Inhomogeneous Background Error Modeling and Estimation over Antarctica

YANN MICHEL AND THOMAS AULIGNÉ
National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 9 July 2009, in final form 11 December 2009)

ABSTRACT

The structure of the analysis increments in a variational data assimilation scheme is strongly driven by the formulation of the background error covariance matrix, especially in data-sparse areas such as the Antarctic region. The gridpoint background error modeling in this study makes use of regression-based balance operators between variables, empirical orthogonal function decomposition to define the vertical correlations, gridpoint variances, and high-order efficient recursive filters to impose horizontal correlations. A particularity is that the regression operators and the recursive filters have been made spatially inhomogeneous. The computation of the background error statistics is performed with the Weather Research and Forecast (WRF) model from a set of forecast differences. The mesoscale limited-area domains of interest cover Antarctica. Inhomogeneities of background errors are shown to be related to the particular orography and physics of the area. Differences seem particularly pronounced between ocean and land boundary layers.

1. Introduction

Producing accurate forecasts over the Antarctic continent is a challenge because of the sparsity of available conventional observations and the difficulties encountered in resolving the effect of the steep topography on the atmospheric flow. Less than a dozen of radiosonde stations operate over the Antarctic continent (Xiao et al. 2008). Since 1980 there has been a significant increase of the number of Automatic Weather Stations (AWS) that provide measurements of wind speed and directions, air temperature and pressure, as well as relative humidity. The spatial distribution of those 100 AWSs is uneven, being mostly in the periphery of the Antarctic continent, as a result of particular projects and national programs. Some drifting buoys data are also available, but a number of them do not disseminate the data early enough for real-time operational applications. The main source of data therefore comes from the satellites, including passive microwave imagery, sounder infrared data, and wind scatterometers. Some of these data are contaminated by the surface ice over the continent. A detailed overview of the observations available over the Antarctic region can be found in King and Turner (1997).

Global models have shown increasing skill in predicting weather at southern latitudes, which is believed to be linked with the increasing amount of satellite data that is assimilated in global models (Simmons and Hollingsworth 2002). However, resolving the steep topography of the Antarctic continent may require the use of regional models to achieve higher spatial resolution. Pendlebury et al. (2003) cite five examples of obvious limitations where global models are not able to resolve key topographic features because of coarse horizontal resolution. It may also be necessary to include a special representation of the physical properties unique to the Antarctic troposphere. The continental boundary layer shows unusual persistent strong winds that can be partly explained through the katabatic wind theory (Ball 1956; Pettré et al. 1990; Parish and Cassano 2003). These surface winds probably induce larger-scale convergence in the troposphere (King and Turner 1997). Wind barrier effects are also frequently reported in case studies. Therefore, limited area modeling may be seen as having potential for complementing global forecasts with higher resolution and physics adaptation to Antarctica. The Antarctic Mesoscale Prediction System (AMPS) has been designed to overcome the difficulties in numerical weather modeling at the Poles. It is originally based on a modified “polar” version of the fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5). The key physical schemes that were...
modified for polar regions include (Bromwich et al. 2001):

- accounting for sea ice with specified thermal properties;
- representing fractional sea ice coverage in grid cells;
- using the latent heat of sublimation for calculations of latent heat flux over ice surface, and assuming ice saturation when calculating surface saturation mixing ratios over ice;
- improving the radiation scheme to include the radiative properties of clouds from the microphysical species;
- modifying the thermal diffusivity for snow-covered permanent ice, and sea-ice grid points;
- increasing levels in the soil model to better represent heat transfer through ice sheets.

Although developed originally with MM5, AMPS now uses the Weather Research and Forecast (WRF) model (Skamarock et al. 2008). The first guess and the lateral boundary conditions are derived from the Global Forecasting System (GFS) developed at the National Centers for Environmental Prediction (NCEP). Over the last few years, AMPS has been shown to provide relevant meteorological guidance for the forecasting in the Antarctic region (Powers et al. 2003). The performance of this model has been statistically evaluated versus observations for a 2-yr period by Bromwich et al. (2005). AMPS has also been successfully used to study the prediction of some severe synoptic events such as the May 2004 McMurdo, Antarctica, storm (Powers 2007; Steinhoff et al. 2008).

The accurate specification of the model initial state is especially important when the sensitivity of the prediction to this initial state is found to be high (Xiao et al. 2008). This can be achieved through advanced data assimilation schemes that are able to handle the special properties of background errors over Antarctica. Two main schemes are available for the AMPS model: an ensemble square root filter (Barker 2005) and a three-dimensional variational assimilation (Barker et al. 2004). In the latter case, error covariances are typically based on offline computations of simplified statistics. The background error samples are often approximated through forecast differences (Parrish and Derber 1992), or through ensembles of forecasts from data assimilation with perturbed observations (Pereira and Berre 2006). The dimension of the full covariance of background errors is far too large to be either stored or even to be estimated. Therefore, the problem is made manageable by reducing the number of degrees of freedom in the background error covariance matrix B. Usually, this matrix is taken to be diagonal, or sparse, in some carefully chosen space. In a variational assimilation scheme, B is often modeled as a sequence of operators using control variable transforms (Derber and Bouttier 1999). This has several advantages: reducing the dimension of B, ensuring physical balance constraints, and improving the conditioning of the minimization (Courtier and Talagrand 1990). Bannister (2008b) reviews the different choices that have been made for the balance part of the transform and for the modeling of spatial correlation in different variational schemes. As it will be further described in section 2, the WRF control variable transform (CVT) uses regressions for the balance, empirical orthogonal functions (EOF) for the vertical covariances, and recursive filters for the horizontal correlations. The main advantage of this gridpoint formulation is to more easily incorporate geographical variations of the covariances. It has however the drawback of reducing the freedom to specify the shape of error statistics in wavenumber space compared to spectral formulations of the background term (Rabier et al. 1998; Berre 2000). Wu et al. (2002) incorporate a latitude dependence of the covariances in a global model, and show this results in a positive impact in the tropics compared to the previous Spectral Statistical-Interpolation (SSI) data assimilation scheme.

