System of 3D Geopotential Height, Temperature and Wind Correlations with Respect to Forecast Errors[1]

Oleg A. Alduchov[2], Vladimir A. Gordin[3]

The correlation model, proposed by [1-5], was applied for 3-years time series of forecast errors (differences observation-forecast) with respect to the forecast fields used in Hydrometeorological Center of Russia, Moscow. The subset of 843 (from available 1055 during three years) upper-air stations over globe were used for calculations. Even three years is too short period for reliable evaluations of correlations, some very preliminary results may be obtained.

The correlations were calculated under following conditions:

a) all differences were normalized (see our previous article here) by local (for each upper-air station) mean and sigma:

$$f(\vec{x},p)=\frac{F(\vec{x},p)-\overline{F}(\vec{x},p)}{\sigma_{F(\vec{x},p)}};$$

b) for each 30-degrees latitudinal zone the fields of such normalized forecast errors are considered as homogeneous and isotropic with respect to horizontal coordinates;

c) it was assumed (because of too short period of accumulated time series) that normalized forecast errors are homogeneous during whole three years (without months partitioning).

The full set of auto- and cross-correlations $\{\mu(r_k,P_i,P_j)\}$ for geopotential height (*H*), temperature (*T*), transversal (*N*) and longitudinal (*L*) wind's components forecast errors at *15* standard pressure levels from *925* to *10* hPa and for *50*-km distance intervals from *0* to *3000* km were calculated.

Resulted correlations were approximated by expressions

$$\mu(r_k,p_i,p_j)=a_0(p_i,p_j)\cdot\delta(r_k)+\sum_{l=1}^{N}a_l(p_i,p_j)\cdot\mathbf{J_0}(r_k/d_l) \qquad (1)$$

where $\delta(r_k)$ is the Dirac-function equal to *0* everywhere except point $r=0$, where it is equal to 1; $\mathbf{J_0}$ - Bessel-function; $\mathbf{A}_l=\{a_l(P_i,P_j)\}, l=0,...,N$ are positively defined matrices of coefficients to be defined; $d_l, l=1,...,N$ are some scale coefficients to be defined too. (*N=4* for the current case).

Being defined as result of the task solution, the positively defined matrixes $\mathbf{A}_l$ guarantee the positively definiteness of approximated functions μ and its closeness to original coefficients μ. So, RMS differences between originally calculated correlations and their positively defined approximations are about *0.01*. Again, approximation (1) allows distinguishing between local errors (including observation errors and local field disturbances) and macro scale errors by very simple technique under reasonable assumptions.

Some cross-sections of approximated correlations of forecast errors with comparison to original coefficients fore middle latitudes of Northern hemisphere are shown on the Figure. Correlations *HxL, TxL* and *NxL* are theoretically equal to 0. It is obvious that correlations are significantly varying for different levels. It is easy to see also that correlations for stratospheric levels (above *100* hPa level) are much more significant comparing troposphere levels. It means that forecast into troposphere is much more close to really observed values against the stratosphere. Comparison with correlations of observed upper-air parameters [3,5], shows that correlation radius for forecast errors is remarkable smaller.

1. Alduchov O.A., Gordin V.A. (AG): *3-D Correlation Functions of Upper-Air Parameters.* Research Activities in Atmospheric and Oceanic Modeling. 1998, N 27/865, pp.1.1-1.2.
2. AG: *3D Correlation Functions - Variational Problem.* Ibid, N 28, 1999, pp.1.1-1.2.
3. AG: *Complete system of 3-D geop. height, temperature and wind correlations.* Ibid, N 29, 2000, pp.1.1-1.2.
4. Gordin V.A.: *Mathematical Problems and Methods in Hydrodynamic Weather Forecasting.* Gordon Breach Publ. House, 2000, 842p.
5. AG: *3-Dimentional Correlation Functions of Basic Upper-Air Parameters.* , 2001, Izvestia of Russian Academy of Sciences. ser. ``The Physics of Atmosphere and Ocean'', 37(1), pp.3-23.

[1] The work was partly supported by the grant 01-05-64748.

[2] Russian Inst. of Hydrometeorological Information-Word Data Center, Obninsk, Russia, aoa@wdc.meteo.ru
[3] Hydrometeorological Center of Russia, Moscow, Russia, vagordin@vagordin.mccme.ru

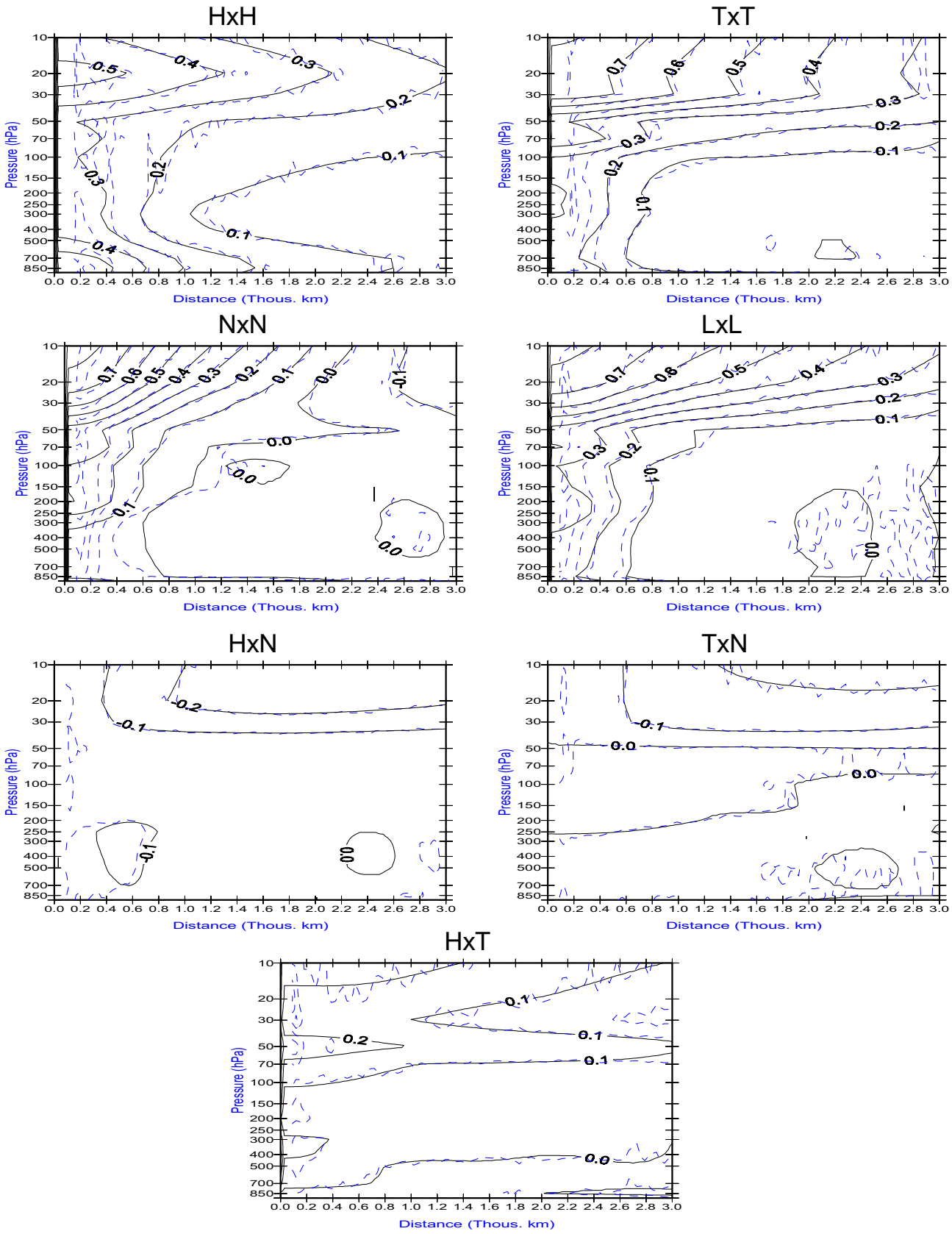


Figure. Auto- (HxH, TxT, NxN and LxL) and cross- (HxN, TxN and HxT) correlation functions for 30-60 degree latitudes of Northern hemisphere, 1998-2000 years.
Dashed line - originally calculated correlations, solid line - approximated.

Bias, Variability, and Hydrostatic Adjustment for 3D Forecast Fields of Geopotential Height, Temperature and Wind[1]

Oleg A. Alduchov[2], Alexei N. Bagrov[3], Vladimir A. Gordin

Forecast meteorological fields are used as first guess (FG) for objective analysis (OA) procedures. Our verification of the fields shows that they are not fulfilled to hydrostatic equation:

$$\frac{dH}{d\xi} = T, \quad \text{where} \quad \xi = -\frac{R}{g} \cdot \ln p \qquad (1)$$

The differential equation can be discretized (by cubic splines) for 16 mandatory levels as

$$\frac{3 \cdot H_{i-1}}{h_{i-1}^2} - \frac{3 \cdot H_i}{h_{i-1}^2} + \frac{3 \cdot H_i}{h_i^2} - \frac{3 \cdot H_{i+1}}{h_i^2} + \frac{T_{i-1}}{h_{i-1}} + \frac{2 \cdot T_i}{h_{i-1}} + \frac{2 \cdot T_i}{h_i} + \frac{T_{i+1}}{h_i} = 0, \quad i = 2,...,15, \qquad (2)$$

where $h_i = \xi_{i+1} - \xi_i$, $i = 1,...,15$. To damp the residuals of the hydrostatic equation we make a projection (see [3]) of any twin vertical profile of the 32D-field $\left\langle \{H_j\}_{j=1}^{j=16}, \{T_j\}_{j=1}^{j=16} \right\rangle$ onto the 18 dimensional subspace composed from all vectors that satisfy to (2).

Evaluations over four years of three-dimensional fields of bias and variability of the difference between some observed upper-air values and forecast fields (Bracknell's) are shown on **Fig.1-5.** The variability here i) is smaller the climatic one, [2]; ii) strongly depends on horizontal coordinates. The mean values of differences are remarkable large. The application of corresponded corrections to the first guess fields during objective analysis procedure leads to improvements of the resulted objective analysis geopotential height fields [1,3].

1. Alduchov O.A., Bagrov A.N., Gordin V.A.: *Statistical characteristics of forecast meteorological fields and their application for objective analysis. Bias and hydrostatic departures*. Submitted to "Meteorology and Hydrology".
2. Alduchov O.A., Gordin V.A : *3-Dimentional Correlation Functions of Basic Upper-Air Parameters.* , 2001, Izvestia of Russian Academy of Sciences. ser. ``The Physics of Atmosphere and Ocean'', 37(1), pp.3-23.
3. Gordin V.A.: *Mathematical Problems and Methods in Hydrodynamic Weather Forecasting.* Gordon & Breach Publ. House, 2000, 842p.

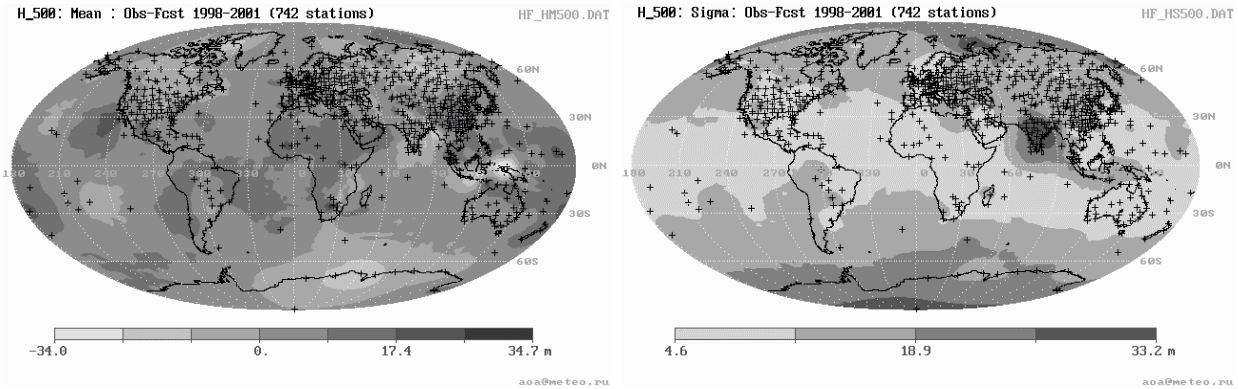


Fig. 1. Mean (left) and variability (right) of differences observations-forecast for geopotential height at level 500 hPa over 1998-2001 years.

