Introducing Climatological Flow-Dependence in the WRFVAR Background Error Model For Variational Data Assimilation. Application to Antarctica.

YANN MICHEL
National Center for Atmospheric Research, Boulder, CO, USA. yann@ucar.edu

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 grid-point modeling in this study makes use of regression-based balance operators between variables, empirical orthogonal function decomposition to define the vertical correlations and high order efficient recursive filters to impose horizontal correlations. A particularity is that the regression operator 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 domain of interest covers the Antarctica, where the inhomogeneity of background errors are expected to be important due to the particular orography, physics, and contrast between ice, land and sea.

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 over the atmospheric flow. It also may 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).

The Antarctic Mesoscale Prediction System (AMPS; Powers et al. (2003)) has been designed to overcome the difficulties in numerical weather modeling at the poles. AMPS is based on a modified "polar" version of the fifth-generation Pennsylvania State University - National Center for Atmospheric Research numerical Mesoscale Model (MM5). Although developed originally with MM5, AMPS now uses the WRF model (Skamarock et al. 2008)). First guess and lateral boundary conditions are derived from the Global Forecasting System (GFS) developed at the National Center for Environmental Prediction (NCEP). AMPS has been shown to provide relevant meteorological guidance for the forecasting in the Antarctic region over the last few years (Powers et al. 2003). It has been successfully used to study the prediction of some severe synoptic events such as the May 2004 McMurdo storm (Powers 2007; Steinhoff et al. 2008). The AMPS modeling system features currently 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 towards McMurdo (Powers et al. 2003), and has currently 45 km resolution. The second domain covers the whole Antarctic continent with an improved spatial resolution of 15 km. Other domains include local areas near the Ross Ice shelf and McMurdo, or near the South Pole around Amundsen-Scott (Powers et al. 2003).

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 achieve through advanced data assimilation schemes that are able to cope with the special properties of background errors over Antarctica, and that are related to the sparsity of observations and the physical properties briefly described above. Two main schemes are available for the AMPS model, namely an ensemble square root filter (Barker 2005) and a three-dimensional variational assimilation (Barker et al. 2004). In this 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 data assimilation ensembles with perturbed observations (Pereira and Berre 2006). The specification of the background error covariance matrix $\mathbf{B}$ is modeled as a sequence of operators using control variable transforms (CVT) (Derber and Bouttier 1999). This is has several advantages in addition to reducing the dimension of $\mathbf{B}$, namely ensuring physical balance constraints and improving the conditioning of the minimization (Courtier and Talagrand 1990).

The goal of this paper is to extend the currently homogeneous WRF scheme to an inhomogeneous formulation in order to make benefit of the grid space formulation of the CVT (Barker et al. 2004). Inhomogeneity in the horizontal correlations is incorporated through the use of inhomogeneous recursive filters Purser et al. (2003). This is illustrated with computations of the background error statistics over the Antarctica.

2. Background error modeling in WRF

a. The Control Variable Transform in 3D-Var

In general, variational assimilation schemes are designed to provide an analysis $\mathbf{x}_a$ that minimizes a cost function $J(\mathbf{x})$:

$$\mathbf{x}_a = \text{Arg min } J \tag{1}$$

$$J(\mathbf{x}) = \frac{1}{2}\big(\mathbf{x}-\mathbf{x}_b\big)^{\mathrm{T}}\mathbf{B}^{-1}\big(\mathbf{x}-\mathbf{x}_b\big) \tag{2}$$

$$+\frac{1}{2}\big(\mathbf{y}-\mathcal{H}(\mathbf{x})\big)^{\mathrm{T}}\mathbf{R}^{-1}\big(\mathbf{y}-\mathcal{H}(\mathbf{x})\big) \tag{3}$$

where, denoting by n the dimension of x and by p the dimension of the vector of observations $\mathbf{y}$:

- $\mathcal{H}$ is the non-linear observation operator;
- $\mathbf{H}$ of dimensions $p \times n$ is the linearized observation operator;
- $\mathbf{B}$ of dimensions $n \times n$ is the background error covariance matrix;
- $\mathbf{R}$ of dimensions $p \times p$ is the observation error covariance matrix.