The numerical weather prediction over Antarctica and the understanding of its climate have recently been key topics of the Fourth International Polar Year (March 2007 to March 2009), organized through the International Council for Science and the World Meteorological Organization. The meteorology part of the experiment aims to learn more about the Antarctic climate, particularly snowfall, and to improve the predictions of weather models in this region of the globe where weather stations are few and far between. The Concordiasi project (Rabier et al. 2007) will also provide validation data to improve the usage of polar-orbiting satellite data over Antarctica, in particular from the Infrared Atmospheric Sounding Interferometer. Good impact of these radiances will certainly require background error statistics that are tuned for Antarctica.

The second section of this paper describes some general aspects of numerical modeling for weather prediction over the polar regions and presents the AMPS models. The third section covers the extension of the grid space formulation of the CVT to the inhomogeneous case. The balance is performed using local regression, and an improved binning strategy allows to represent geographical variations. Inhomogeneity in the horizontal correlations is incorporated through the use of inhomogeneous recursive filters (Purser et al. 2003b). Specific issues such as the normalization and the boundary conditions are highlighted. The fourth section illustrates the computations of the background error statistics over Antarctica with the two AMPS domains used in this study. Focus is put
on the depiction of inhomogeneity for mesoscale data assimilation. The extensive testing of this new formulation on cycled data assimilation experiments is planned to be reported on in a future paper. Perspectives for Antarctica numerical weather prediction and mesoscale data assimilation are discussed in the conclusions.

2. Numerical weather prediction over Antarctica

The background errors may depend on a variety of factors, including data density and distribution, but the main dependence is on the region of interest. The coupling between variables by the dynamics is very different in the tropics from the midlatitude for instance, but it may also largely vary within those regions especially for the mesoscale features. This is made very clear in the comparisons made by Sadiki et al. (2000) and Montmerle et al. (2006) with a regional model. It is therefore important to summarize the main characteristics of the weather over Antarctica, as it is very likely that they will be reflected in the background errors statistics.

a. Climatological weather over Antarctica

Recent studies have benefited from the use of long-term reanalysis to produce climatologies over Antarctica. Simmonds et al. (2003) used the NCEP reanalysis set over an extended period (1979–2000) to diagnose Eulerian climatological properties of the circulation. Insight into synoptic weather is further gained through the application of an objective tracking algorithm that depicts cyclone trajectories and their Lagrangian characteristics. The Antarctic coast has been found to be the site of significant cyclogenetic activity. Hoskins and Hodges (2005) applied another modern feature tracking technique to the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40). They found greater asymmetry of the cyclone tracks in wintertime than in summertime. The storm track has a spiraling shape toward the Antarctic continent. The genesis and lysis of cyclones are shown to be influenced by the Andes. Idealized atmospheric experiments performed by Inatsu and Hoskins (2004) suggest that the sea surface temperature, the topography, and tropical convection asymmetries influence the storm track intensity and position. The topography of the Antarctic plateau acts as a natural barrier for low-level cyclones, and anticyclonic conditions prevail in the interior.

Part of the coastal region of Antarctica encounters strong and persistent surface winds that are directed along topographic pathways. Since the first application of katabatic wind theory to these surface winds (Ball 1956), the subject has received considerable attention. It is often considered that this phenomenon is not restricted to the boundary layer, but that it also may strengthen the anticyclone over the continent through large-scale convergence in the upper troposphere (King and Turner 1997).

b. The Antarctic Mesoscale Prediction System

The AMPS modeling system currently features six grids of various horizontal spacing ranging from 45 to 1.6 km. Specific initialization through data assimilation is only performed for the two largest domains, shown in Fig. 1. The first domain extends up to New Zealand, to cover meteorological conditions for the flights toward McMurdo, and has currently 45-km resolution. The second domain covers the whole Antarctic continent with an improved spatial resolution of 15 km. Other domains cover local areas near the Ross Ice Shelf and McMurdo, or near the South Pole around Amundsen–Scott (Powers et al. 2003). All nestings are two-way interactive. The vertical resolution has 44 hybrid pressure levels, reaching 10 hPa at the top (Fig. 2).

Various verification studies attest the skill of the model in predicting key features of Antarctic weather. Guo et al. (2003) performed a complete annual cycle of forecasts and compared it to various observations and global analysis. They found consistent good performance in predicting both large- and regional-scale features. The main caveats concern biases found for near-surface temperature and cloud cover. Further validation in an operational context is provided by Bromwich et al. (2005). A recent increase in model vertical and horizontal resolution has been done to reduce model errors.