[1] The work was partly supported by the grant 01-05-64748.
[2] Russian Inst. of Hydrometeorological Information-Word Data Center, Obninsk, Russia, aoa@wdc.meteo.ru
[3] Hydrometeorological Center of Russia, Moscow, Russia, vagordin@vagordin.mccme.ru

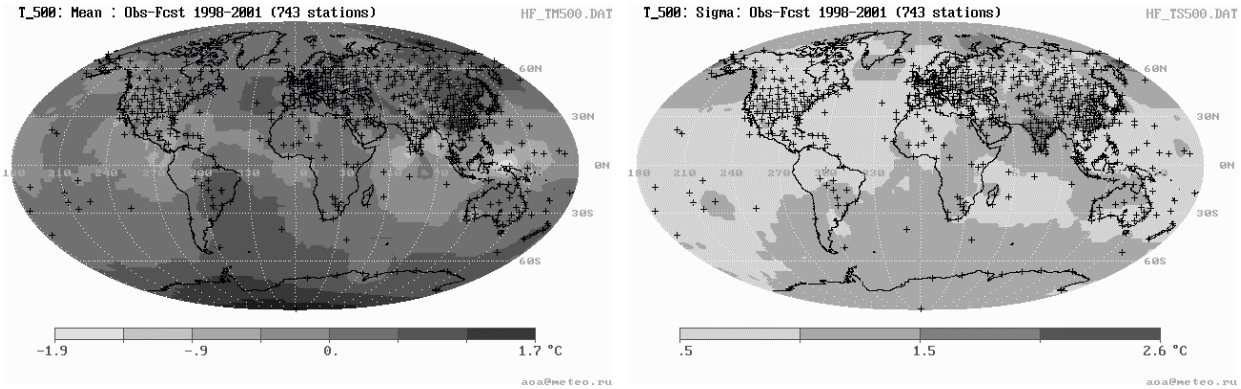


Fig. 2. Mean (left) and variability (right) of differences observations-forecast for temperature at level 500 hPa over 1998-2001 years.

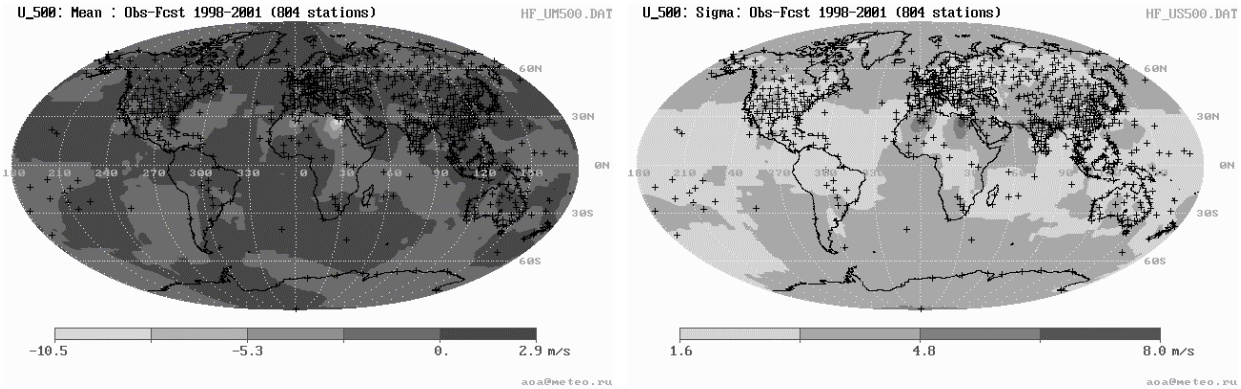


Fig. 3. Mean (left) and variability (right) of differences observations-forecast for zonal wind at level 500 hPa over 1998-2001 years.

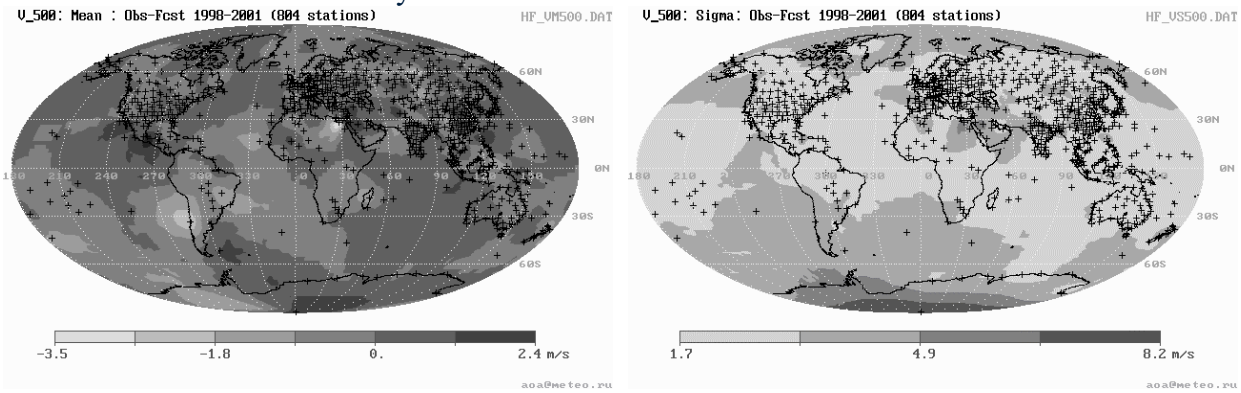


Fig. 4. Mean (left) and variability (right) of differences observations-forecast for meridional wind at level 500 hPa over 1998-2001 years.

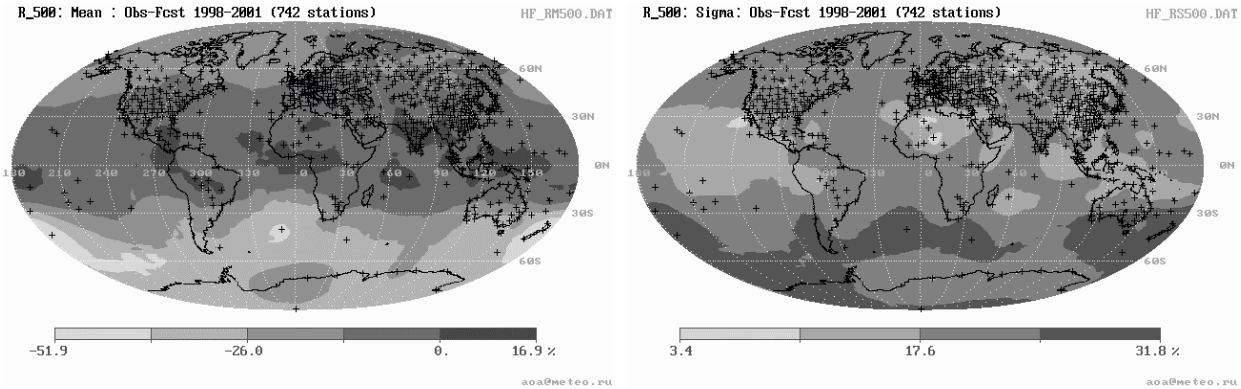


Fig. 5. Mean (left) and variability (right) of differences observations-forecast for relative humidity at level 500 hPa over 1998-2001 years.

Impact of ATOVS AMSU-A radiance data in the DMI-HIRLAM 3D-Var analysis and forecasting system

Bjarne Amstrup
Danish Meteorological Institute
Lyngbyvej 100, DK-2100 Copenhagen Ø
Denmark
Email: bja@dmi.dk

Since May 2001 an ongoing observing system experiment (OSE) using Advanced TIROS (Television Infra-Red Observation Satellite) Operational Vertical Sounder (ATOVS) brightness temperatures from the polar orbiting satellite NOAA16 (National Oceanic and Atmospheric Administration) in near real-time has been run at DMI (Danish Meteorological Institute). All AMSU-A (Advanced Microwave Sounding Unit-A) level 1c data (channels 1 to 10, only) available from locally received data from the DMI Smidsbjerg antenna as well as from the DMI Sdr. Strømfjord/Kangerlussuaq (Greenland) antenna have been used.

Initially an observation error covariance matrix was calculated using data from a two month period with passive inclusion of NOAA16 data received via the local DMI Smidsbjerg antenna. Subsequently, new observation error covariance matrices have been calculated using data from later one month periods including data from the DMI Sdr. Strømfjord antenna.

The setup used for the first impact studies from June through September 2001 as well as some results can be found in Amstrup, 2001. Further details concerning the forecast model can be found in Sass *et al.*, 2002. The basic model is the DMI-HIRLAM-G (DMI HIgh Resolution Limited Area Model - G) for which the horizontal resolution is 0.45°, the number of vertical levels 31, the number of grid points is 190×202, the time step is 240 s, and the lateral boundary values are from ECMWF forecasts.

Based on observation (obs-) and field verification the impact is basically neutral in this period. There is a marginal positive impact on some parameters based on obs-verification.

Subsequently, some changes were made, including modifications in the observation covariance matrix used. Results for December 2001 and January 2002 can be found in Schyberg *et al.*, 2002. The results for December were neutral, but for January a positive impact was found. An illustration of this is shown in terms of observation verification results in Figure 1. In this figure WIA denotes the run with ATOVS AMSU-A data included and REF denotes the reference run.

The plans at DMI involve usage of ATOVS AMSU-A data in the operational assimilation system in the autumn of 2002 unless unforeseen problems are revealed during the continuous pre-operational runs.

References

Amstrup, Bjarne. 2001. *Impact of ATOVS AMSU-A radiance data in the DMI-HIRLAM 3D-Var analysis and forecasting system.* DMI Scientific Report 01-06. Danish Meteorological Institute.

Sass, Bent Hansen, Nielsen, Niels Woetmann, Jørgensen, Jess U., Amstrup, Bjarne and Kmit, Maryanne. 2002. *The Operational DMI-HIRLAM System - 2002 version.* DMI Technical Report 02-5. Danish Meteorological Institute.

Schyberg, Harald, Landelius, Tomas, Thorsteinsson, Sigurdur, Tveter, Frank Thomas, Vignes, Ole, Amstrup, Bjarne, Gustafsson, Nils, Järvinen, Heikki and Lindskog, Magnus. 2002. Assimilation of ATOVS data in the HIRLAM 3D-Var System. *HIRLAM Technical Report*, **to be published**.

