a gradient exists along these directions. Therefore, looking at the components of VVar would uncover the directions that are responsible for the resulting effect in VVar.

[16] Figure 3 lists different VVar components above 10 hPa, where the greatest differences occur. As before, we will notate the latitudinal component of VarV as VarY, the longitudinal component as VarX and the vertical components as VarZ. Figure 3a shows that all three reanalysis systems have very similar VarX, especially MERRA and JRA25. On the other hand, Figure 3b shows that the VarY is very similar between MERRA and ERA40, while JRA25 shows a large variability during the so-called "Pinatubo period" from 1992 to 1996. Figure 3c shows that the greatest differences exist between the models for VarZ: MERRA displays the highest variability level, followed by ERA40, then by JRA25. The JAR25's VarZ, which is the lowest one, would mean that JRA25's upper atmosphere is better vertically coupled than the other systems. At the same time, JRA25 shows the highest VarY, which would mean that in the JRA25 upper atmosphere there is less mixing in the latitudinal direction (more isolation) than in the other systems. The findings related to behavior of both the VarY and VarZ in JRA25 (the small variance in the vertical direction and the large one in the latitudinal direction) are consistent with a more isolated and colder stratospheric polar vortex than in the other systems. Indeed, the range of the temperatures within 40S–90S averaged over vertical and temporal domains is (185K–220K), while the corresponding range showed by the ERA40 is (210K–235K). In MERRA, the highest magnitude of VarZ and the smallest magnitude of VarY would mean that in its upper atmosphere mixing in the latitudinal directions dominates. In the MERRA the range of the temperatures within 40S–90S averaged over vertical and temporal domains is (220K–235K).

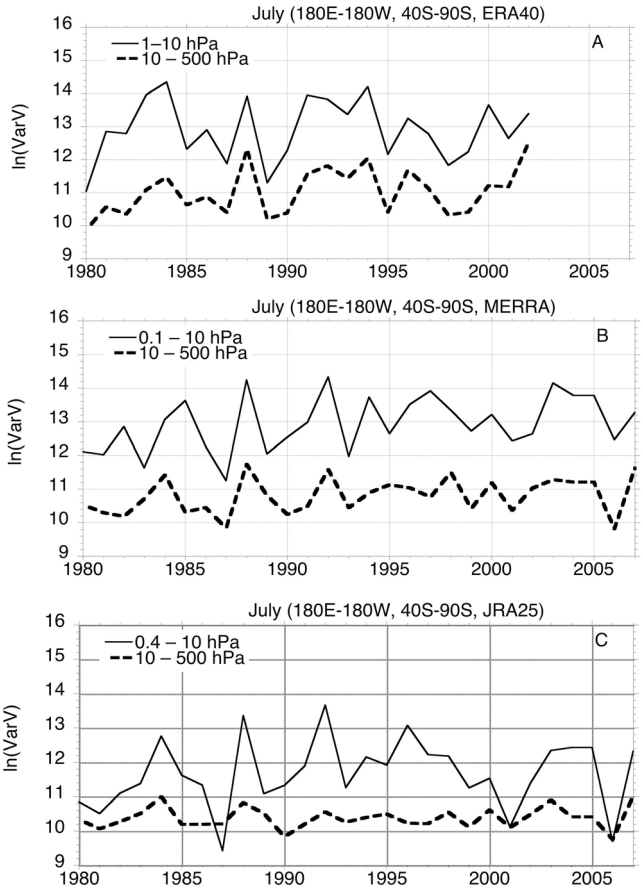

Figure 2. The July VVar calculated for different parts of the southern hemisphere cap atmosphere: (a) ERA40, (b) MERRA, and (c) JRA25.