The analysis is a weighted average of the background and of the observations, and the weights depend on the covariances of their respective errors. The specification of $\mathbf{B}$ is achieved through the change of variable (CVT)

$$\mathbf{v} = \mathbf{B}^{1/2}\chi \tag{4}$$

The choice made for the WRF 3D-Var (Barker et al. 2004) was adapted from the UK-MetOffice Control Variable Transform (Lorenc et al. 2000) but the horizontal correlations were prescribed through homogeneous recursive filters (Hayden and Purser 1995) rather than through a spectral space decomposition. Our update over the formulation of Barker et al. (2004) includes the multiplication of a variance rescaling factor and the use of spatially inhomogeneous recursive filters following Purser et al. (2003). The new CVT may be written

$$\mathbf{v} = \mathbf{U}_p\mathbf{U}_v\mathbf{U}_{\text{ih}}\mathbf{S}\chi \tag{5}$$

where $\mathbf{S}$ is a multiplicative variance scaling factor, $\mathbf{U}_{\text{ih}}$ is the application of high order, fully inhomogeneous recursive filters to impose horizontal correlations, $\mathbf{U}_v$ is the application of vertical correlations through EOF and $\mathbf{U}_p$ changes the control variables to model variables through physical relationships.

3. Background error statistics over the Antarctica

a. Methodology

The development of an advanced background matrix aims to better represent the features of errors. The Antarctic region provides a challenging environment to design and test new data assimilation techniques (Barker et al. 2004) owing to the sparsity of available observations and to the unusual atmospheric processes occurring (King and Turner 1997). The study of background errors has drawn considerable attention in the community, given the usually large impact that their representation have on NWP performance. Few studies have been devoted to the examination of background error covariances over the Antarctic region. However, the covariances are expected to be specific in this region. For instance, using a global model, Ingleby (2001) reported also a change a sign in the large-scale cross-covariance between temperature and surface pressure. As in Berre (2000), vertical cross-covariances can be computed to infer some physical meaning of the balance relationships. Their statistical significance can be inferred from the explained ratio of variance, obtained through covariances between unbalanced and total variables.

b. *Inhomogeneous balance operators: an illustration*

The second 15-km resolution AMPS domain is used to compute the balance operators. As a first step, the inhomogeneity of the balance will be shown by computing the statistics with geographical masks. The inner AMPS domain (Fig. 1) is split between the oceanic and continental parts to look for possible differences. The coupling between temperature and streamfunction and unbalanced velocity potential is depicted in Fig. 2. The geostrophic coupling (panels a and b) is very similar in the troposphere, but of opposite sign in the boundary layer. This can be explained through temperature inversions. Differences between temperature and unbalanced velocity potential are even more striking. The low level covariance maximum at the surface over Land might be linked with the frequent occurrence of katabatic winds, whose strength is linked with temperature and topography.

c. *Standard deviations*

Background error standard deviations (or variances) seem clearly linked with the circulation of synoptic systems on the seas around the continent, and are heavily influenced by the contrast between seas and land. For small scale variables (unbalanced temperature and humidity), an increase of variance along the boundary of domain 2 is sometimes visible in domain 1, as a consequence of nesting (not shown). For streamfunction, an increase of variance can be seen towards the circumpolar vortex (panel a in Fig. 3). Humidity shows lower variance over the Antarctic plateau (panel d). Unbalanced velocity potential and unbalanced temperature rather exhibit opposite behavior with increased variance over the Antarctic plateau (panels b and c). For all variables, there seem to be a significant reduction of the variance towards the border that is linked with the larger scale common boundary conditions from the lagged NMC method (Pereira and Berre 2006).

d. *Horizontal lengthscales*

Horizontal lengthscales 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. As the latter authors, we estimate the (local) lengthscales through the ratio of the variance of a field and the variance of the Laplacian of this field. For instance, the correlation lengthscale of streamfunction is estimated *via*:

$$L = \left(8\frac{V(\psi)}{V(\xi)}\right)^{1/4} \tag{6}$$

where ξ is the vorticity (the Laplacian of streamfunction) and is computed through spectral transform, taking into account the map projection factor, and V is the variance over time in the NMC method.