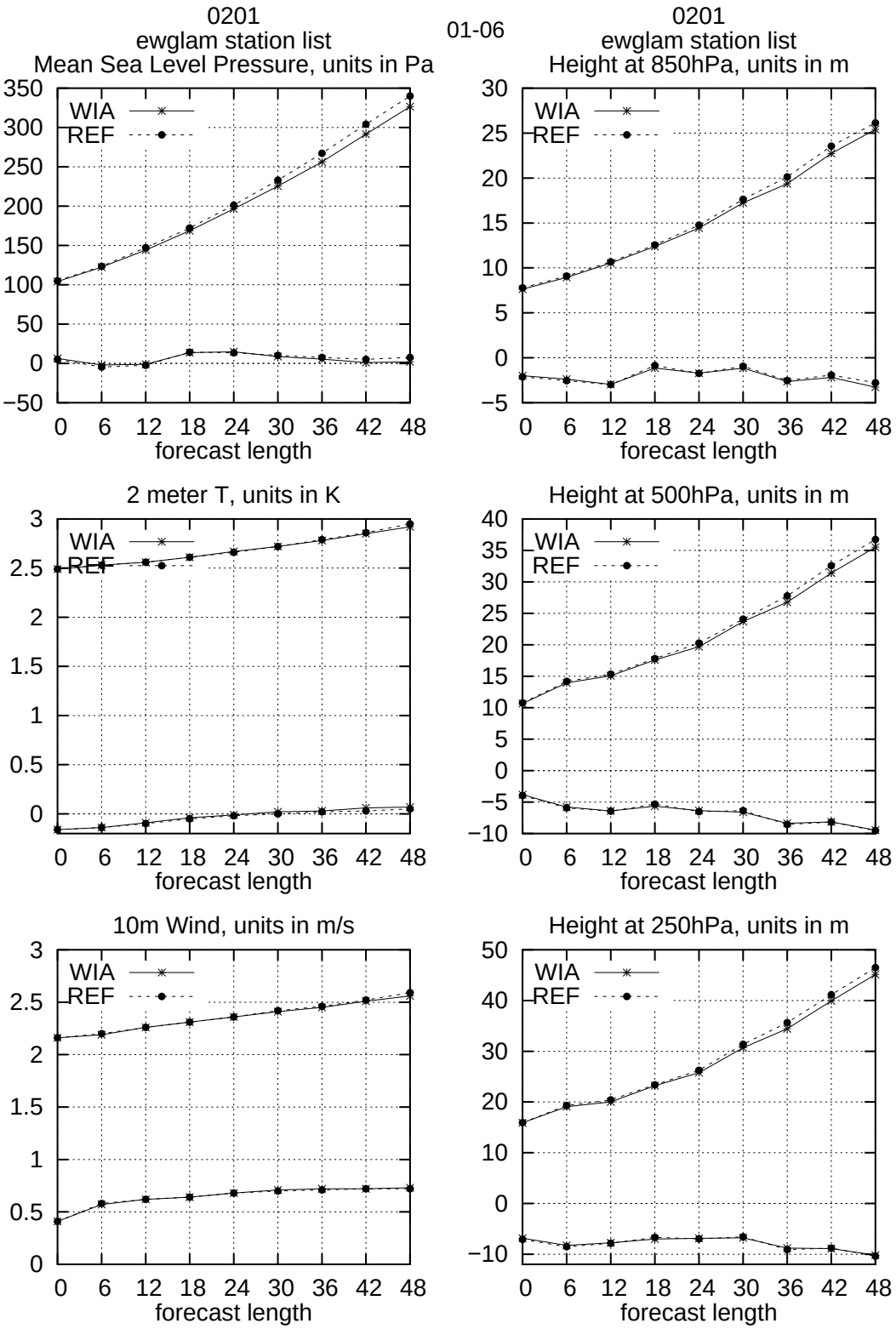


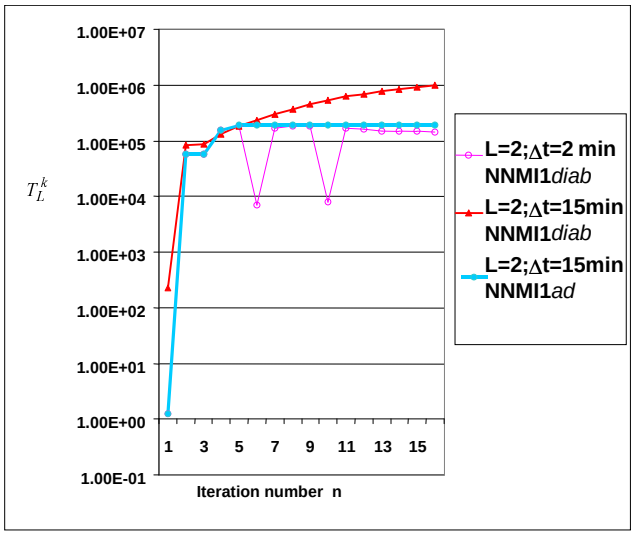
Figure 1: Observation verification results of surface parameters (left) and geopotential heights (rigt) for January 2002 for WIA and REF of surface and upper level parameters as specified in the plot. EWGLAM station list. The upper curves are root mean square errors and the lower curves are bias.

Damping of fast growing gravity waves generated by imbalance of initial mass and wind fields and diabatic forcing

Astakhova E.D., Frolov A.V., Tsvetkov V.I. (E-mail: tsvetkov@rhmc.mecom.ru)

Hydrometeorological Center of Russia, 9-13 Bolshoy Predtechenskii per., 123242, Moscow, Russia

Numerical solutions of the equations of atmospheric hydro- and thermodynamics contain high-frequency noise, which arises, firstly, because of the inconsistency of initial meteorological fields and, secondly, due to diabatic forcing. To damp the noise, nonlinear normal mode initialization (NNMI) procedures (Machenhauer, 1982) are widely used now. They assume that time derivatives of the amplitudes of "fast" modes with frequencies above some given threshold must be set zero at the initial time moment. Then we come to nonlinear equations solved by iterations. However, to ensure a convergence of this iterative process quite strong conditions must be satisfied, e.g., only of a limited number of vertical modes can be initialized or time-averaged diabatic fluxes must be used in a model. At the same time, the diabatic processes can affect the divergent component of the atmospheric circulation (vertical motions) considerably and increase the effect of gravitational waves onto Rossby components in the numerical solution of the primitive equation diabatic model. Thus, the diabatic forcing must be taken into account in order to obtain the correctly balanced initial fields. With this purpose we developed a new diabatic nonlinear normal mode initialization procedure based on a specific selection of most rapidly growing gravitational waves. The procedure can be applied to the entire spectrum of gravitational modes with allowance for the real order of magnitude of physical phenomena present in the atmosphere. A series of numerical experiments was performed with the Hydrometcenter of Russia global spectral model T85L31 with different initialization procedures starting from the 1996 ECMWF objective analysis data. A convergence of the iterative process during initialization and coincidence of the initial data fields were estimated with the help of a characteristic T_L^k (*L* is the number of the vertical mode, *k* the iteration number) suggested by Rasch, 1985. T_L^k can be considered as a typical time scale of the gravitational waves corresponding to the L*th* gravitational mode), and if it decreases from iteration to iteration, it means that the iterative process does not converge. It can be assumed that the initial data are balanced well and do not generate high-frequency modes if T_L^k is high enough (~$3\cdot10^4$ c). As to the convergence of the iterative processes, the numerical experiments showed the following (see the Figure, where NNMI0 denotes the experiment with the Machenhauer initialization procedure; NNMI1*ad* and NNMI1*diab* stand for the experiments with new initialization procedure, adiabatic and diabatic versions, respectively). First, .the time step Δt is important for the NNMI1 performance. At small Δt~2 min, NNMI0 and NNMI1*ad* are equally successful in balancing the initial pressure and wind fields (the values of T_L^k are of the same order). In this case, a change of the initialization procedure practically does not affect the forecast skill. At Δt equal to the model step (15 min), the NNMI0 procedure diverges, while NNMI1 improves the balance between wind and pressure initial fields (T_L^k increases) and, hence, the forecast skill. Second, an application of NNMI1*diab* results in a significant growth of T_L^k. The improvement was most striking for the planetary-scale modes (L=1, 2). Third, the best results were obtained by NNMI1*diab* for 5-7 initialized modes, and from 15 to 20 iterations were necessary to converge the process.



Variations of T_L^k in the iterative processes

The numerical experiments with T85L31 model and NNMI1 demonstrate that the most intensive divergent eddies are usually simulated in the regions with great orographic slopes or in zones of intense convective mixing. The adiabatic initialization, on average, leads to weaker divergent circulation within the tropical and baroclinic mid-latitude zones in the vicinity of large sources of convective heating. It retains main features of the atmospheric circulation above the mountains without considerable decrease of its intensity. In its turn, the diabatic initialization gives divergent circulation components that are closer to the actual ones in the regions with intense diabatic fluxes.

An application of the proposed initialization procedure ensures smooth numerical solutions of the model with minimal changes of initial fields, especially, the divergent components of the atmospheric circulation. Important, that the introduction of diabatic factors, firstly, convective heating, into initialization procedure gives a noticeable improvement of initial field balance (as compared to that obtained with the adiabatic initialization) within mid-latitude baroclinic zones.

The study was financially supported by the Russian Foundation for the Basic Research, projects 00-05-64803, 01-05-65493, 01-05-65400.

References

Machenhauer B., 1982. Fundamentals of nonlinear normal mode initialization.- In: The interaction between objective analysis and initialization/R.Daley, J. Derom, D. Williamson, Eds., NCAR Tech. Note, NCAR/TN – 204 +PROC.

Rasch P.J.,1985. Development in normal initialization. Part 1: A simple interpretation for normal mode initialization, Mon. Wea. Rev., v. 113, No.10.

Impact of QuikSCAT in the GEOS GCM

Robert Atlas*, E. Brin, S.C. Bloom, J. Ardizzone, J. Terry, J.C. Jusem, and D. Bungato
Data Assimilation Office
NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771

The evaluation of QuikSCAT data at the DAO consisted of both subjective and objective comparisons of QuikSCAT winds to ship and buoy observations, GEOS and NCEP wind analyses, ERS–2 wind vectors, and SSM/I wind speeds. This was then followed by a series of data assimilation and forecast experiments using the GEOS DAS. The experiments were aimed at comparing the impact of QuikSCAT with that previously obtained with NSCAT, and assessing the relative utility of QuikSCAT, SSM/I, and ERS–2 winds, the relative contributions of QuikSCAT directional and speed information, and the effectiveness of the QuikSCAT ambiguity removal algorithms.

A Control assimilation was generated using all available data with the exception of satellite surface winds. Then assimilations were generated that added either SSM/I wind speeds, QuikSCAT wind speeds, ERS–2 unique wind vectors, QuikSCAT ambiguous wind vectors, QuikSCAT unique wind vectors, or the combination of QuikSCAT with ERS–2 and SSM/I. The results of this initial evaluation of QuikSCAT demonstrated potential for QuikSCAT data to improve meteorological analyses and forecasts, but also indicated ambiguity removal, and rain contamination problems that were limiting the application of QuikSCAT winds to data assimilation.

As an illustration of the impact of QuikSCAT data, Figure 1 shows anomaly correlations for a limited sample of GEOS–3 Control and QuikSCAT $500\ hPa$ height forecasts for the Northern Hemisphere extratropics and Southern Hemisphere extratropics. From this figure, it can be seen that there is a slight positive impact of QuikSCAT in the Northern Hemisphere extratropics and a larger positve impact in the Southern Hemisphere extratropics using this DAS.

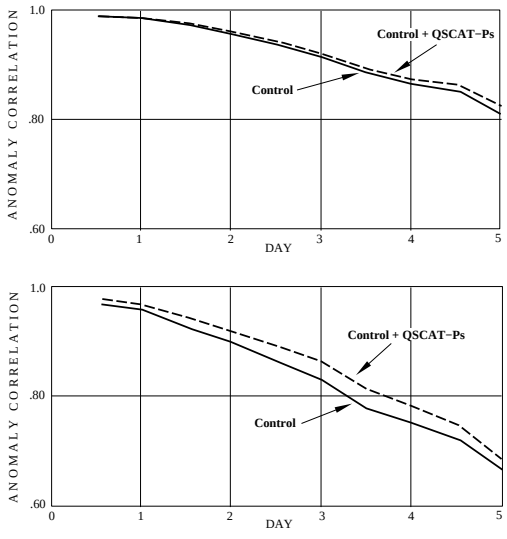


Figure 1 Relative impact of QuikSCAT data on GEOS–3 model forecasts. The $500\ hPa$ geopotential height anomaly correlations, averaged over four forecasts are shown for the (top) Northern Hemisphere extratropics and (bottom) Southern Hemisphere extratropics.

*e-mail: ratlas@dao.gsfc.nasa.gov.

Computation of background error statistics of a double nested limited area model

Gergely Bölöni and András Horányi

Hungarian Meteorological Service, H-1525 Budapest, P.O. Box 38, Hungary
e-mail:boloni.g@met.hu, horanyi.a@met.hu

The presented paper will summarize the work on background error statistics to be used in the data assimilation system of the limited area model (LAM) ALADIN (Horányi et al., 1996). Beside a brief overview of the general developments, we would like to focus on the work performed with the Hungarian domain of the ALADIN model (ALADIN/HU[1]).