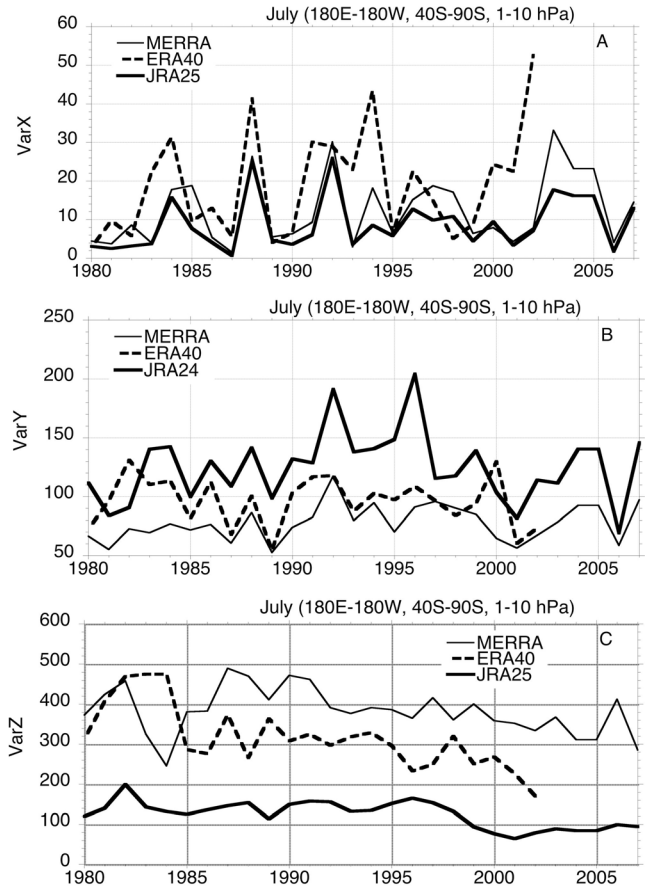

Figure 3. Different components of the July VVar in the upper atmosphere: (a) VarX, (b) VarY, and (c) VarZ.

[17] In the next section, based on the analysis of GPH field, we show another possible interpretation and application for the components of the VVar index.

5. Comparison With the Southern Annular Mode (SAM) Index

[18] A generalized index of the Southern Hemisphere annular mode (SAM) is characterized by synchronous fluctuations in pressure of one sign over the polar caps and of the opposite sign at lower latitudes. High values of the index correspond to an anomalously strong polar vortex, low values to a weak one. As by *Thompson and Wallace* [1998] and *Baldwin et al.* [1994] the SAM index was obtained by calculating the leading modes of the empirical orthogonal function (EOF) of the monthly mean geopotential height anomalies at 20°–90°S from the NCEP reanalysis system. Therefore, according to the definition of our VVar metric, we expect that the SAM index would be comparable with the VVar latitudinal component, VarY. Figure 4 shows a comparison of the July SAM index, averaged over 10–500 hPa, with the anomaly of ln(VarY), calculated from the NCEP GPH field for the domain (180E–180W, 20S–90S, 10–500 hPa). As expected, Figure 4a demonstrates high agreement (correlation of 0.92) between these two metrics. Therefore, each component of the VVar can be taken as a "leading mode" of atmospheric variability in a particular direction.

[19] Figure 4b shows the GPH based anomaly of ln(VarY), calculated from other three reanalysis systems, MERRA, ERA40 and JRA25. It can be seen that MERRA and JRA25 have very similar temporal evolution to SAM (correlation 0.82 and 0.84, correspondingly), while ERA40 is not (correlation is less than 0.2). In addition, the absolute values of the ERA40 (not shown) of the GPH based ln(VarY) is considerably larger (~18) than the VarY calculated from MERRA and JRA25 (they are at the level of 13.4 and 12.8, respectively). This indicates that in the domain (180E–180W, 20S–90S, 10–500 hPa) the ERA40 GPH field (averaged over vertical direction and time, 1979–2002) has the largest gradient along the latitudinal direction, which is also evident (not shown) from the geographical distribution of the GPH directly.

6. Conclusions

[20] In this paper we have presented and applied a simple methodology to explore the evolution of the temperature and geopotential height (GPH) variability patterns over the Southern Hemisphere cap region in terms of the variances of variables using four reanalysis systems (ERA40, NCEP, MERRA, and JRA25). The SH polar cap region still lacks comprehensive atmospheric observations, especially during polar nights, and therefore, the reanalysis systems serve as an important source of information, in particular for numerical model verification. We calculated the temporal evolution of a quantity, which we named atmospheric "volume variability", VVar, and tracked it based on the July atmospheric temperature and 3D geopotential height distributions.

[21] The purpose of the VVar diagnostic is not only to find a correlation between the capacity for fluctuation (expansion-contraction tendencies) of the given quantities' VVar,