The estimates of unbalanced surface pressure error local lengthscales for AMPS domains 1 and 2 are shown in Fig. 4. Small scale noise is noticeable in the raw estimates (panels a and c), consistent with the findings of Pannekoucke et al. (2008). There is a clear contrast between the continent and the seas, with much larger lengthscales over the Antarctic plateau, especially in the Eastern part. The effect of nesting is also visible, with smaller lengthscales at the inner boundary. Over domain 2, we can also see a strong reduction of the lengthscale along the coastline. It is necessary to filter the lengthscales such that the correlation functions keep a quasi-Gaussian shape. This allows a good amplitude correction and lengthscales representation (Pannekoucke and Massart 2008) as well as the additional constraint, for horizontally separable recursive filters, to keep undesirable grid-related anisotropy to a negligible level. The filter is done through convolution (multiplication in spectral space) and allows the main geographic contrasts to be represented (panels b and d).

4. Conclusion

In data sparse areas, the correct specification of the background error covariance matrix is a key element to spread 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).

We address first the issue of inhomogeneous background error modeling. This is achieved through a local balance transform and inhomogeneous recursive filters, which can be achieved at a reasonable cost thanks to the WRF-Var grid point formulation. The balance uses local regressions in grid point space such as in Wu et al. (2002). The inhomogeneous recursive filters are based on the work by Purser et al. (2003).

The second part of this paper describes the covariances of background error over the Antarctic region for the Antarctic Mesoscale Prediction System. Coupling in the boundary layer is shown to differ over the plateau with respect to over the ocean, which is believed to be due to the frequent occurrence of temperature inversions and katabatic winds over the continent. Otherwise, the obtained background error characteristics share strong similarities with the ones computed in the mid-latitude band, namely geostrophic coupling between temperature and streamfunction.

Variations of 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 unbalanced velocity potential and relative humidity errors, and on the contrary increased variance over the continent for streamfunction and unbalanced temperature. Unbalanced surface pressure shows increased variance around the coastlines. The diagnosed local lengthscales also exhibit significant geographical variations; for instance the surface pressure error lengthscale is multiplied by a factor of 3 between its minimum value over seas and its maximum value over the Antarctic plateau.

REFERENCES

Ball, F., 1956: The theory of strong katabatic winds. *Australian J. Phys.*, **9**, 373–386.

Barker, D., 2005: Southern high-latitude ensemble data assimilation in the Antarctic Mesoscale Prediction System. *Mon. Wea. Rev.*, **133**, 3431–3449.

Barker, D., W. Huang, Y.-R. Guo, A. Bourgeois, and Q. Xiao, 2004: A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Mon. Wea. Rev.*, **132**, 897–914.

Berre, L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Mon. Wea. Rev.*, **138**, 644–667.

Bouttier, F., 1994: A dynamical estimation of forecast error covariances in an assimilation system. *Mon. Wea. Rev.*, **122**, 2376–2390.

Courtier, P. and O. Talagrand, 1990: Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus*, **42**, 531–549.

Derber, J. and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195–221.

Hayden, C. and R. Purser, 1995: Recursive filter objective analysis of meteorological fields: Applications to NESDIS operational processing. *J. Appl. Meteor.*, **34**, 3–15.

Ingleby, N., 2001: The statistical structure of forecast errors and its representation in the Met. Office Global 3-D variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **127**, 209–231.

King, J. and J. Turner, 1997: *Antarctic meteorology and climatology*. Cambridge University Press.

Lorenc, A. C., et al., 2000: The Met Office global 3-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **126**, 2991–3012.

Pannekoucke, O., L. Berre, and G. Desroziers, 2008: Background-error correlation length-scale estimates and their sampling statistics. *Quart. J. Roy. Meteor. Soc.*, **134**, 497–508.

Pannekoucke, O. and S. Massart, 2008: Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. *Quart. J. Roy. Meteor. Soc.*, **134**, 1425–1438.

Parish, T. and J. Cassano, 2003: The role of katabatic winds on the antarctic surface wind regime. *Mon. Wea. Rev.*, **131**, 317–333.

Parrish, D. and J. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Pereira, M. and L. Berre, 2006: The use of an ensemble approach to study the background error covariances in a global NWP. *Mon. Wea. Rev.*, **134**, 2466–2489.

Pettré, P., M. F. Renaud, R. Renaud, M. Déqué, S. Planton, and J. C. André, 1990: Study of the influence of katabatic flows on the antarctic circulation using GCM simulations. *Meteorology and Atmospheric Physics*, **43**, 187–195.

Powers, J., 2007: Numerical prediction of an antarctic severe wind event with the Weather Research and Forecasting (WRF) model. *Mon. Wea. Rev.*, **135**, 3134–3157.