Computation of background error covariances is required in the context of data assimilation in order to estimate the solidity of first guess information obtained by an earlier model forecast. The so called NMC method is a widely used algorithm for the definition of background errors in numerical weather prediction models (Parrish and Derber, 1992). This method estimates the background errors as the deviation of model forecasts with different forecast ranges (usually 36 and 12 hours) but valid for the same time. The background error statistics are computed thus on forecast differences accumulated for a few months (typically 3). As the straightforward application of the method (originally used in global models) didn't prove to be convenient to provide mesoscale analysis in the ALADIN model, the so called lagged-NMC method was developed (Široká et al., 2001). The lagged-NMC method is very similar to the original one (standard-NMC) described above with the modification that the lateral boundary conditions are exactly the same for the two forecasts taking part in the computation of the departures. The background error statistics computed using the lagged method put indeed the emphasis on the mesoscale features in the analysis through an error variance maxima shifted towards small scales and through sharper spatial structure functions.

Recently an extensive study of lagged background error statistics was performed in Budapest at the Hungarian Meteorological Service, exploring the sensitivity of the statistics to the forecast range. It means that not only 36h-12h, but all possible combination of forecast differences were created and accumulated as a base of the statistics. While the model is integrated until 48 hours the forecast range can vary between 6 and 48 hours and the time gap between the subtracted forecasts from 6 to 42 hours. The main goal of this study was to choose the optimal statistics for the 3D-VAR scheme used in ALADIN and also to obtain some information about the predictability properties of the model. Many investigations were made in order to digest the various impacts on the statistics while playing with the forecast ranges taking part in the NMC differences. The most important remarks can be summarized as follows: increasing the forecast ranges but keeping a constant time gap, both the total error variance reduction and the shift of the error variance maxima towards smaller scales is enlarging (illustrated on fig.2). Both phenomena are due to the rising number of common lateral boundary conditions that results an increasing loss of information in the context of forecast differences. While an overshot error variance reduction leads to unsatisfactory spatial structure functions, one has to be careful when choosing the optimal background statistics. A compromise must be done to have the most shifted variance maxima still keeping a necessary amount of the total error variance. In order to make an optimal choice of the total variance reduction, 3D-VAR single observation experiments were run with the full set of statistics. Surprisingly, the experiments produced very weak analysis increments compared to the standard-NMC statistics, even using those lagged-statistics with highest error variances. On fig.1 one can compare the impact of the lagged method in case of ALADIN/HU and ALADIN/LACE[2] models (which is the driving model of ALADIN/HU). We would like to point on the reduction of the absolute values of analysis increments due to the lagged method, which is around 6 times stronger in case of the Hungarian domain. This new result is thought to be the consequence of the double nesting which is in fact done in the case of ALADIN/HU model. Namely, it is supposed that the small

[1] A version of the ALADIN model covering the Carpathian basin with $1600km * 1152km$ domain size and $\delta x = 8km$ resolution.

[2] A version of the ALADIN model covering Central-Europe with $2922km * 2630km$ domain size and $\delta x = 12.2km$ resolution.

difference in the resolution and geometry of ALADIN/HU and its driving model ALADIN/LACE allows a strong influence even on smaller scales of the previous through the lateral boundary conditions coming from the latter. This strong influence was recently supported by a parallel verification of the two model versions leading to very similar scores. In the context of lagged-NMC forecast differences it means a loss of information not only on larger, but on small scales as well, resulting in a reduction even on the small-scale part of the error variance spectra. This small-scale variance reduction, appearing also on fig.2, was not found in the case of ALADIN/LACE model (Široká et al., 2001) probably because of more emphasized differences between its resolution and the resolution of its driving model, ARPEGE[3].

Plans for the future are to obtain more information on the applicability of lagged-NMC statistics in the double nested model and the tuning of a scalar factor in front of the background error cost function, which is a possibility to avoid the unwanted impacts of exaggerated total error variance reduction.

References

[1] Horányi A., Ihász I., Radnóti G., 1996: ARPEGE/ALADIN: A numerical weather prediction model for Central-Europe with the participation of the Hungarian Meteorological Service. *Időjárás*, 100, 277-301

[2] Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Centre's spectral statistical interpolation system. *Mon. Wea. Rev.*, 1747-1763

[3] Široká, M., C. Fischer, V. Cassé and J-F. Geleyn, 2001: The definition of mesoscale selective forecast error covariances for a limited area variational analysis. *Meteor. Atmos. Phys. special issue of the SRNWP Workshop on high-resolution modeling, Offenbach, 25.-27. 10. 1999*, submitted.

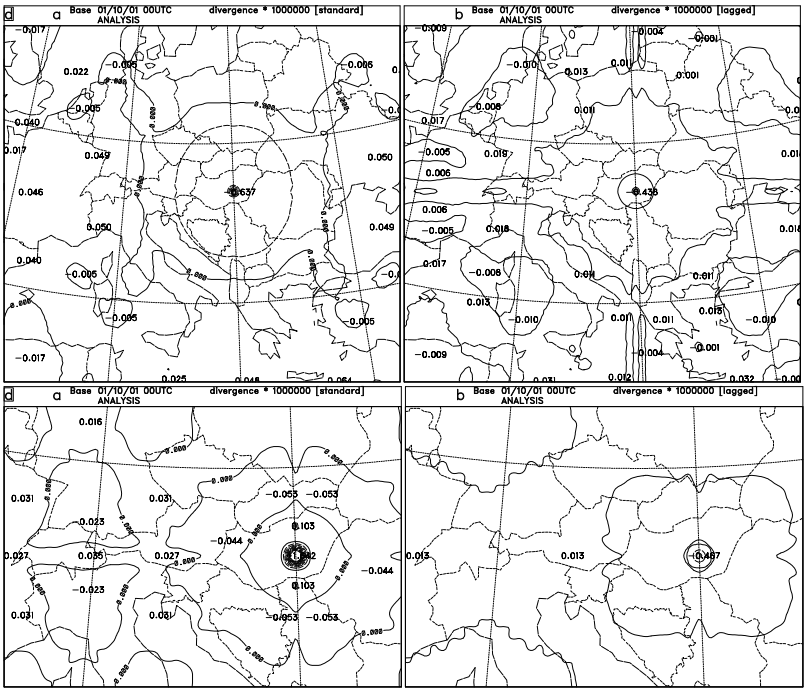


Figure 1: Divergence analysis increment of a single observation experiment ($\delta T = 1°$) on model level 16. Top-left: ALADIN/LACE with standard statistics ($min = -0.637$), top-right: ALADIN/LACE with lagged statistics ($min = -0.438$), bottom-left: ALADIN/HU with standard statistics ($min = -1.642$), bottom-right: ALADIN/HU with lagged statistics ($min = -0.467$).

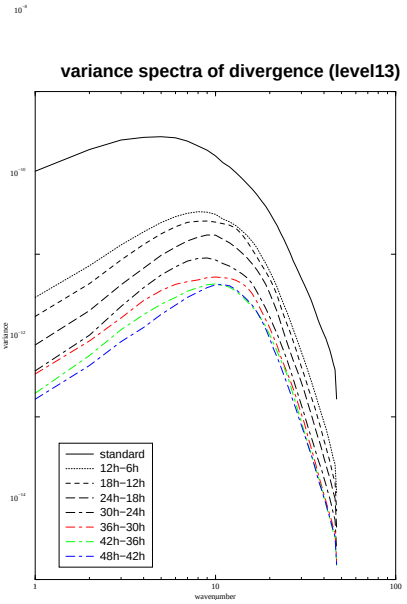


Figure 2: Background error variance spectra of divergence on model level 13. The spectra of lagged-NMC statistics while the time gap is 6 hour in the NMC differences and the spectra of standard statistics as a comparison.

[3] A global model using stretched geometry with an average around $20km$ resolution over the ALADIN/LACE domain.

Background Error Covariances in a Quasigeostrophic Reduced Rank Kalman Filter

Alexander Beck[1], Martin Ehrendorfer[2] and Patrick Haas[1]
[1] *Institute for Meteorology and Geophysics, University of Vienna, Althanstrasse 14, A–1090 Wien, Austria*
[2] *Institute for Meteorology and Geophysics, University of Innsbruck, Innrain 52, A–6020 Innsbruck, Austria*
email: alexander.beck@univie.ac.at

The purpose of this work is, in continuation of Beck and Haas (2001), to study the benefits on analysis quality of including *dynamical* (i.e., state–dependent) background information in a simple, yet reasonably realistic environment. The impact of flow–dependent background error covariances is studied within a four–dimensional variational data assimilation (4D–Var) system based on the quasigeostrophic model of Marshall and Molteni (1993), carrying 1449 degrees of freedom. This 4D–Var system allows for assimilating a given set of synthetic observations in cycling experiments with different specifications for the background error covariance matrix B. Specifically, B may be kept either static, or fully dynamic as the entire analysis error covariance matrix is carried forward in time according to the Extended Kalman Filter (EKF) equations (Ehrendorfer and Bouttier 1998). Further, B may be specified as a combination of dynamically evolved analysis error covariances and static background according to the theory developed for the Reduced Rank Kalman Filter (RRKF). The implementation of the RRKF closely follows the formulation at the European Centre for Medium–Range Weather Forecasts (see, Fisher 1998; Fisher and Andersson 2001). The basic idea of the RRKF is that the dynamical propagation of the background errors is only applied in a subspace of relatively small dimension (say, k) that is defined by the so–called Hessian singular vectors (HSV; Barkmeijer et al. 1998). Preliminary results from four cycling experiments are presented that are designed to investigate the performance of the RRKF – in different configurations (i.e., different dimension of the subspace k) – in comparison to either a static background formulation, or to the EKF. Each of the cycling experiments covers 12 assimilation intervals with a window length of 12 hours each. These preliminary results suggest that the performance of this 4D–Var system is sensitive to the specification of the background.

As an example, Fig. 1 shows the analysis and the forecast error (fc; as a function of lead time), for a static B (solid curve), an RRKF formulation with k=10 (dotted) and k=100 (dashed) HSVs, and the full EKF (chain–dashed) corresponding to an RRKF with k=1449 HSVs. Errors are measured in terms of the total energy (TE) metric (see, Ehrendorfer 2000), in units of J / kg. Each of the curves is the mean over the 12 subsequent assimilation intervals. Thinned observations on an irregular grid are sampled every 6 hours from a "truth run" with pre–specified observation error variances. It is this "truth run" that has been used in the computation of the forecast (and analysis) errors. Hence, these errors can be quantified exactly. Fig. 1 suggests that incorporating dynamical background error information is beneficial in terms of reducing analysis and subsequent forecast errors. Such a benefit is also expected from theoretical considerations (e.g., Fisher 1998).

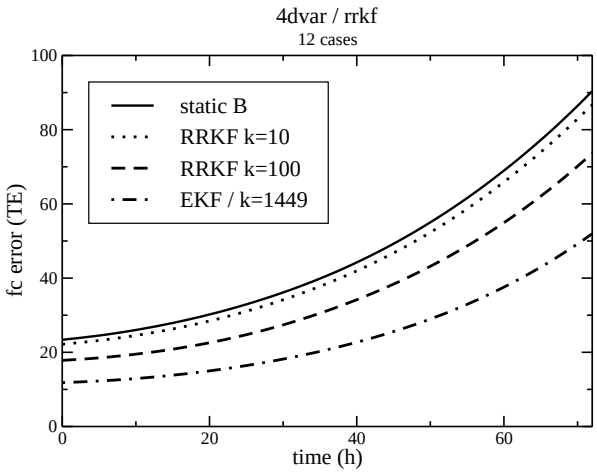


Fig. 1: Performance of the 4D–Var system in terms of analysis and fc error as a function of lead time.

However, it is also evident that a significant improvement is only achieved if O(100) HSVs are used. Finally, it is necessary to emphasize that these results are dependent on various other specifications of the 4D-Var system studied here, and are thus of a very preliminary nature. Additional investigation is necessary to further confirm the benefits of a dynamical B in data assimilation.

References

Barkmeijer, J., M. van Gijzen, and F. Bouttier, 1998: Singular vectors and estimates of the analysis error covariance metric. *Quart. J. Roy. Meteor. Soc.*, **124**, 1695–1713.

Beck, A., and P. Haas, 2001: Reduced rank Kalman filtering based on quasigeostrophic dynamics. In: *Research Activities in Atmospheric and Oceanic Modeling*, H. Ritchie (Ed.), Volume No. 31, WMO/TD-No. 1064, pp. 1.1–1.2. WMO.