The next step in numerical weather prediction (NWP) consists in trying to improve the model background with available observations through data assimilation. AMPS uses a three-dimensional variational data assimilation (3DVAR) similar to the one developed for MM5 (Barker et al. 2004), and is able to assimilate various kinds of data, including conventional data (reports from surface stations, surface automatic weather stations, upper-air stations, ships, buoys, pilot and aircraft reports, etc.), satellite cloud track winds (Powers 2007), and Global Positioning System Radio Occultation Refractivity (Wee et al. 2008). Radiiances or retrievals from satellite sounders are likely to be added in a near future together with a variational bias correction scheme (Au- ligné and McNally 2007). Advanced background error modeling is needed over Antarctica for some specific reasons: in data-rich areas, background error covariances generally exhibit simpler patterns and less spatial variability (including less anisotropy) than in data-sparse areas and near continent borders. Clearly, the Antarctic region features both a lack of observations and a complicated topography near coastlines.
3. Background error modeling in WRF

a. The control variable transform in 3DVAR

In general, variational assimilation schemes are designed to provide an analysis \( x_a \) that minimizes a cost function \( J(x) \):

\[
x_a = \text{Arg min } J, \tag{1}
\]

\[
J(x) = \frac{1}{2} (x - x_b)^T B^{-1} (x - x_b) + \frac{1}{2} [y - \mathcal{H}(x)]^T R^{-1} [y - \mathcal{H}(x)], \tag{2}
\]

where, denoting by \( n \) the dimension of \( x \) and by \( p \) the dimension of the observations \( y \):

- \( \mathcal{H} \) is the nonlinear observation operator;
- \( B \) of dimensions \( n \times n \) is the background error covariance matrix; and
- \( R \) of dimensions \( p \times p \) is the observation error covariance matrix.

We also define the innovation vector to be the departure between observation and background: \( d = y - \mathcal{H}(x_b) \). The incremental approach (Courtier et al. 1994)
solves the problem with respect to $v = x - x_b$ using linearization:

$$J(v) = \frac{1}{2} v^T B^{-1} v + \frac{1}{2} (d - Hv)^T R^{-1} (d - Hv),$$

(3)

where $H$ of dimensions $p \times n$ is the linearized observation operator and where the minimization is easier to accomplish because the cost function is quadratic. The change of variable $v = B^{1/2} u$, where $B^{1/2}$ is a square root of $B$, reduces the cost function to be

$$J(u) = \frac{1}{2} u^T u + \frac{1}{2} (d - HB^{1/2} u)^T R^{-1} (d - HB^{1/2} u).$$

(4)

Different choices for the construction of $B^{1/2}$ are reviewed by Bannister (2008b).

The choice made for the WRF 3DVAR (Barker et al. 2004) was adapted from the Met Office CVT defined by Lorenc et al. (2000), but the horizontal correlations were prescribed with homogeneous recursive filters for the regional models, based on Hayden and Purser (1995).

The update over the formulation of Barker et al. (2004) consists in a new CVT:

$$v = U_p u U_{ih} S u,$$

(5)

where $S$ is a multiplicative variance scaling factor, $U_{ih}$ is the application of high-order, fully inhomogeneous recursive filters to impose horizontal correlations (Purser et al. 2003b), $U_v$ is the application of vertical correlations through homogeneous EOF, and $U_p$ changes the control variables to model variables through statistical relationships.

b. The variance scaling

The variance scaling transform $v_s = Su$ is a simple multiplication in gridpoint space. The elements of the diagonal matrix $S$ are the background error standard deviations of the unbalanced variables projected onto EOF modes. Small-scale noise is apparent in variance (or standard deviations) maps even using 100 realizations (Berre et al. 2007). In operational implementations of variational assimilation, variances maps are often spatially filtered. Raynaud et al. (2009) described an objective filter when variances are estimated from a very small ensemble. In our study, a crude spatial filtering is applied to variance maps through convolution with an exponentially decaying function with typical length scales of 450 and 150 km over domains 1 and 2, respectively. No attempt has been made yet to make this filtering adaptive to either the vertical EOF level or the variable.

c. The recursive filters

Recursive filters are used to impose horizontal correlations, using the following CVT:

$$v_v = U_{ih} v_s.$$  

(6)

1) The basic algorithm and its computational cost

In the absence of precise information, it is natural to consider isotropic spatial correlation functions. Applying a convolution with an isotropic function can be done in physical grid space, or in spectral space where it takes the form of a multiplication (Gaspari and Cohn 1999). In one dimension, with a grid size denoted by $N$, the direct convolution with an arbitrary function uses $O(N^2)$ operations, whereas fast Fourier transforms achieve a computational cost of $O(N \log(N))$. Therefore, the spectral transform becomes a more computationally efficient way to impose homogeneous correlations as soon as the grid size is larger than a few hundred (Smith 2007).

Recursive filters are also an efficient technique in correlation modeling. The basic steps of recursive filters of order $n$ applied to a one-dimensional field $p = [p_1, \ldots, p_i, \ldots, p_N]^T$ transformed into $s$ may be written (Purser et al. 2003a) as

$$q_i = \beta p_i + \sum_{j=1}^{n} \alpha_j q_{i-j},$$

(7)

$$s_i = \beta q_i + \sum_{j=1}^{n} \alpha_j s_{i+j},$$

(8)

which correspond to the advancing and backing steps, respectively. The number of operations needed is therefore
The obtained correlation is quasi-Gaussian, but a wider range of shapes can be achieved through linear combinations at the expense of computational cost. The coefficients $a_j$ of the recursive filters can be built from the length scale normalized by the grid spacing. This procedure involves the construction of a properly weighted differential operator acting on the grid, and then the Choleski decomposition of its matrix representation. Suitable boundary conditions have also to be provided when working with a noncycling grid as in a regional model (Purser et al. 2003a).

2) INHOMOGENEOUS BIDIMENSIONAL QUASI-GAUSSIAN RECURSIVE FILTERS

Inhomogeneous bidimensional quasi-Gaussian recursive filters have been designed to take into account the metrical properties of the grid and to incorporate varying length scales (Purser et al. 2003b). The weighted differential operator is in this case much more costly to compute, and Choleski decomposition has to be performed now for every line of the two-dimensional domain. This may prove prohibitive for the computations of the correlations of the day, where the $B$ matrix is recomputed or updated at each analysis cycle.