Powers, J., A. Monaghan, A. Cayette, D. Bromwich, Y.-H. Kuo, and K. Manning, 2003: Real-time mesoscale modeling over Antarctica: The Antarctic Mesoscale Prediction System. *Bull. Amer. Meteor. Soc.*, **84**, 1533–1545.

Purser, R., W. Wu, D. Parrish, and N. Roberts, 2003: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536–1548.

Skamarock, W., et al., 2008: A description of the advanced research WRF version 3. *NCAR Technical Note.*

Steinhoff, D., D. Bromwich, M. Lambertson, S. Knuth, and M. Lazzara, 2008: A dynamical investigation of the may 2004 McMurdo Antarctica severe wind event using AMPS. *Mon. Wea. Rev.*, **136**, 7–26.

Wu, W., R. Purser, and D. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916.

Xiao, Q., et al., 2008: Application of an adiabatic WRF adjoint to the investigation of the May 2004 McMurdo, Antarctica, severe wind event. *Mon. Wea. Rev.*, **136**, 3696–3713.

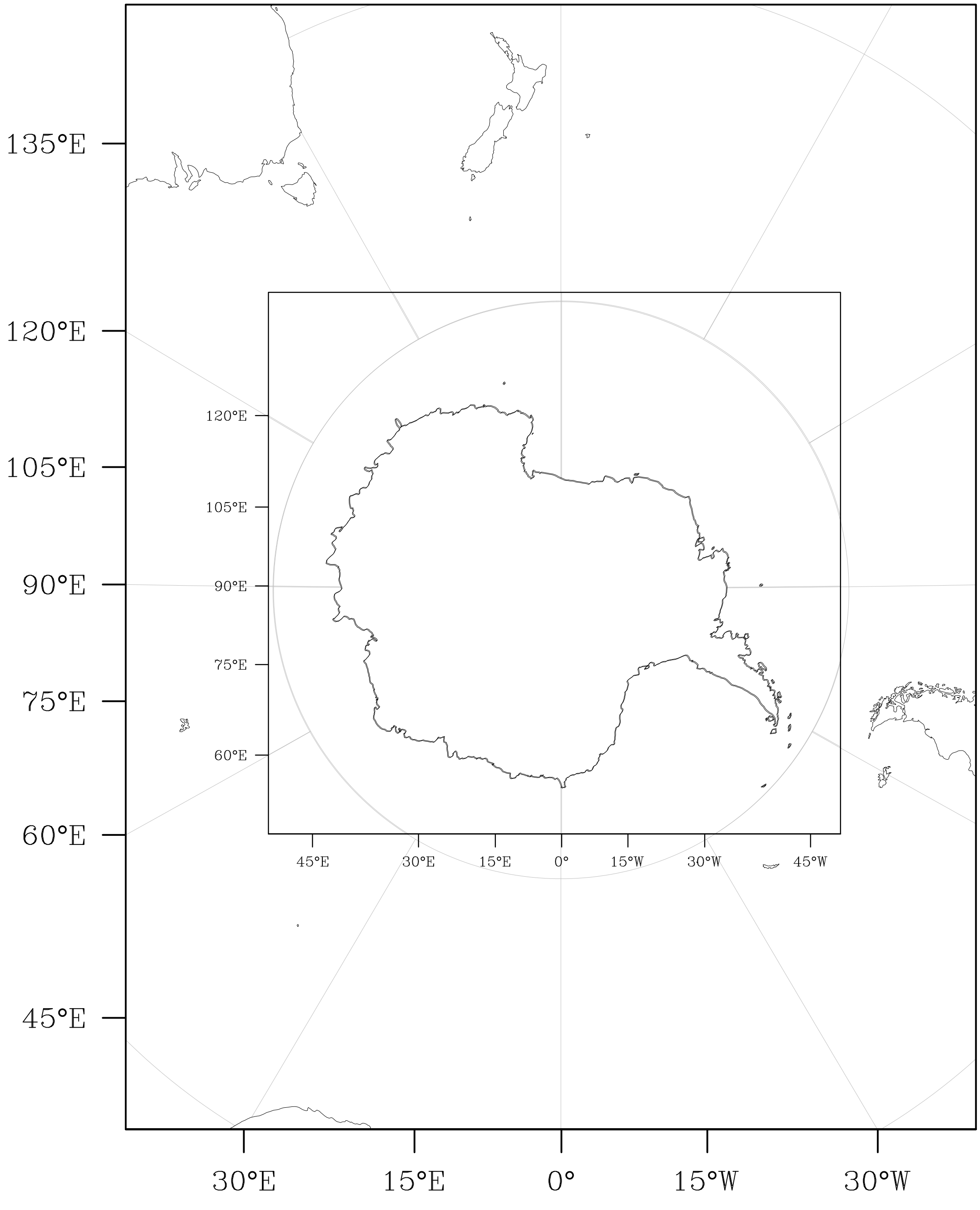

FIG. 1. AMPS domains for the 45 km resolution configuration (outer box) and the nested 15 km resolution configuration (inner box)