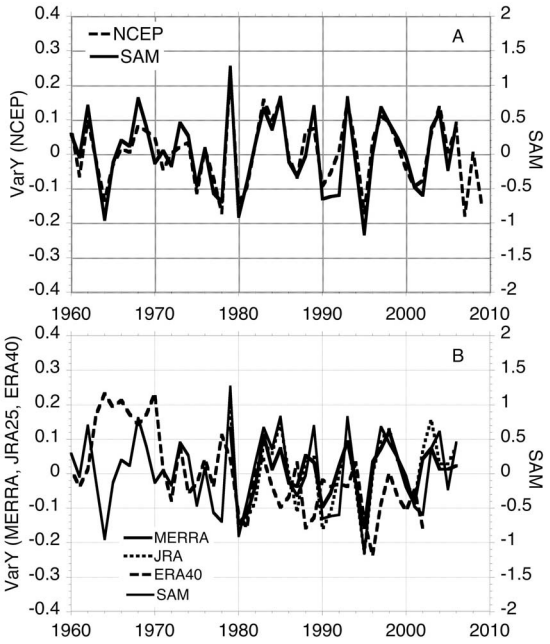

Figure 4. The July Southern Annular Mode (SAM) index, averaged over 10–500 hPa, and the anomaly of ln(VarY), calculated from (a) the NCEP geopotential field for the geometric domain (180E–180W, 20S–90S, 10–500 hPa); and (b) from other reanalysis systems.

but also to determine whether the cause(s) of the aforementioned fluctuation is identical. For example, based on four reanalysis systems we have found that all four systems agree with each other with the absence of any statistically significant trend in the July temperature variability of VVAr over the SH cap (40S–90S). Also, all four reanalysis systems show a higher inter-annual variability in VVar for the upper atmosphere than for the lower atmosphere. Furthermore, a comparison of the VVar components showed that the greatest differences among the reanalysis systems came from differences in variances in the vertical direction. The latter addresses a degree of the vertical coupling of the atmosphere, and, in combination with the latitudinal VVar component, allows for assessing the polar vortex strength, which by itself is an important diagnostic for the upper/lower atmospheres coupling. A comparison of the latitudinal component of the VVar with the SAM index allows for interpreting of all the VVAr components as "leading modes" of atmospheric variability in a particular direction.

[22] An increasing number of different model simulations as well as an increasing size in models' spatial and temporal domains will dramatically increase the number of datasets and the amount of information within them, and thus make their assessment more difficult. The VVar might serve as an initial bulk diagnostic for defining where the data exhibit the most differences. The VVar aim to help researchers navigate throughout massive influxes of data and handle 4D arrays used in the analysis of models and observations. An advantage of this diagnostic is its embracement of a quantity's state over a number of directions simultaneously, which provides a better insight into any tendencies that occur in quantity's state development. The VVar diagnostic is to be used in conjunction with analysis of the single-direction variances, each of which is a component of VVar itself. Also, if the multivariate data are presented in the NetCDF format, then calculation of the new diagnostic will be trivial with the use of scientific visualization software (e.g., IDL, Ferret, NCS, MatLab).

[23] **Acknowledgments.** This work was supported by the NASA Science Mission Directorate (MAP) under Contract No. NNX09AJ42G. We are very thankful to Mary Nehls-Frumkin, Samuel Schlesinger and anonymous reviewers for the comments, which lead to considerable improvement of the paper.

[24] The Editor thanks the two anonymous reviewers for their assistance in evaluating this paper.

References

Baldwin, M. P., X. Cheng, and T. J. Dunkerton (1994), Observed correlations between winter-mean tropospheric and stratospheric circulation anomalies, *Geophys. Res. Lett.*, *21*, 1141–1144, doi:10.1029/94GL01010.

Bosilovich, M., J. Chen, F. R. Robertson, and R. F. Adler (2008), Evaluation of global precipitation in reanalyses, *J. Appl. Meteorol. Climatol.*, *47*, 2279–2299, doi:10.1175/2008JAMC1921.1.

Brasseur, G. P., and S. Solomon (2006), *Aeronomy of the Middle Atmosphere*, *Chem. Phys. Stratos. Mesos. Ser.*, Springer, New York.

Bromwich, D. H., and R. L. Fogt (2004), Strong trends in the skill of the ERA-40 and NCEP/NCAR reanalyses in the high and middle latitudes of the Southern Hemisphere, 1958–2001, *J. Clim.*, *17*, 4603–4619, doi:10.1175/3241.1.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kistler, R., et al. (2001), The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, *Bull. Am. Meteorol. Soc.*, *82*, 247–267, doi:10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2.

Onogi, K., et al. (2007), The JRA-25 reanalysis, *J. Meteorol. Soc. Jpn.*, *85*(3), 369–432, doi:10.2151/jmsj.85.369.

Thompson, D. W. J., and J. M. Wallace (1998), The Arctic Oscillation signature in the wintertime geopotential height and temperature fields, *Geophys. Res. Lett.*, *25*, 1297–1300, doi:10.1029/98GL00950.

Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012, doi:10.1256/qj.04.176.

N. Andronova and S. Boland, Atmospheric, Oceanic and Space Sciences, University of Michigan, 1541D Space Research Bldg., 2455 Hayward St., Ann Arbor, MI 48109-2143, USA. (natand@umich.edu)