The AMPS stereographic polar projection for the first domain is large enough to require inhomogeneous recursive filters even with homogeneous length scales because the map scaling factor becomes nonnegligible. Figure 3 illustrates the application of the inhomogeneous recursive filters when the map scaling factor is taken into account through the metric term, resulting in apparently larger increments in the center of the domain compared to the borders. No undesirable anisotropy is noticeable, despite the fact that the scale of the increment diminishes by 30% near the corner of the domain compared to the center.

Another issue is the computation of the normalization factor. The same problem is found in correlation modeling with a diffusion equation. Weaver and Courtier (2001) suggested using either a direct amplitude estimation by applying the filter to an impulse on each grid point or the randomized estimation of Andersson et al. (2000). Purser et al. (2003b) showed that an asymptotic analysis could provide a good estimation of the normalization factor. Pannekoucke and Massart (2008) applied this approach to study the error made on the normalization in an idealized one-dimensional diffusion context. The estimation of the normalization was found to be accurate in a case of smooth enough length scale fields. The filtering of background error statistics should also take into account the limited size of the sampling. We choose to make the spatial filtering scale of each variable and vertical mode to be proportional to the median length scale computed over the domain. A consequence is that wind length scales maps will be more heavily filtered than mass length scales maps. Figure 4 shows the relative error made on the normalization factor versus the filtering index (which is a cutoff wavenumber inversely proportional to the scale of filtering), for different variables and vertical modes (see next paragraph for the definition of the vertical transform). The error is computed as the root-mean-square error between the asymptotic analysis using a filtered diffusion and the real estimation using impulse responses at each grid point. The normalization error does not vary much among the vertical modes. Different variables have however different error slopes despite the fact that the filtering scale has been made dependent on the median length scale, which may be related to probable differences in length scales power spectra. The error generally increases when the filtering scale decreases and increasing in cases of strong normalization errors occur (as depicted by the uncertainty interval). A filtering index of 4 is used, which guarantees that the normalization error is lower than 7%. The spatial filtering has been set up as the maximum of a fixed value that intends to remove
the sampling noise (and which is the same as the one for the variances) and of a value made proportional to the median length scale of the variable. Practical consequences of this filtering will be further described in section 4, together with the depiction of Antarctica background error covariances.

d. The vertical transform

The vertical part of the transform is applied via an empirical orthogonal function decomposition of the vertical component of background error \( \mathbf{B}_v \) on model levels (Barker et al. 2004). This is a widely used choice (Bannister 2008b). A time- and domain-averaged estimate of \( \mathbf{B}_v \) is chosen to take the form \( \mathbf{B}_v = \mathbf{E} \Lambda \mathbf{E}^T \), where \( \mathbf{E} \) represents the eigenvectors and the diagonal matrix \( \Lambda \) contains the associated eigenvalues. The vertical part of the CVT takes the simple form of the change of basis:

\[
\mathbf{v}_p = \mathbf{U}_v \mathbf{v}_v = \mathbf{E} \Lambda^{1/2} \mathbf{v}_v.
\] (9)

Fig. 4. Mean standard deviation of relative error on the recursive filters normalization vs the filtering index for (a) streamfunction, (b) unbalanced velocity potential, (c) unbalanced temperature, and (d) relative humidity. The relative error for unbalanced surface pressure, not shown, is similar to the wind case [(a),(b)]. Results are shown for the first three EOF. The 90% Gaussian confidence interval on this relative error for the first EOF (obtained from 200 samples) is also shown in gray shading.
e. The balance transform

1) IMPLEMENTATION

The current balance transform implemented in WRF-Var is statistical and defined in physical space. The control variables are the streamfunction $\psi$, the unbalanced part of velocity potential is $\chi_u$, the unbalanced part of temperature is $t_u$, the unbalanced part of surface pressure is $P_s$, and the relative humidity is $\text{rh}$. The balance transform is defined by

$$
\begin{bmatrix}
\psi \\
\chi \\
t \\
P_s \\
\text{rh}
\end{bmatrix}
= 
\begin{bmatrix}
I & 0 & 0 & 0 \\
M & I & 0 & 0 & 0 \\
N & P & I & 0 & 0 \\
Q & R & 0 & I & 0 \\
0 & 0 & 0 & 0 & I
\end{bmatrix}
\begin{bmatrix}
\psi \\
\chi_u \\
t_u \\
P_{s_u} \\
\text{rh}
\end{bmatrix},
$$

(10)

where $I$ stands for the identity matrix, and $M, N, P, Q, R$ are regression matrices. The block lower-triangular shape of the complete balance operator is designed such as it is easily invertible, which is required in the calibration part (Derber and Bouttier 1999). The matrices are computed from statistical regression in gridpoint space (Wu et al. 2002).

The correlation between streamfunction and velocity potential defined by $M$ accounts for the friction effect at lower levels in cyclones (Berre 2000). Here $M$ is assumed diagonal, reflecting that the correlation is localized, and depends on the vertical level. It is also varying with geographical position for the new inhomogeneous formulation.

The mass–wind coupling is taken into account by the $N$ and $P$ operators that project streamfunction and unbalanced velocity potential increments to a vertical profile of temperature increments. The $P$ operator links temperature with the divergent part of the flux, and is known to be more statistically significant as the resolution increases (Berre 2000). Both $N$ and $P$ depend on the geographical position for the inhomogeneous formulation.

The humidity is still treated as univariate, although there is ongoing work to couple it with other variables (Krysta et al. 2009). The increase in memory storage to compute the regression coefficients for humidity is significant, and may be prohibitive in a fully inhomogeneous formulation. Therefore, we use coarse binning strategies to compute and store the regression coefficients.