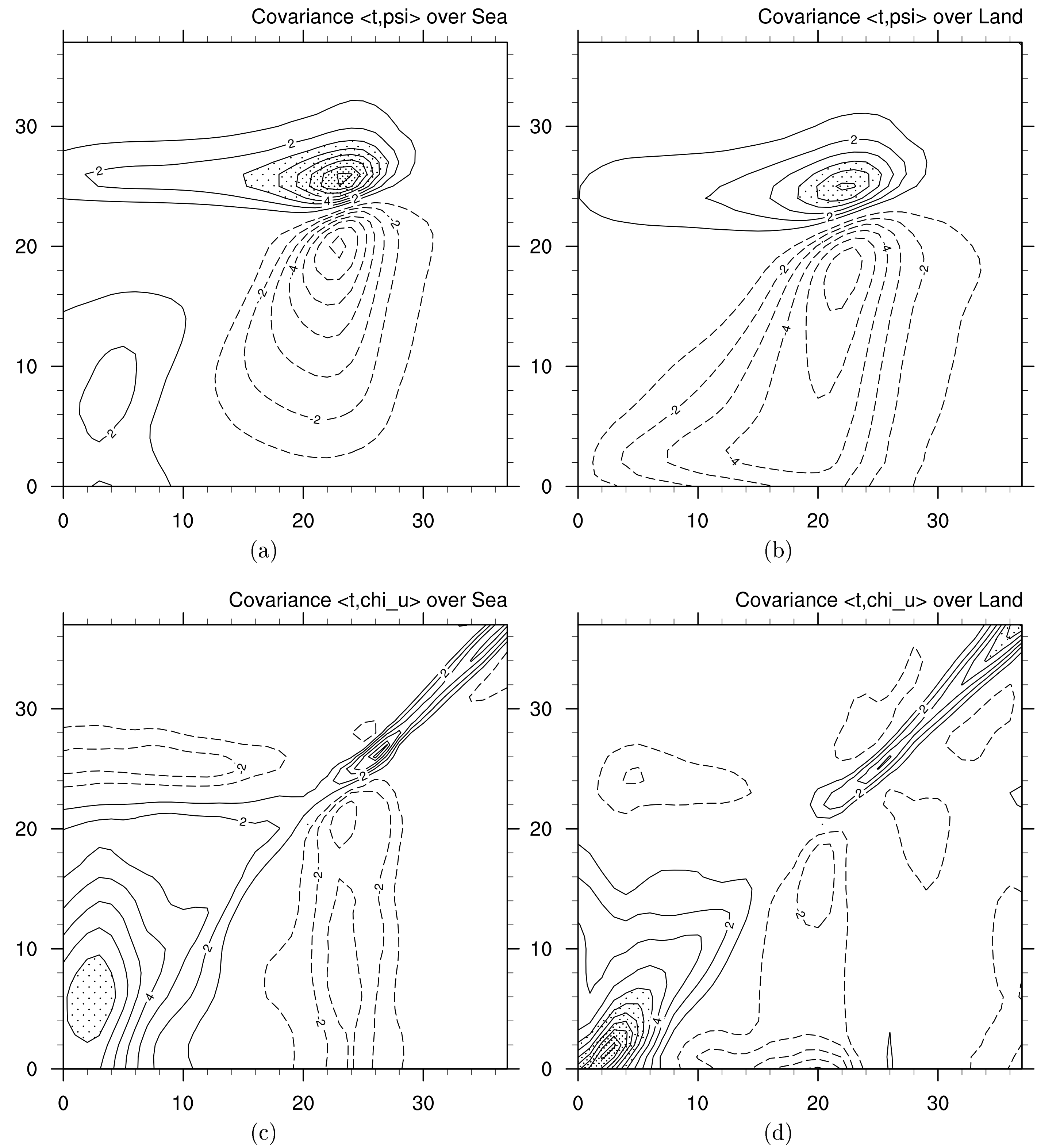

FIG. 2. Diagnosed vertical covariances between temperature and streamfunction errors (panel a and b, units: $10^5\mathrm{m}^2\ \mathrm{s}^{-2}$ K) and velocity potential errors (panel c and d, units: $10^4\mathrm{m}^2\ \mathrm{s}^{-2}$ K) drawn with solid (dashed) lines for positive (negative) values.