Ehrendorfer, M., 2000: The total energy norm in a quasigeostrophic model. *J. Atmos. Sci.*, **57**, 3443–3451.

Ehrendorfer, M., and F. Bouttier, 1998: *An explicit low–resolution extended Kalman filter: Implementation and preliminary experimentation*. ECMWF Tech. Memo. No. 259, 21 pp. [Available from ECMWF, Shinfield Park, Reading RG2 9AX, England].

Fisher, M., 1998: *Development of a simplified Kalman filter*. ECMWF Tech. Memo. No. 260, 16 pp. [Available from ECMWF, Shinfield Park, Reading RG2 9AX, England].

Fisher, M., and E. Andersson, 2001: *Developments in 4D–Var and Kalman Filtering*. ECMWF Tech. Memo. No. 347.

Marshall, J., and F. Molteni, 1993: Toward a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.*, **50**, 1792–1818.

HOURLY MESO-SCALE SURFACE DATA ANALYSIS FOR NOWCASTING

F. Bouyssel, V. Ducrocq, F. Favot, M. Nuret, F. Taillefer, G. Therry
Météo-France/CNRM, 42 Avenue Coriolis, 31057 Toulouse, France
E-mail: Francois.Bouyssel@meteo.fr

An hourly meso-scale analysis of near surface parameters (T2m, Hu2m, V10m) has been developed over France using the high density (spatial and temporal) surface observations. This analysis uses the optimal interpolation technique and has been tuned to represent meso-scale events for forecasters rather than for initializing a numerical model. The first guess is the more recent forecast available from the operational French meso-scale model (ALADIN), which is comprised between 4 and 9 hours range. The horizontal spatial resolution of the first guess and the analysis is 9 km. The density of the hourly surface observation network over France is 40 km in average. The quality control of observations consists in a comparison to the first guess and in a test of spatial consistency. The test tunings are not too severe in order to keep meso-scale features like for instance those associated with thunderstorms. Some specific controls are applied to avoid that mountainous or coastal observations degrade the analysis since the structure functions of forecast errors are supposed homogeneous and isotropic. Additional diagnostics, such as the convective available potential energy (CAPE) of the 2m air parcel, or the low-level convergence moisture (MOCON), are computed from the meso-scale analysis outputs.

The analysis has been first evaluated on several convective case studies. Then a real time experimentation has been conducted during summer 2001 where a large part of the Météo-France meso-scale surface observations have been concentrated on an hourly basis. In order to improve the analysis on the frontier regions, the available surface observations from the neighbouring countries have been also taken into account in the analysis. During the experimentation, the advice of 'forecasters' has been sought on the operational use of these meso-scale analyses. The analysed fields were available 40 mn after observation time on the forecaster's workstation. Several possible uses have been identified. The meso-scale analysis is interesting as an interpolator tool for observations and is able to represent some meso-scale features like wind gustiness or strong cooling under thunderstorms. The differences between the meso-scale analysis and the ALADIN forecast are useful to confirm or invalidate the forecast scenario. Finally the potential of the diagnostics computed from analyses for thunderstorm nowcasting is still under investigation.

Assimilation and Forecast of Hurricane Floyd with the DAO Finite Volume Data Assimilation System

Jiun-Dar Chern* and Shian-Jiann Lin
NASA/GSFC, Greenbelt, MD 20771
E-mail: jchern@dao.gsfc.nasa.gov (6)

1. Introduction

The last decade has seen a growing interest among the major operational weather prediction centers in the application of global models for tropical cyclone track and intensity predictions. Current generation of global models with resolution in the range of 50 to 100 km is generally sufficient for resolving the synoptic scale steering flow. But it is still too coarse to resolve the inner core of tropic cyclone for reliable intensity predictions. The primary goal of present study is to examine the analysis quality and forecast skill of Hurricane Floyd with the new DAO global Data Assimilation System.

2. The fvDAS/model based forecast system

In recent years, a major effort has been undertaken at Data Assimilation Office (DAO), NASA Goddard Space Flight Center to develop the next-generation data assimilation system — the Finite Volume Data Assimilation System (fvDAS). The system consists of the joint NASA/NCAR general circulation model (fvGCM, Lin and Rood 2002) and the physical-space Statistical Analysis System (PSAS, Cohn et. al 1998). Preliminary results from forecast experiments using the fvGCM with initial conditions produced by fvDAS have shown that in general there is significant improvement in the forecast skill over DAO's operational Goddard Earth Observing System Data Assimilation System (GEOS-3 DAS). There are still many aspects of the new system need to be tested and evaluated. In this study, we have carried out a series of assimilation and forecast experiments of Hurricane Floyd with the fvDAS to evaluate the analysis quality and forecast capability.

3. Experimental design and results

Hurricane Floyd passed fairly close to the entire U.S. east coast from September 14 to September 17, 1999 and resulted in one of the largest evacuation in U.S. history: an estimated two million people were evacuated. It was blamed for 56 deaths and about $3-6 billion in total damage. In this study, a series of assimilation and 5-day forecast have been carried out from September 9 to September 15 using a pre-release version of fvDAS. By default, the fvDAS has 1° x 1.25° horizontal resolution with 55 vertical levels from surface to 1 Pascal. Tropical cyclones develop over the tropical oceans, and these are the data-sparse areas of convectional surface-based observations. Remote sensing provides the most reliable and often the only data source about the structure and position of tropical cyclone. It is very important to use satellite data to improve the initial analysis of the hurricane. In addition to the convectional data, we have utilized the SSM/I total precipitable water, TIROS Operational Vertical Sounder (DAOTOVS), cloud tracking

wind, and QuickSCAT SeaWinds Scatterometer surface wind in the data assimilation system.

Figure 1 shows Hurricane Floyd tracks form National Hurricane Center observed best track, fvDAS analysis, and fvGCM 5-day forecast starting at 0000 UTC 12 September 1999. In spite of the fact that the 1° x 1.25° resolution can only represent the broad features of tropical cyclones and its environment, the hurricane track from fvDAS analysis is in excellent agreement with the best track. The 5-day forecast with initial condition from fvDAS analysis also shows good skill in forecasting the track of Floyd, especially the first 48 –hour forecast.

4. References

Cohn, S. E., A. da Silva, J. Guo, M. Sienkiewicz, D. Lamich. 1998: Assessing the effects of data selection with the DAO physical-space statistical analysis system. *Mon. Wea. Rev.*, **126**, 2913-26.

Lin, S. J. and R. B. Rood. 2002: A vertically Lagrangian finite-volume dynamical core for global models. (to be submitted to *Mon. Wea. Rev.*)

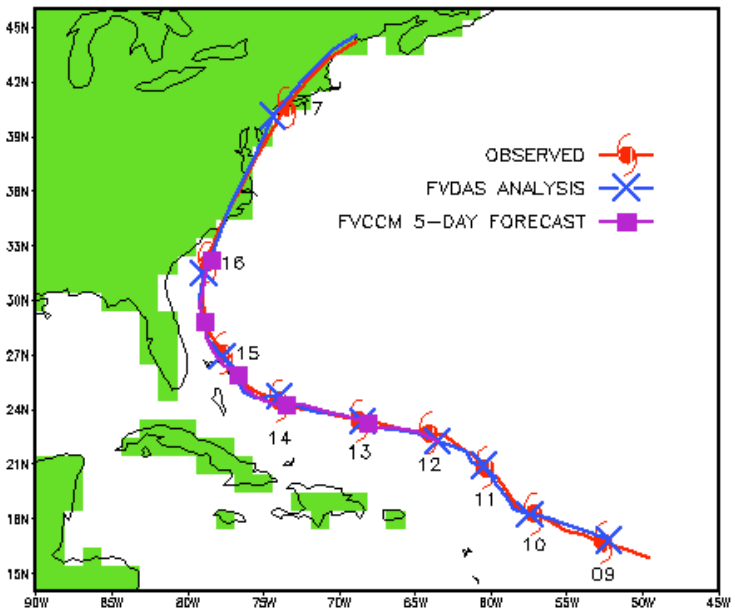


Figure 1. Hurricane Floyd tracks form National Hurricane Center observed best track, NASA DAO Finite Volume Data Assimilation System (fvDAS) 1º x 1.25º analysis, and fvGCM model 5-day forecast starting at 0000 UTC 12 September 1999.

THE CANADIAN 3D-VAR ANALYSIS SCHEME ON MODEL VERTICAL COORDINATE: RECENT AND FUTURE WORK

Clément Chouinard, Cécilien Charette, Jacques Hallé, and Réal Sarrazin.
MSC, Dorval, Canada
e-mail: clement.chouinard@ec.gc.ca

1. INTRODUCTION

The operational analysis system of the Canadian Meteorological Centre (CMC) has undergone major revisions in the last two years. The changes described here are really a follow-up to the June 1997 implementation of the first Canadian 3D variational system (3D-var) described in Gauthier et al. 1999. First, in June 2000, 3D-var was converted from a 16-pressure to a 28 terrain-following-η level system (3D-Var-η) including a complete revision of background and observational errors. In September 2000, the use of satellite data was updated to directly assimilate TOVS radiances, in replacement of SATEM thickness data. In December 2001, the 3D-var's use geopotential data from radiosondes was abandoned in favor of using temperature and surface pressure. The same was done with the synoptic surface data and both have had a very positive impact on analyses.

Some years ago, the CMC developed unified model codes for both the global and regional forecast models (Côté et al. 1998). Similarly, the 3D-var system was also coded to support both the regional and global pressure systems as described in Gauthier et al. 1999, and Laroche et al. 1999. In January 2001, the regional spin-up system was fully upgraded from the 16 pressure levels to 28 terrain-following η levels including the direct assimilation of TOVS radiances as in the global analysis system. In December 2001, its use of geopotential data from radiosonde was also abandoned in favor of using temperature and surface pressure.

2. GENERAL FORMULATION OF THE 3D-VAR, AND ESTIMATION OF NEW STATISTICS

Several operational NWP centres currently employ, or have employed a 3D-Var system. The 3D-var formulation of this study is based on the incremental approach, and the analysis variables are actually the incremental corrections. The initial misfits between the observations and the short-term forecast or innovations are computed in observation space using the full resolution background state whereas the analysis increments are calculated at lower resolution. In the global system, the trial fields are used at the full resolution (0.9 degree grid) of the GEM model and the analysis increments are calculated at the lower T108 spectral resolution.

The covariance statistics of the 3D-Var-η system were redesigned based on a 3-month ensemble of 24 and 48-h forecasts valid at the same time. The 24-48-h forecast differences are not entirely representative of 6 hour forecast error statistics, and for this reason, the estimated variance fields are assumed to be zonally invariant and were also scaled down using information from the variances of radiosonde observations minus 6h-forecast wind and temperature averaged over broad latitude bands.

The 3D-Var-η system is multivariate and like its predecessor imposes a simplified balance constraint between mass and wind analysis increments. The complete balance operator is actually the product of local horizontal and vertical balance operators as discussed in Gauthier et al. (1999). The vertical operator which transforms vertical profiles of the balanced mass variable into temperature profiles is estimated using a regression analysis over the ensemble of error samples between temperature profiles and profiles of linearly balanced stream function increments in grid-point space. This approach is used to avoid problems of increased noise in the vertical structure and the null space associated with using a theoretically based inverse hydrostatic operator. Typically, the horizontal and vertical spread of balanced is larger than those of the unbalanced component (Chouinard et al. 2001).

3. NEW SOURCES OF DATA; SATELLITE AND AIRCRAFT

In 3D-Var-η, observations are assimilated in their raw or unprocessed form with the use of so-called "observation operators" thereby avoiding interpolating data to and from a fixed pressure grid prior to their assimilation. For example, TOVS radiances are directly assimilated with the help of a radiance transfer model (RTM). The advanced microwave sounding unit (AMSU) of the recent NOAA series produces data of exceptional quality that are relatively easier to use than the IR data in most sky conditions. In the current 3D-Var-η, only the microwave data of channels 3-10 of AMSU-A from NOAA-15 and 16 platforms are used. All radiances used are quality controlled

prior to the monitoring and assimilation steps as described in Chouinard et al. (1999, 2000). The impact of TOVS radiances is very large and positive everywhere, but exceptionally so in the SH, where the predictability of the system was improved by close to 24 hours (Chouinard et al. 2001).