2) BINNING STRATEGIES

The dimensions of the AMPS grids are denoted $n_i, n_j$ on the horizontal, and $n_k = 43$ on the vertical. For the domain 1, we have $n_i = 220$ and $n_j = 290$; and the domain 2 is about 4 times more computationally expensive with $n_i = 442$ and $n_j = 418$. In a fully spatially inhomogeneous framework, Operators $M, Q, R$ require the computation and storage of $n_i n_j n_k$ coefficients, which size scales as the control variable. Operators $N$ and $P$ require the computation and storage of $n_i n_j n_k^2$ coefficients (typically $10^8$ for the second domain), which is too expensive to store yet. Moreover, when computed this way, the regression coefficients exhibit significant sampling noise. In the framework of WRF-Var, we need spatial averaging (to improve covariance estimation) as well as compacted storage (to reduce memory cost). This is performed thanks to a binning strategy, where the regression coefficients are computed and stored over bins, typically on a coarser horizontal grid. Spatial filtering is also necessary to avoid numerical artifacts in $B$.

The geostrophic coupling is mainly latitude dependent, which naturally leads to a latitude binning. This is very similar to the degree of geographical variations that global models can currently describe (Derber and Bouttier 1999; Wu et al. 2002). We have designed such a binning for the AMPS domain 1, taking into account the polar stereographic grid, and partitioning the domain into 11 latitude bands. For higher resolution, we may expect boundary processes to be important in those correlations (Berre 2000), such as more advanced binnings should be used to study and describe geographical variations of regression operators. We will now describe the results and the physical meanings of the relationships implied by these operators.

4. Background error statistics over Antarctica

a. Methodology

The Antarctic region provides a challenging environment to design and test new data assimilation techniques owing to the sparsity of available observations and to the unusual atmospheric processes occurring. The study of background errors has drawn considerable attention in the community, given the usually large impact that their representation has on NWP performance. The study of low-resolution Kalman filters that propagate covariances has been useful in understanding the dynamical properties of errors (Bouttier 1994). It is often very difficult to distinguish sampling errors from properties of background errors, as shown by the ensemble Kalman filter framework. Authors often put a stress on features that have a physical meaning, such as the non-separability (Ingleby 2001), or the geostrophic coupling (Derber and Bouttier 1999). Similarly, we will try to physically describe the relationships implied by $B$.

Measuring $B$ and calibrating the associated CVT can be achieved through several strategies. First, time averages are often performed to increase the sample size,
which is an assumption of ergodicity. Sample of errors can be produced through forecast differences valid at different ranges, which is known as the National Meteorological Center (NMC, currently NCEP) method (Parrish and Derber 1992; Rabier et al. 1998). An alternative favored method is to rely on ensembles of forecasts. The ensemble can be build from variational assimilations with perturbed observations. Berre et al. (2006) showed that the cycle of errors was better represented in the ensemble approach. Some systematic differences have been found between the two methods, especially regarding length scales (Buehner 2005; Pereira and Berre 2006). Recently, Météo-France and the ECMWF have moved to an operational ensemble-based method (Bannister 2008a).

We will however use the convenient NMC method, as few differences have been encountered yet with the ensemble-based method within the WRF-Var context (H. Huang 2009, personal communication) maybe because of the youth of this data assimilation scheme. In AMPS operational system, B is recomputed every month over a 1-month dataset to catch seasonal features of errors. We have kept this general framework but use a 2-month period (1 April–1 June 2009). Statistics are computed for the first two AMPS domains of respective resolutions 45 and 15 km (see Fig. 1 for their spatial extensions), through differences between 24- and 12-h forecasts verifying at the same time.

To our knowledge, this paper is the first study devoted to the examination of background error covariances over the Antarctic region. Some of the mesoscale features of the cross covariances described by Berre (2000) can be understood as the result of thermodynamic processes, which are expected to be different over the seas and the Antarctic plateau. Ingleby (2001) reported also a change in sign for the large-scale cross covariance between temperature and surface pressure, proving that large-scale balance may also be inhomogeneous over the domain. Physical meaning of the balance relationships can be inferred with the computation of vertical cross covariances. Their statistical impact can be deduced from the ratio of the variance of the balanced variables over the total variance.

b. Background error statistics for the first domain

We describe now the structure of the balance operators on the 45-km resolution domain. Cross covariances appear quite similar in structure to the ones of Berre (2000), but our control variable is different. The streamfunction is proportional to geopotential height, thus proportional but of opposite sign to the linearized mass variable of Derber and Bouttier (1999). The velocity potential is the inverse Laplacian of the divergence, thus larger scale and negatively correlated to divergence.

1) Balance statistical coupling

The statistical coupling $M$ between streamfunction and velocity potential is shown in Fig. 5. Its precise structure depends on latitude (Figs. 5a–c). At lower levels, ageostrophic flow is caused by the frictional drag of the underlying surface. This is the well-known Ekman effect in ocean and atmospheric dynamics: low-level inflow (outflow) is observed for cyclones (anticyclones) in the boundary layer. The ageostrophy causes upward (downward) vertical motion in cyclones (anticyclones), which is compensated by divergence (convergence) at upper levels, thus the negative covariance with streamfunction above (Fig. 5). As shown by Figs. 5d, this effect has however a statistical impact only at the lower levels (up to model level 10, about 850 hPa). Consistent with the decrease of geostrophy, the explained velocity potential variance ratio decreases equatorward. There is also a decrease of statistical impact at higher latitude over the Antarctic plateau. It is known that the low-level wind is strongly influenced by topography and katabatic phenomena, which may in turn decrease the correlation between streamfunction and velocity potential background errors.