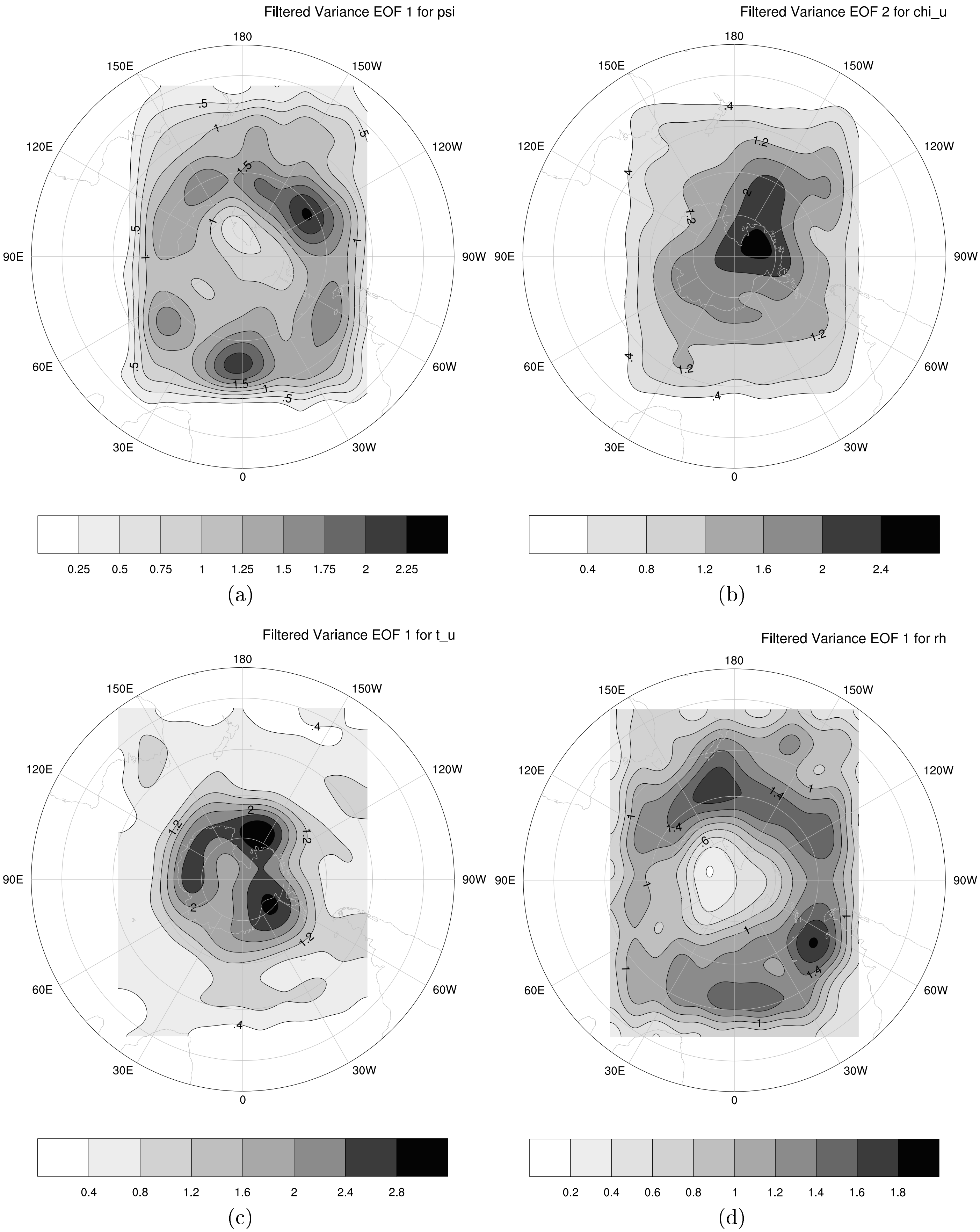

FIG. 3. Rescaling factor of variance for streamfunction (EOF1, panel a), unbalanced velocity potential (EOF 2, panel b), unbalanced temperature (EOF1, panel c) and relative humidity (EOF 1, panel d).