In recent years, a very important source of wind data has become available in the form of automated aircraft wind reports (ACARS/AMDAR). The impact of these winds when added to the AIREP data was shown to be marginally negative in the previous 16-pressure level 3D-Var system and consequently was never implemented even though monitoring indicated the data was of very good quality. Because the correlation structures of the 3D-Var-η system are much improved, the impact of the additional ACARS data has now been shown to produce very positive impacts particularly when accompanied by TOVS radiances (Chouinard et al. 2001).

3.1. Direct assimilation of temperature (including significant levels) and surface pressure

Even though the analysed mass variables of the current 3D-Var-η regional and global systems are temperature and surface pressure, geopotential has remained the main source of observations from RAOBS and similarly the surface pressure was assimilated indirectly as a proximity to surface geopotential datum. Because of this, significant level temperatures from RAOBS and aircraft temperature reports were never assimilated in the current system. Recently, we have introduced the direct assimilation of temperatures and surface pressure from RAOBS instead of geopotential, and similarly, temperature and moisture observations from the surface synoptic meteorological (SM) network are now directly assimilated producing larger and more consistent corrections to the trial field thereby improving the surface and PBL structures. Parallel suites prior to the December 2001 implementation clearly indicate the positive impact of direct assimilation of surface and upper air temperature data on analyses and 10-day forecasts.

3.2. Application to the regional analysis system

Like its previous pressure system, the 3D-Var-η system was adapted to the regional model to produce analyses directly on its model levels. As in the global, the analysis increments are calculated at the low-resolution horizontal and vertical resolution of the analysis grid of the global statistics. The increments are interpolated to and from the regional model grid during the 12-h spin-up to arrive at the final analysis on the higher resolution regional model grid. As expected, the same improvements obtained with the global 3D-Var-η were obtained in the regional system, but most significantly, the moisture analyses and the precipitation forecasts were significantly improved (Chouinard et al. 2001).

4. FUTURE WORK

In preparation for the assimilation of satellite data from future platforms (NOAA-17), we have started testing with AMSU-B TOVS radiances from current NOAA polar orbiters. The model and analysis top are being raised to 0.1 hPa to accommodate the higher peaking radiances of the AIRS and IASI instruments.

5. REFERENCES

Chouinard C. And J. Hallé: The impact of TOVS radiances in the CMC 3D-Var analysis system. ITSC-X proceedings, Boulder Colorado, February 1999, p92-98.

Chouinard C., J. Hallé, and R. Sarrazin: Recent results with TOVS data in the new CMC 3D-Var-analysis system: the combined and separate impact of microwave radiance observations with aircraft wind data. ITSC-XI proceedings, Hungary, Budapest, September 2000, p53-57.

Chouinard C., C. Charette, J. Hallé, P. Gauthier, J. Morneau, and R. Sarrazin: The Canadian 3D-Var analysis scheme on model vertical coordinate. 18th Conference On Weather Analysis and Forecasting, 30 July-2 August 2001, Fort Lauderdale, Florida.

Côté, J.; S. Gravel, A. Méthot, A. Patoine, M. Roch and A. N. Staniforth. 1998. The operational CMC/MRB global environmental multiscale (GEM) model: Part I - Design considerations and formulation. Mon. Weather Rev. 20: 1373-1395.

Gauthier P., C. Charette, L. Fillion, P Koclas, and S. Laroche, 1999: Implementation of a 3D variational analysis at the Canadian Meteorological Centre. Part I: The global analysis. Atmosphere-Ocean. Vol. XXXVII, No. 2 pp 103-156.

Laroche S.,P. Gauthier , J. St-James, and J. Morneau, 1999: Implementation of a 3D variational analysis at the Canadian Meteorological Centre. Part I: The regional analysis. Atmosphere-Ocean. Atmosphere-Ocean. Vol. XXXVII, No. 3 pp 281-307.

USE OF THE BREEDING TECHNIQUE IN THE ESTIMATION OF THE BACKGROUND COVARIANCE MATRIX FOR A QUASI-GEOSTROPHIC MODEL

M. Corazza[1,2], E. Kalnay[1], D. J. Patil[1], E. Ott[1], J. A. Yorke[1], B. R. Hunt[1], I. Szunyogh[1] and M. Cai[1]
[1] University of Maryland, College Park, 20742 MD, USA
[2] INFM – DIFI, Università di Genova, 16146 Genova, Italy

It is well known that numerical weather predictions are sensitive to small changes in the initial conditions, i.e., a rapid growth of the initial errors can lead in a relatively short time to large forecast errors. During the last decades much effort has been devoted to study and improve the methods used for the preparation of the initial conditions for numerical atmospheric models (the so-called analysis), as well as to understand the mechanisms involved in the growth of the initial errors.
The analysis is obtained as a statistical interpolation of short-range numerical forecasts (known as background) with new observations. The weight given to each of these contributions is essentially proportional to the inverse of their error covariance. It follows that a good representation of the observation and background error covariances is one of the major goals in the development of data assimilation systems (e.g. Klinker et al., 2000; Bennet et al., 1996; Houtekamer and Mitchell, 1998; Hamill and Snyder, 2000). In 3D-Variational schemes the background error covariance matrix is statistically derived from long term statistical estimations and it is maintained *constant in time* during the assimilation cycle. This implies that the large time dependence of the errors ("errors of the day") is neglected, despite its large variability (Corazza et al, 2001).

Kalnay and Toth (1994) argued that the similarity between breeding (Toth and Kalnay, 1993, 1997) and data assimilation suggests that the background errors should have a structure similar to those of bred vectors. Here we take advantage of the relationship between the bred vectors and the analysis and background errors demonstrated by Corazza et al., 2001, for a simple quasi-geostrophic model (Morss, 1999), and test whether it is possible to augment the constant forecast error covariance used in 3D-Var with "errors of the day" derived from the breeding method. We present results obtained with different methods aimed to include in the data assimilation scheme (i.e., in the representation of the background error covariance matrix) the information given by the bred vectors.

The numerical model is a quasi-geostrophic (QG) mid-latitude flow in a channel discretized by finite differences both in horizontal and vertical directions. The simulated data assimilation is performed with an algorithm similar to the operational Spectral Statistical Interpolation (SSI) at NCEP (Parrish and Derber, 1992). "Rawinsonde observations" are generated every 12 hours by randomly perturbing the true state at fixed observation locations. Bred vectors are produced using a method similar to that adopted at NCEP (Toth and Kalnay 1993,1997), rescaling the difference between the perturbed runs and the control forecast every 12 hours.

Since this is a simulation system, we can explicitly define the "true state of the atmosphere" (by integrating the model from a given initial state) and therefore study the analysis and forecast errors. A perfect model assumption is made so that our conclusions are not necessarily valid for more complex models with model errors, and similar tests have to be made with more general simulation systems and with real forecast systems.

Corazza et al. (2001) found that bred vectors in this QG simulation system are indeed closely related to the background errors, suggesting that the bred vectors can be useful in specifying the part of the background error covariance matrix that corresponds to the "errors of the day". In particular we found the following properties of the bred vectors:

- Convergence to well organized structures in the bred vectors occurs within a few (3-5) days. This indicates that it is possible to operationally use the information given by the bred vectors without waiting an infinite time for asymptotic convergence.
- Bred vectors obtained using normalizations based on the potential vorticity and on the stream function are virtually indistinguishable.
- Bred vectors obtained using the "true" atmosphere are very similar to those obtained using the "analysis" atmosphere. This is true even if we use a low density observing network, suggesting that the bred vectors are not too sensitive to the details of the flow and that the errors themselves are more likely dependent on the large scale nature of the flow.

The original data assimilation cycle (referred to as the *regular* data assimilation system) is based on the NCEP 3D-Var scheme, and is solved iteratively for the analysis state $\mathbf{x}_a$ (Morss, 1999). Given the background state (or first guess - the 12 hour forecast from the previous analysis) $\mathbf{x}_b$, and the set of observations $\mathbf{y}_o$, the equation can be written as follows:

$$(\mathbf{I}+\mathbf{B}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})(\mathbf{x}_a-\mathbf{x}_b)=\mathbf{B}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}_o-H(\mathbf{x}_b))$$

where **B** is the background error covariance matrix, **R** is the observation error covariance matrix, *H* is the observation operator and **H**, $\mathbf{H}^{\mathrm{T}}$ are the matrices that represent the linearized *H* and its transpose respectively. The right part of the equation is computed at the beginning of the process, and the equation is then iteratively solved for $(\mathbf{x}_a-\mathbf{x}_b)$ until the equation is satisfied with an error smaller than a given threshold.

The easiest way to introduce the bred vectors in this equation is to globally substitute **B** with the ensemble average of the outer product of the bred vectors. We can build a new background covariance matrix as $\sum_{i=1}^{k}\mathbf{b}_i\mathbf{b}_i^{\mathrm{T}}/k$ where $\mathbf{b}_i$ is the i^{th}

bred vector defined over the entire domain. The substitution of **B** with the new matrix can be done at a negligible computational cost. Moreover, this implementation allows to apply $\sum_{i=1}^{k} \mathbf{b}_i \mathbf{b}_i^{\mathrm{T}} / k$ and **B** simultaneously (Hamill and Snyder, 2000) so that the data assimilation scheme can be generalized to:

$$\left(\mathbf{I} + \left(\alpha \frac{c}{k} \sum_{i=1}^{k} \mathbf{b}_i \mathbf{b}_i^{\mathrm{T}} + (1-\alpha)\mathbf{B} \right) \mathbf{H}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{H} \right) (\mathbf{x}_a - \mathbf{x}_b) = \left(\alpha \frac{c}{k} \sum_{i=1}^{k} \mathbf{b}_i \mathbf{b}_i^{\mathrm{T}} + (1-\alpha)\mathbf{B} \right) \mathbf{H}^{\mathrm{T}} \mathbf{R}^{-1} \left(\mathbf{y}_o - H(\mathbf{x}_b) \right)$$

where α is a number between 0 (for the regular system) and 1 (background error covariance matrix based fully on bred vectors) and *c* is a normalization factor kept constant in the results presented here. It should be noted that the covariances in **B** were tuned to optimize the regular 3D-Var.

The use of the bred vectors allows decreasing the squared error of the analysis (averaged over the horizontal domain) by a factor between 15 and 20% (around 8-10% in the error) for α equal to 0.4. The percentage improvement continues throughout the 72 hour forecast, suggesting that the correction to the analysis due to the bred vectors affects, at least in part, the growing errors. This method is able to reduce the squared error in the analysis and forecasts up to a factor of 40% considering a modified version of the bred vectors aimed to take into account random observational errors introduced in the analysis step. Simulations using these vectors show a remarkable improvement of the performance of the assimilation cycle with respect to the one based on the standard bred vectors without random "reseeding". It is interesting to note that for large values of α, when the role of the statistically derived **B** is small, the augmented system is not able to maintain the error small. This indicates that the space spanned by the bred vectors is not large enough to represent all the error directions, and that the contribution of the regular part of the assimilation scheme cannot be neglected when using global methods to include bred vectors in the data assimilation cycle. This was also observed using local methods (not shown).

Motivated by the studies performed by Patil et al. (2001) we are presently testing new methods to locally use the bred vectors in the data assimilation system. The use of local methods is desirable in order to optimise the information given by two or more bred vectors, which may be, for example, positively correlated in one area and negatively correlated in another area far away. The background error correlations should vanish beyond a limited horizontal extent, whereas our use of global bred vectors implies correlations over the global domain. An example of local use of the bred vectors can be derived as a generalization of the method proposed by Kalnay and Toth (1994).

REFERENCES

Bennett, A. F., B. S. Chua, and L. M. Leslie, 1996: Generalized inversion of a global NWP model., *Meteor. Atmos. Phys.*, **60**, 165-178.

Corazza, M., E. Kalnay, D. J. Patil, R. Morss, I. Szunyogh, B. R. Hunt, E. Ott, and M. Cai, 2001: Use of the breeding technique to estimate the structure of the analysis "errors of the day". Submitted to Nonlinear Processes in Geophysics.