The mass–wind coupling $N$ is presented in Fig. 6. At midlatitude over seas, a positive error in temperature at lower levels is correlated with a negative error in streamfunction in the midtroposphere (Fig. 6b). This is consistent with a potential vorticity view of the circulation (Hoskins 1997; Fig. 6c). In the boundary layer, the covariance rather tends to be positive, which is probably the effect of surface heating. Over the continent (Fig. 6a) this covariance tends to be negative. This may indicate that stronger anticyclonic conditions are associated with temperature radiative cooling. Therefore, the land contrast (sea versus ice) may lead to inhomogeneous correlations between temperature and streamfunction. The change of sign of the covariance below and above the tropopause (model level 26, about 300 hPa) may be explained through the observed vertical autocovariance structure of temperature (Berre 2000). The statistical impact of this coupling (shown in Fig. 6d) is maximum toward the midtroposphere where it reaches 0.5, consistent with the values found for large-scale models (Derber and Bouttier 1999; Wu et al. 2002), and increases poleward, consistent with geostrophy.

The coupling between temperature and unbalanced velocity potential is depicted in Fig. 7. Figure 6b shows the vertical cross covariances obtained at midlatitude over seas, and is very similar in structure to the cross covariance between temperature and divergence (but of opposite sign) depicted by (Berre 2000, his Fig. 11b). Convergence at the lower levels induces upward vertical motion, thus a decrease of temperature, which is vertically
stratified. This coupling is important at the lower levels, where it explains about 5%–10% of the variance (Fig. 7d), and in the lower stratosphere where it explains about 10%–20% of the variance. Over the Antarctic plateau, strong negative correlations exist between boundary layer temperature and midlevel unbalanced velocity potential (panel a). This may be related to the fact that low level katabatic winds may be strong enough to force upper-level convergence (King and Turner 1997). There seems to be a slight increase of the amount of temperature variance explained by unbalanced velocity potential at 90°S in the midtroposphere (Fig. 7d), reaching about 10% in the upper troposphere.

Surface pressure errors are also balanced with streamfunction and unbalanced velocity potential, as shown in Fig. 8. The statistical significance generally increases poleward, except when crossing the high topography over the continent (near 70°S) where the balance decreases.

So far, the balance transforms typically resemble those of midlatitude, as they have been described by numerous authors (Bouttier et al. 1997; Derber and Bouttier 1999; Berre 2000; Wu et al. 2002). Some differences can

---

**Fig. 5.** Statistical relationship between streamfunction and velocity potential background errors over AMPS domain 1. (a),(b),(c) Vertical cross-covariance matrices at polar, mid-, and subtropical latitudes, drawn with solid (dashed) lines for positive (negative) values (units: $10^{10}$ m$^4$ s$^{-4}$, contour interval 2.0). (d) Ratio of explained variance of velocity potential as a function of model level (contour interval 0.1).
however be seen in the balance above the continent, including
- a lower friction coupling between streamfunction and velocity potential,
- a lower surface pressure balance,
- a change of sign of the correlations between temperature and streamfunction in the boundary layer, and
- a slight increase of the coupling between temperature and unbalanced velocity potential.

These features are interesting in the scope of regional data assimilation and will be further described over the higher-resolution second domain.

2) VERTICAL CORRELATIONS

The vertical correlations of control variables are depicted in Fig. 9. The streamfunction correlations (Fig. 9a) are generally positive and quite broad (except at the upper levels), with a local widening at the tropopause. The unbalanced velocity potential correlations (Fig. 9b)
are much sharper than the streamfunction correlations, and a negative lobe, which is consistent with the fact that integrated divergence should be close to zero. These vertical correlations are very similar to the ones described by other authors at the midlatitude (Rabier et al. 1998; Ingleby 2001; Montmerle et al. 2006).

Humidity correlations (Fig. 9d) are sharp on the vertical, with a noticeable tightening just at the boundary layer top (toward model level 9). This partial decorrelation between humidity errors in the boundary layer and above may be due again to the special surface conditions over the Antarctic plateau. Surprisingly, this is a characteristic shared with humidity errors in convective areas (Montmerle et al. 2006, their Fig. 4), where boundary conditions are also of great importance.

The vertical part of the CVT uses EOF decomposition rather than those correlations to impose vertical covariances. The first three leading eigenvectors are drawn in Fig. 10. They have similar structures than the ones presented in Ingleby (2001) and Barker et al. (2004). One noticeable exception is the first mode of unbalanced temperature that peaks at lower levels. This may be due
to the fact that the vertical background error covariance matrix is not weighted by the pressure height of the levels, giving larger weights to the boundary layer. The first mode of unbalanced velocity potential has an oscillating structure consistent with the fact that total column divergence should be close to zero. The first mode of relative humidity peaks in the midtroposphere with monotonic correlations along the whole vertical column.

3) LOCAL VARIANCES

Background error variance rescaling factors are shown in Fig. 11. They seem clearly linked with the circulation of synoptic systems on the seas around the continent, especially for the streamfunction (Fig. 11a), the relative humidity (Fig. 11d), and the unbalanced surface pressure (Fig. 11e). Unbalanced temperature (Fig. 11c) is heavily influenced by the contrast between seas and land, consistent with the peak of the associated EOF. For all variables, there seems to be a significant reduction of the variance toward the border that is linked with the larger-scale common boundary conditions from the lagged NMC method (Pereira and Berre 2006).

4) LOCAL HORIZONTAL LENGTH SCALES

Horizontal length scales are a simple diagnostic of the often complicated shape of real background error correlation functions. They are mostly estimated through simple local formulas (Pannekoucke et al. 2008). This simplification can make it doubtful whether such simple estimates can be used within an inhomogeneous data assimilation scheme, but the results of the NCEP global assimilation scheme from Wu et al. (2002) were positive. We also estimate the (local) length scales through the ratio of the variance of a field and the variance of the Laplacian of this field. For instance, the correlation length scale of streamfunction is estimated via

$$L = \left[ \frac{V(\psi)}{V(\xi)} \right]^{1/4},$$

where $\xi$ is the vorticity (the Laplacian of streamfunction) and is computed through spectral transform, taking into account the map projection factor, and $V$ is the sample averaged variance in the NMC method.