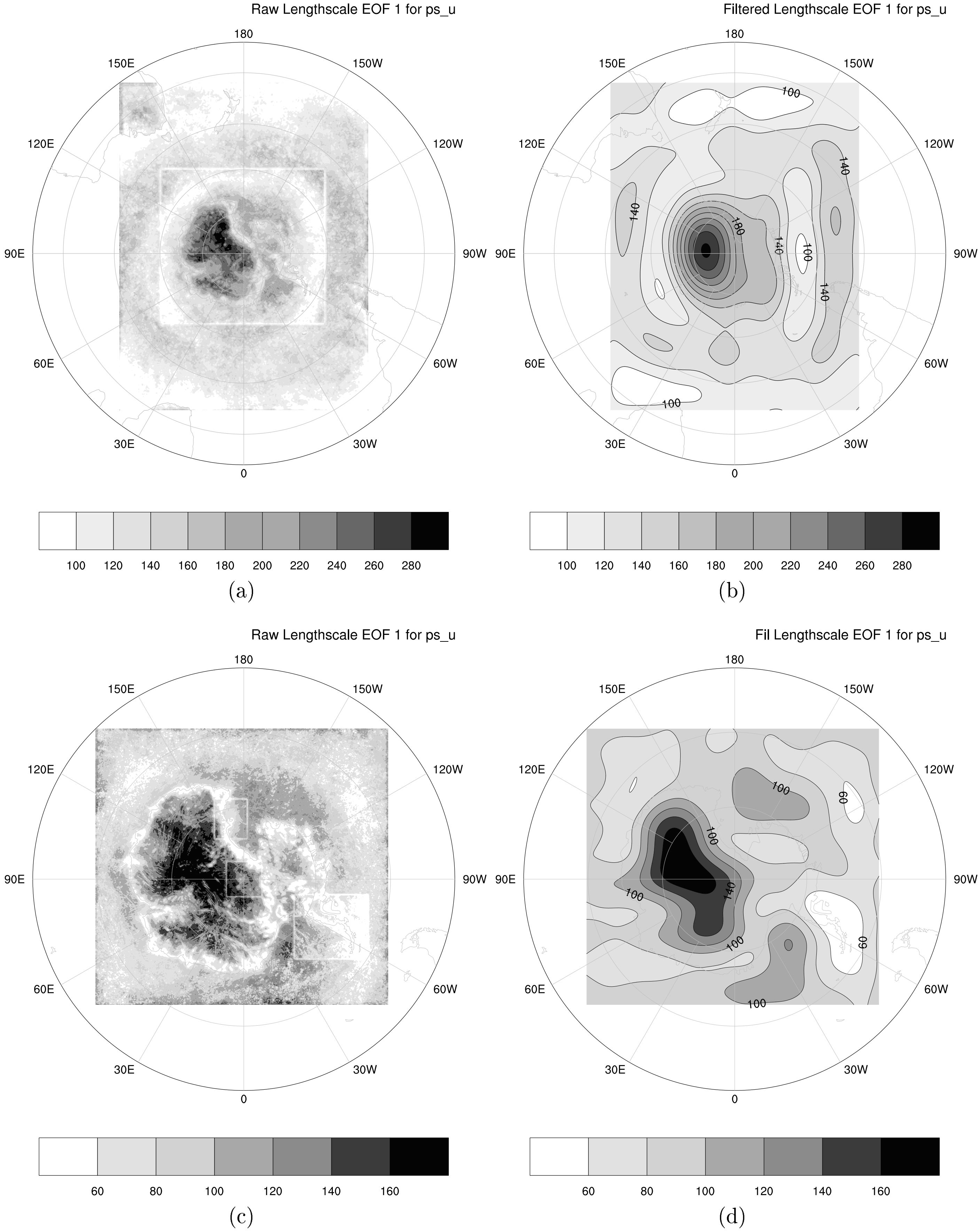

FIG. 4. Local lengthscales (km) of unbalanced surface pressure background error for domain 1 (upper panels) and 2 (lower panels). Raw (spatially filtered) estimates are shown in left (right) panels.