Hamill, T. M., and C. Snyder, 2000: A Hybrid Ensemble Kalman Filter – 3D Variational Analysis Scheme., *Mon. Wea. Rev.*, **128**, 2905-2919.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique., *Mon. Wea. Rev.*, **126**, 796-811.

Kalnay, E., and Z. Toth, 1994: Removing growing errors in the analysis cycle., *Tenth Conference on Numerical Weather Prediction – Amer. Meteor. Soc.*, pp. 212-215.

Klinker, E., F. Rabier, G. Kelly, and J.-F. Mahfouf, 2000: The ECMWF operational implementation of four dimensional variational assimilation. III: Experimental results and diagnostics with operational configuration., *Quart. J. Roy. Meteor. Soc.*, **126**, 1191.

Morss, R. E., 1999: *Adaptative observations: Idealized sampling strategies for improving numerical weather prediction.*, Ph. D. thesis, Massachusetts Institute of Technology, 225 pp.

Parrish, D. F., and J. D. Derber, 1992: The National Meteorological Center spectral statistical interpolation analysis system., *Mon. Wea. Rev.*, **120**, 1747-1763.

Patil, D. J. S., B.R. Hunt, E. Kalnay, J. A. Yorke, and E. Ott, 2001: Local Low Dimensionality of Atmospheric Dynamics, *Phys. Rev. Lett.*, **86**, 5878.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NCEP: the generation of perturbations., *Bull. Amer. Meteor. Soc.*, **74**, 2317-2330.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP: the breeding method., *Mon. Wea. Rev.*, **125**, 3297- 3318.

Near-Realtime Sea Surface Pressure Fields from NASA's SeaWinds Scatterometer and Their Impact in NWP

Shannon R. Davis, Mark A. Bourassa, R. Atlas, J. Ardizzone, Eugenia Brin
James J. O'Brien, David F. Zierden

1. Introduction

Spaceborne scatterometers provide marine surface wind vector measurements with unprecedented coverage in space and time. As a result, these instruments can potentially serve as invaluable tools in operational meteorology as well as in the realm of numerical weather prediction. The most recently launched satellite scatterometer, SeaWinds, relays data from over 94% of the planet's ice-free oceans in a 24-hour period. Here the impact of assimilated SeaWinds data upon a numerical weather models performance is examined in a joint collaboration betweent the Center for Ocean and Atmospheric Prediction Studies and NASA's Data Assimilation office at the Goddard Space Flight Center. Particular emphasis devoted to the impact of assimilated scatterometer-derived sea surface pressures. Several experiments are conducted using a new NASA/NCAR finite volume model to separately assess the impact of assimilated sea-surface pressure fields derived from SeaWinds and assimilated SeaWinds winds themselves. The results from these experiments are objectively analyzed and the individual impacts of assimilated SeaWinds variables is compared.

2. Background

This study involving the assessment impact of assimilated scatterometer-derived pressures in addition to assimilated scatterometer winds is a significant departure from many other earlier works. With preliminary impact studies made by Baker et al. (1984) and Duffy et al (1984), it was established that satellite scatterometer wind data (SeaSAT) alone could provide significant improvements in the surface analyses for major synoptic events but less of an impact on the analyses and forecasts for the upper atmospheric levels. In subsequent studies, several initiatives by Duffy and Atlas and more recently by Atlas(2001), demonstrated that the impact of scatterometer surface winds from could be extrapolated vertically by adjusting the mass of upper levels relative to the adjusted features observed at the surface. Here, an alternative technique to enhancement of scatterometer data impact is advanced. Sea-surface pressures derived from the scatterometer data are assimilated into the model in place of the wind data itself. Assuming a relatively hydrostatic state in the atmosphere, surface pressures represent a three dimensional column of the atmosphere rather than the two-dimensional equivalent a the sea surface presented by the wind vectors alone. Surface pressure fields additionally affect the mass fields of the atmosphere directly and thus the impact of their assimilation into the model should be much stronger than that of the winds alone. With the use of pressures, the issues regarding complicated boundary layer physics in a model can be circumvented and so too the need for any vertical extrapolation scheme to adjust the model's upper atmospheric levels.

The calculation of sea-surface pressures from the scatterometer wind vectors in this study is made following the works of Harlan and O'Brien (1986), Brown and Zeng (1994), and ,ost recently Zierdent et. al. (2000). Wind vectors from the SeaWinds instrument are produced along a regularly gridded swath that flows. From these wind vectors observations, relative vorticity is determined following a centered difference scheme and blended via a variational technique with a geostrophic vorticity derived from an initial guess/model pressure field. A new surface pressure field is solved numerically using a method of successive overelaxation.

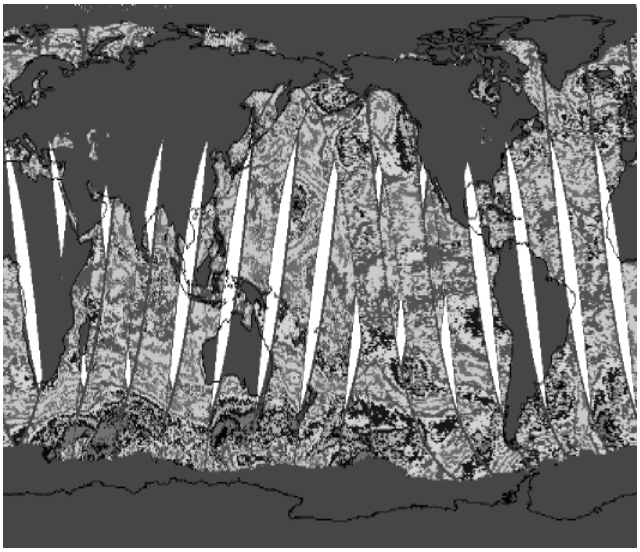
3. Data & Instrument

The SeaWinds instrument was launched by NASA in 1999 aboard the sun-synchronous, polar-orbiting satellite, QuikSCAT. It is a Ku-band scatterometer in the tradition of its predecessors SeaSAT and NSCAT, and functions as an active microwave sensor operating at 13.2 GHz Unlike its predecessors however, SeaWinds employs twin conically rotating pencil beams in its design, thereby improving the quantity and quality of its observations. With an orbital period of 101 minutes, the coverage of SeaWinds is typically 94% of the ice free oceans in one day.

Since its launch several different data sets have been made available to researchers and operational weather centers by NASA-JPL and NOAA-NESDIS. These include processed science quality and near-realtime wind vectors. Both are employed in this study to be consistent with previous impact studies and to suggest the operational viability of using the SeaWinds winds and wind-derived surface pressures.

4. Model

A new Physical Space/ Finite-Volume global circulation model was employed in these experiments This model was developed jointly by NCAR and the NASA Data Assimillation Office. Detailed information about the model is available on the web at:
http://dao.gsfc.nasa.gov/pages/atbd.html



Sample Coverage for one Day By the QSCAT SeaWinds Scatterometer

SeaWinds Pressures and Winds Impact Study Methodology

5. Methodology

A series of individual experiments are conducted to assess and compare the impact of assimilated SeaWinds-derived pressures and assimilated SeaWinds surface wind vectors. The methodology for this impact study follows most closely that established by Atlas et al (2001) and is outlined graphically in Figure 2.

6. Results

Results from the assimilation of one-month of global Seawinds Pressures are presently being generated at NASA's Data Assimilation Office. Impacts of these pressures on the surface analysis alone are shown in figure 3, 4, 5,6, and 7.

7. Acknowledgements

SeaWinds Data used is this experiment was processed and made available by several essential research groups. NASA support was directed from the Ocean Vector Wind Science Team. Data sets were produced by Remote Sensing Systems and sponsored in part by NASA's Earth Science Information Partnership, an association of technical information sites for the Earth sciences and by the NOAA/NASA pathfinder Program for early EOS products; Principal investigator: Frank Wentz.

The Center for Ocean Atmospheric Prediction Studies receive its base funding from ONR's Secretary of the Navy grant.

8. References

Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. J. Meteor. Soc. Japan, 75(1B),111-130.

Atlas,R., S. C. Bloom, R. N. Hoffman, E. Brin, J. Ardizzone, J. Terry, D. Bungato, and J. C. Jusem,1999: Geophysical validation of NSCAT winds using atmospheric data and analyses. J. Geophys. Res., 104(C5),11405-11424.

Atlas, R. and R.N. Hoffman. 2001: The use of satellite surface wind data to improve weather analysis and forecasting. in Satellites and Society, Elsevier, 57-79.

Bourassa,M.A., D. M. Legler, J.J. O'Brien, and S. Smith, 2001:SeaWinds validation with research vessels. J. Geophys. Res., submitted.

Brown, R.A., and L. Zeng, 1994: Estimating central pressures of oceanic mid-latitude cyclones. J. Appl. Meteor., 33,1088-1095

Hoffman, R.N., 1993:A preliminary study of the impact of ERS-1 C-band scatterometer wind data on the European Center for Medium Range Forecasts global data assimilation system. J. Geophys. Res., 98, 10233-10244

Liu, W.T., W.Q. Tang, and P.S. Polito,1998:NASA scatterometer provides global ocean-surface wind fields with more structures than numerical weather prediction, Geophys. Res. Lett., 25 ,761-764

Zierden, D. F., Bourassa,M.A., and J.J. O'Brien, 2000: Cyclone surface pressure fields and frontogenesis from NASA scatterometer(NSCAT) winds. J. Geophys. Res.,105, 23967-23981

A VARIABLE-RESOLUTION STRETCHED-GRID GENERAL CIRCULATION MODEL AND DATA ASSIMILATION SYSTEM WITH MULTIPLE AREAS OF INTEREST: STUDYING THE ANOMALOUS REGIONAL CLIMATE EVENTS OF 1998

Michael S. Fox-Rabinovitz, ESSIC (Earth System Sciences Interdisciplinary Center), University of Maryland, College Park, MD, and Data Assimilation Office, NASA/Goddard Space Flight Center, Greenbelt, MD

Lawrence L. Takacs, and Ravi C. Govindaraju SAIC (Science Applications International Corporation), 4600 Powder Mill Road, Beltsville, MD 20705-2675, and Data Assimilation Office, NASA/Goddard Space Flight Center, Greenbelt, MD

The GEOS (Goddard Earth Observing System) stretched-grid (SG) GCM and the GEOS SG-DAS have been developed and thoroughly tested over the last few years (Fox-Rabinovitz et al. 1997, 2000, 2001, 2002, Fox-Rabinovitz 2000). The model and system are used for regional climate experiments for seasonal, annual, and multiyear time scales. The following major results have been recently obtained.

Introduction of the new design of the stretched grid with multiple areas of interest allowed us to study simultaneously the variety of anomalous regional climate events of 1998 at mesoscale resolution. Both the SG-GCM simulation and SG-DAS assimilation products obtained with enhanced regional resolution are used in the study.

The new stretched-grid design with multiple (four) areas of interest, one at each global quadrant, is implemented into both a stretched-grid GCM and a stretched-grid data assimilation system (DAS). The four areas of interest include: the U.S./Northern Mexico, the El-Nino area/Central South America, India/China, and the Eastern Indian Ocean/Australia. Both the stretched-grid GCM and DAS annual (November 1997 through December 1998) integrations are performed with 50 km regional resolution while the maximum grid interval is about 3 degrees. The moderate stretching is used. The efficient downscaling to mesoscales is obtained for both the SG-GCM and SG-DAS for each of the four areas of interest while the consistent interactions between regional and global scales and the quality of global circulation is preserved. This is the advantage of the stretched-grid approach.

Also, the hurricanes, typhoons, and severe storms for the areas of interest and their vicinities are produced with the GEOS SG-DAS.

The areas of interest with enhanced regional resolution are located within each global quadrant that makes the grid-point global distribution more homogeneous than that of the original stretched grid design with one area of interest. Such homogeneity of a grid-point distribution affects positively the overall quality of global simulated and assimilated products.