The horizontal length scales of background error correlations are known to be strongly influenced by dynamics first, and also by data density effects (Bouttier 1994) and topography (Deckmyn and Berre 2005). For large-scale balanced variables, there is a geostrophic relationship between horizontal scale $\Delta L$ and vertical scale $\Delta Z$:

$$\Delta L = \frac{N}{f_0} \Delta Z,$$

where $N$ is the Brunt–Väisälä frequency and $f_0$ is the Coriolis parameter (Lindzen and Fox-Rabinovitz 1989). The tropopause is lower and the Coriolis parameter is larger over the polar region, such as one would expect smaller length scales for streamfunction poleward. This is effectively the case as shown in Fig. 12a. Relative humidity errors also exhibit this poleward decrease of horizontal length scales. Unbalanced surface pressure seems to be correlated over large scale in the Eastern part of the Antarctic plateau (Fig. 12e). This local maximum of length scales is also visible for unbalanced temperature (Fig. 12c), unbalanced divergence, and maybe streamfunction. This may be related to large-scale surface winds that are driven by thermal and orography effects. It is also possible that the coarse resolution of domain 1 would lead to significant model errors in the orography that are reflected in the length scales maps of surface pressure background error. On the contrary, Deckmyn and Berre (2005) have noticed a reduction of length scales over the mountains at lower levels. Some orographic effects may also be seen for unbalanced surface pressure and unbalanced temperature, with minima over New Zealand, over South America, and along the coastline of Antarctica.

It is also clear that filtering has a strong effect on length scales maps. Wind length scales are generally about 3 times larger than unbalanced mass length scales, thus requiring heavy spatial filtering because of the recursive filter normalization issue. Whereas filtering of unbalanced
temperature or relative humidity length scales was found to keep the main geographical variations observed in unfiltered maps (not shown), this was not the case for streamfunction and unbalanced velocity potential, where strong filtering proved necessary (Figs. 12a,b). This may be a motivation for changing the control variable to vorticity and divergence, at the expense of a Laplacian inversion.

Figure 13 shows the application of inhomogeneous recursive filters where both the inhomogeneous length scales from Fig. 12 and the inhomogeneous map scaling factor from Fig. 3 are taken into account. This simulates unbalanced surface pressure increments. The main geographical variations previously described are reflected, with noticeable wider increments over the eastern part of the continent. Undesirable anisotropy is barely visible, implying that single-pass recursive filters are able to represent a fair degree of spatial inhomogeneity without any grid-oriented artifact (which otherwise would require several passes of recursive filters to disappear).

c. Background error statistics for the second domain

The preceding depiction has pointed out some special properties of background errors over the high southern
latitudes. The effect of increasing the horizontal resolution will be assessed with the second AMPS domain. Background errors are computed over Antarctica at 15-km resolution, and the results confirm the need for inhomogeneous modeling.

1) BALANCE STATISTICAL COUPLING

The geographical variations of the main statistical couplings are depicted in Fig. 14. They are generally consistent with the findings over domain 1 but the new binning allows the investigation of the longitude dependence. The balance between streamfunction and velocity potential (Fig. 14a) decreases over the continent and this seems to be linked with orography (Fig. 1). As a result, the eastern part of the continent shows weaker Ekman coupling than the western part. The geostrophic coupling between temperature and streamfunction, shown in Fig. 14b at a midtropospheric level, shows the poleward increase observed over domain 1. However, the balance is also significantly weaker over high slopes, and especially over the eastern coastline. This is believed to be related to the $\sigma$ formulation of vertical levels (Berre 2000). The coupling between temperature and unbalanced velocity potential (Fig. 14c) seems to be higher over the slopes, however, which may be related to katabatic flows. The geographical variations of surface pressure balance (Fig. 14d) are close to the ones of temperature, with general poleward increase but local decrease over slopes.

FIG. 10. First three eigenvectors of the EOF decomposition of the vertical background error covariance matrix for control variables over AMPS domain 1: (a) streamfunction, (b) unbalanced velocity potential, (c) unbalanced temperature, and (d) relative humidity, as a function of model level.
FIG. 11. Rescaling factor of variance for the first EOF for (a) streamfunction, (b) unbalanced velocity potential, (c) unbalanced temperature, (d) relative humidity, and (e) unbalanced surface pressure.
FIG. 12. Local length scales (km) for the first EOF for (a) streamfunction, (b) unbalanced velocity potential, (c) unbalanced temperature, (d) relative humidity, and (e) unbalanced surface pressure.
2) VERTICAL CORRELATIONS

The vertical correlations over the second domain are hardly different from the ones over the first domain (not shown). The modest increase of horizontal resolution does not lead to an obvious vertical sharpening. The main difference concerns unbalanced velocity potential, where larger correlations and stronger negative correlations occur in the lower stratosphere over the plateau.

The EOF decomposition obtained over domain 2 (Fig. 15) is also similar to the ones obtained over domain 1 (Fig. 10) except for the unbalanced velocity potential. The first EOF peaks at model top rather than representing errors that are anticorrelated between the lower and upper stratosphere (as the second EOF does). This is different from the results of Barker et al. (2004), where EOF decompositions were found to be robust against varying horizontal resolutions.

3) LOCAL VARIANCES

Background error variance rescaling factors over the second domain are shown in Fig. 16, and prove to be very similar to the variances obtained over the first domain (Fig. 11). The higher resolution does not seem to bring new features. The higher-resolution domain affects the unbalanced temperature variances, which seem to spread across a lower range of values. The lateral boundary conditions seem to increase the variance on the borders for unbalanced surface pressure.

4) LOCAL HORIZONTAL LENGTH SCALES

The diagnosed local horizontal length scales over domain 2 are shown in Fig. 17. The increase of horizontal resolution yields lower length scales than over domain 1. A striking feature is that now all variables show a similar distribution, even if streamfunction and unbalanced velocity potential are larger-scale variables. The correlations are larger in the eastern part of the Antarctic plateau, in clear contrast to the western part. A local minimum of length scales is apparent all along the coastline of the continent, which may indicate partial decorrelation between errors above sea and above the continent. This is somewhat different from the results obtained over domain 1, where the spatial variations of the length scales seemed to be linked with overall latitude gradient. This may be due to the higher resolution and thus better representation of the orography. Using a fully inhomogeneous formulation, rather than a simpler latitude dependence, allows one to represent those east–west asymmetries, which is a desirable feature of mesoscale background error modeling.

5. Conclusions and perspectives

In data-sparse areas, the correct specification of the background error covariance matrix is a key element in spreading the information retrieved from the observations. The Antarctic region still presents some unique challenges for regional numerical weather prediction: difficulties arise from poor first-guess and lateral boundary condition (as global models may be tuned for mid-latitude weather characteristics), shortage of conventional observations, steep and complex topography, and special physical conditions that prevail over the plateau. These conditions may lead to inhomogeneous background covariance matrix, as errors are expected to be strongly driven by dynamics (Bouttier 1994).

This paper addresses the problem of inhomogeneous modeling, including a local balance transform, a local variance scaling and inhomogeneous recursive filters, which can be achieved at a reasonable cost thanks to the WRF-Var gridpoint formulation. The balance uses local regressions in gridpoint space described by Wu et al. (2002). The inhomogeneous recursive filters are based on the work by Purser et al. (2003b), and allow one to...
represent a significant amount of inhomogeneity. Spatial filtering of length scales was found to be necessary to impose good control on amplitudes. We made the filtering intensity proportional to the length scale itself. A practical consequence is that inhomogeneity is better represented for small background scale variables (e.g., unbalanced mass variables) than for wind variables. It is probably possible to solve this problem by choosing vorticity and divergence as wind control variables. It is also necessary to work on formulations that represent the anisotropy of the correlations. Anisotropy may indeed be important over the Southern Oceans where one can expect the circumpolar vortex to stretch errors in the azimuthal direction.

Fig. 14. Explained variance ratios for the main statistical couplings over domain 2: (a) explained variance ratio of velocity potential at the lowest level, (b) explained variance ratio of temperature by streamfunction at a midtropospheric level, (c) explained variance ratio of temperature by unbalanced velocity potential at lowest level 1, and (d) total explained variance ratio of surface pressure. Those fields, defined on a coarse binning grid, have been spatially filtered for convenient display.
The covariances of background error over the Antarctic region for the Antarctic Mesoscale Prediction System have been described. Statistics have been computed for two domains with horizontal resolutions of 45 and 15 km, respectively. The obtained background error characteristics share similarities with the ones computed in the midlatitude band, involving strong geostrophic coupling between temperature or surface pressure with streamfunction. This coupling has been found to be weaker over steep slopes near the Antarctic coastline. Moreover, the correlation changes sign in the boundary layer over the continent, with cold errors being related to anticyclonic errors. The correlations between temperature and unbalanced velocity potential also exhibit significant changes in shape and strength over the continent.

The vertical autocovariances also show similar characteristics to what has been obtained at the midlatitude, with increased vertical extension at the tropopause and at lower levels (although this latter effect may be due to increased model resolution in the boundary layer). Standard deviations are probably linked with dynamics, including the storm tracks and synoptic activity over the seas around Antarctica, as well as with boundary layer processes above the continent. We found increased variance over the surrounding seas for streamfunction, upper-level unbalanced velocity potential, and relative humidity errors, and on the contrary increased variance over the continent for unbalanced temperature and lower-level unbalanced velocity potential. Unbalanced surface pressure shows increased variance around the coastlines. The

**Fig. 15.** First three eigenvectors of the EOF decomposition of the vertical background error covariance matrix for control variables over AMPS domain 2. The legend is as in Fig. 10.
FIG. 16. Rescaling factor of variance over AMPS domain 2. The legend is as in Fig. 11, but that the result for EOF 2 is shown for unbalanced velocity potential.
FIG. 17. Local length scales (km) over AMPS domain 2. The legend is as in Fig. 12, but that the result for EOF 2 is shown for unbalanced velocity potential.
diagnosed local length scales also exhibit significant geographical variations. Over the first domain, they seem larger over seas for streamfunction and relative humidity, which is believed to be consistent with geostrophic scaling. Over Antarctica and especially for the second domain, a contrast has been found between the eastern and western parts of the continent. Whereas these findings are probably physical and relevant to mesoscale data assimilation, their precise impacts still have to be assessed.

The underlying goal behind the development of a more flexible background error covariance matrix is to address higher resolutions for mesoscale modeling. We will investigate the impact of the newly developed and estimated background error covariance matrices for weather forecasting over Antarctica. Another topic of interest is that this inhomogeneous modeling could be used to represent the error of the day: the CVT can be estimated on a limited number of ensemble members with suitably perturbed initial conditions at a single date. This may help to relax the ergodic assumption in variational assimilation schemes. An interesting issue would be to compare the nearly raw background error covariance matrix obtained through direct sampling in the ensemble Kalman filter to this filtered matrix that is, however, able to represent some inhomogeneity. This comparison could be achieved using a hybrid assimilation scheme that blends different background errors.

Acknowledgments. This research has been partly supported by Météo-France, CNRM-GAME, and CNRS, by the EOS Interdisciplinary Science Program, and by the NASA Research and Application Program.

REFERENCES