The global fields and diagnostics are well reproduced by the SG-GCM simulation and SG-DAS assimilation. Their spectra are very close for all spectral ranges: the long-, medium-, and short-wave ones. The spectra are also close to that of the reference1x1 degree ECMWF reanalyses except for the shortest waves or mesoscales for which the stretched-grid spectra show larger energy due to higher resolution used for the large regions of interest.

The global zonal mean vertical distributions of prognostic variables are close to those of the reference 1x1 degree ECMWF reanalyses. The same is true for horizontal distributions of the prognostic and diagnostic fields. All that confirms that the high quality global characteristics are obtained for simulated and assimilated fields.

The simulated and assimilated anomalous regional climate events of 1998 include: the spring (April-June) flooding in the Midwest and Northeast and the drought in southeastern U.S.; the December-1997 - May-1998 Mexican drought; the Indian summer (June-September) monsoon; the severe summer (June-September) flooding in China; anomalous precipitation over Australia; anomalous March-May precipitation over South America; and precipitation over the African Sahel region. The above event simulation and data assimilation captured the major anomalies at medium and mesoscale resolution.

For all of the events the simulated and assimilated precipitation and/or precipitation anomaly patterns appeared to be close to each other and also close in many details to gauge precipitation data. Simulated precipitation is sometimes overestimated compared to that of assimilated or gauge data precipitation especially outside the areas of interest. Assimilated precipitation compares generally well with gauge data. The mesoscale features are adequately produced for both simulated and assimilated precipitation.

Other diagnostic and prognostic variables at different levels compare well with verifying data. All that shows the success of stretched-grid simulation and assimilation in terms of the efficient downscaling to realistic mesoscales.

The overall conclusion is that the SG-GCM simulation and SG-DAS assimilation produced realistic global and especially regional products that adequately represent the various anomalous regional climate events occurred in1998. Evidently, the quality of assimilated products is higher than that of simulated ones.

The obtained results show that both the SG-GCM and SG-DAS with multiple areas of interest are viable practical tools for simultaneous high-resolution simulations and data assimilations of regional climate events in all four global quadrants.

References

Fox-Rabinovitz, M. S., L.V. Stenchikov, M. J. Suarez, L. L. Takacs, 1997: A finite-difference GCM dynamical core with a variable resolution stretched-grid, Mon. Wea. Rev., Vol. 125, No. 11, 943-2968.

Fox-Rabinovitz, M. S., L.V. Stenchikov, M. J. Suarez, L. L.Takacs, and R.C. Govindaraju, 2000: An uniform and variable resolution stretched-grid GCM dynamical core with real orography, Mon. Wea. Rev., Vol. 128, No. 6, 1883-1898.

Fox-Rabinovitz, M. S., 2000: Regional climate simulation of anomalous U.S. summer events with a variable-resolution stretched-grid GCM, J. Geoph. Res., v. 105, No. D24, pp. 29,635-29,646.

Fox-Rabinovitz, M.S., L.L. Takacs, and M.J. Suarez, 2001: A Variable Resolution Stretched Grid GCM: Regional Climate Simulation. Mon. Wea. Rev., Vol. 129, No. 3, pp. 453-469.

Fox-Rabinovitz, M.S., L.L. Takacs, and R.C. Govindaraju, 2002: A stretched-grid GCM and DAS with multiple areas of interest, submitted.

Improving Global Analysis and Forecast Using Microwave-based Rain and Moisture Data

Arthur Y. Hou, Sara Q. Zhang, Arlindo M. da Silva

Data Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
E-mail: arthur.hou@gsfc.nasa.gov

Analyses produced by global data assimilation systems currently contain significant errors in primary hydrological fields such as precipitation and evaporation, especially in the tropics. Rainfall estimates derived from space-borne passive microwave sensors can provide valuable pattern and intensity information on precipitation for improving global analysis and forecast. At the Data Assimilation Office at NASA Goddard Space Flight Center, we have been exploring the use of innovative techniques to assimilate rainfall and total precipitable water (TPW) data provided by the TRMM Microwave Imager (TMI) and Special Sensor Microwave/Imager (SSM/I) instruments. Results show that variational assimilation of 6-h averaged surface rain rate and TPW using the moisture tendency of the forecast model as a control variable significantly improves not only the hydrological fields but also key climate parameters such as clouds, radiation, and tropospheric moisture in the analysis produced by the Goddard Earth Observing System (GEOS) Data Assimilation System. The improved analysis also yields improved short-range forecasts in the tropics.

Figure 1 shows the impact of assimilating 6-hour averaged TMI and SSM/I rainfall and TPW on GEOS analysis at 1^o x 1^o horizontal resolution for January 1998. The improved precipitation in the tropics effectively reduces the monthly-mean bias and standard deviation errors in the outgoing longwave radiation (OLR), which was not assimilated but used for independent verification. Since current global analyses contain significant errors in hydrological parameters, the result that rainfall assimilation improves not only precipitation but also related fields such as cloud and radiation has important implications. It identifies precipitation as a key observation type for improving the quality and usefulness of global analyses for understanding the earth's water and energy cycles.

The improved analysis with rainfall data also provides better initial conditions for storm-track and quantitative precipitation forecasts (QPF), as shown in Fig. 2 for Hurricane Bonnie. Results from 5-day ensemble forecasts show systematic improvements in precipitation, divergent winds and geopotential heights in the tropics. These results suggest that rainfall assimilation has the potential to significantly improve weather forecasting skills.

Reference:

Hou, A. Y., S. Zhang, A. da Silva, W. Olson, C. Kummerow, J. Simpson, 2001: Improving global analysis and short-range forecast using rainfall and moisture observations derived from TRMM and SSM/I passive microwave instruments. *Bulletin of Amer. Meteor. Soc.*, **82**, 659-679.

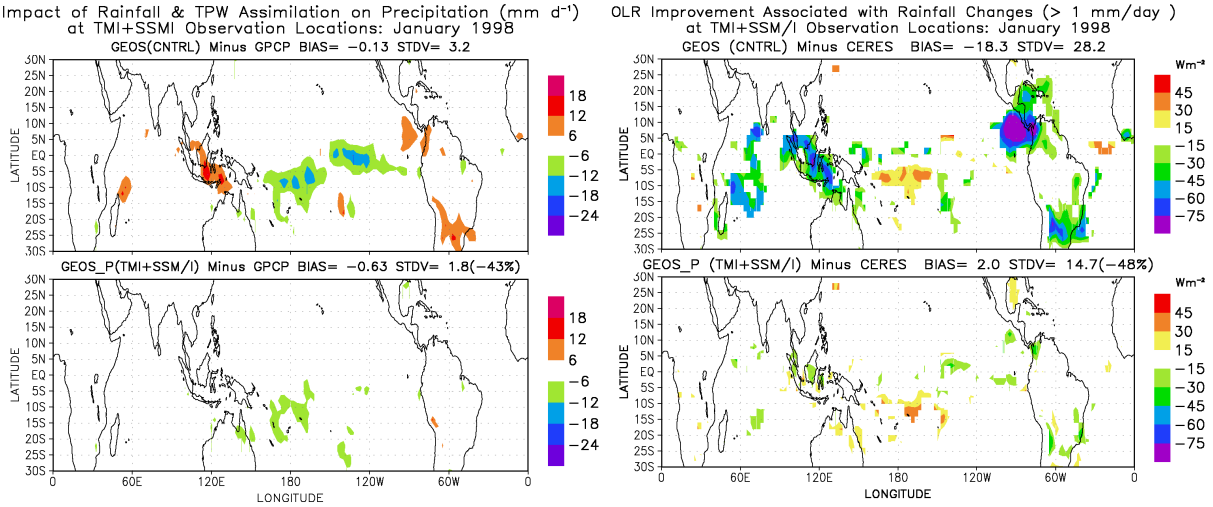


Figure 1. GEOS assimilation results with and without TMI and SSM/I observations for January 1998. Left panels show errors in the monthly-mean tropical precipitation fields verified against GPCP combined Satellite-gauge estimate: Top is the difference between the GEOS control (without rainfall and TPW data) and GPCP. Bottom is the corresponding error in GEOS assimilation with rainfall and TPW data. Right panels show the impact on the outgoing longwave radiation (OLR) verified against CERES/TRMM measurements. Percentage changes in the tropical-mean error standard deviation relative to the GEOS control are given in parentheses.

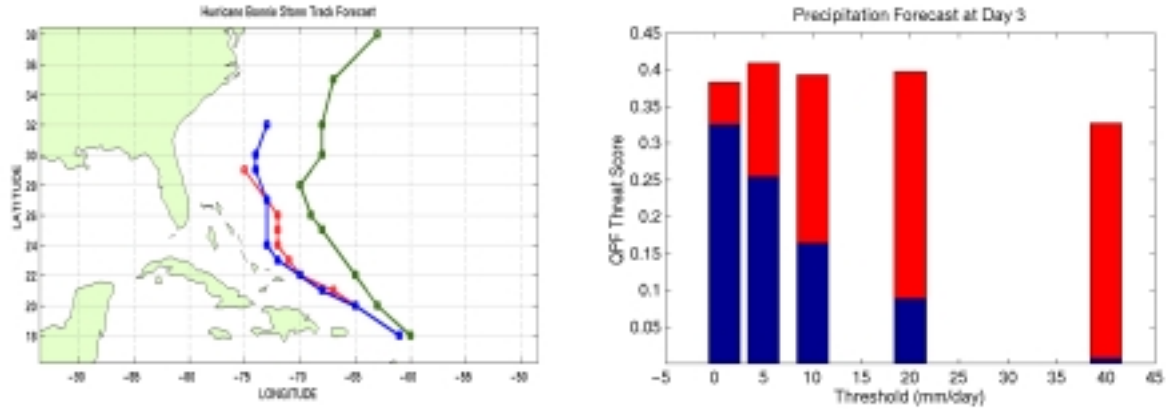


Figure 2. Improved storm track forecasts and QPF Equitable Threat Scores for Hurricane Bonnie. The left panel shows that the 5-day storm track forecast initialized with 1^{o} x 1^{o} GEOS analysis containing TMI and SSM/I rainfall data (blue) is in close agreement with the best track analysis from NOAA. The track from the control experiment is shown in green. The forecasts are initialized at 12:00 on 20 August 1998. The right panel shows the consistently higher Equitable Threat Scores for Day 3 precipitation forecast (red) initialized by the analysis with rainfall data. Results for the control experiment are shown in blue. A higher Threat Score corresponds to greater forecast skills.

One month cycle experiments of the JMA mesoscale 4-dimensonal variational data assimilation (4D-Var) system

Yoshihiro Ishikawa and Ko Koizumi*
Numerical Prediction Division, Japan Meteorological Agency

The JMA is developing a 4-dimensional variational data assimilation (4D-Var) system for a hydrostatic mesoscale model (MSM) with a horizontal resolution of 10km and 40 vertical levels. The system becomes operational from March 2002 with three-hour assimilation windows. Since it is aimed to provide MSM products within one and half hours from observation times, an incremental approach is taken to save computational time, using a 20km version of MSM for inner loop calculation of 4D-Var. The adjoint model includes simplified physics. Assimilated data are radiosonde, synop, ship, buoy, airep, wind-profiler and radar-AMeDAS precipitation data.

In order to evaluate the total performance of 4D-Var in the operational environment, 3-hour forecast-analysis cycle experiments were performed for one month period of June and September 2001. 18-hour forecasts were made four times a day (00, 06, 12 and 18 UTC initials). The root mean square errors (RMSE) of 120 forecasts were calculated for each month against radiosonde observations in Japan and threat scores were also calculated for precipitation forecasts evaluation. These scores were compared with those of the routine forecasts which employ the one-hour cycle optimal interpolation and physical initialization as an analysis method during the last three-hour period preceding to the initial time.

Fig. 1 shows threat scores of 10mm/3hour calculated against 40km-averaged radar-AMeDAS precipitation analysis data. The scores of 4D-Var surpass those of routine forecasts for every forecast time of both June and September.

Fig. 2 shows RMSEs of 500hPa heights and 850hPa temperature of June 2001. The RMSEs of 4D-Var forecasts are smaller except for the 500hPa height error at FT=0, the reason of which is supposed to be that temperature data of radiosondes are used for assimilation, not geopotential height data. Although the error at FT=0 is slightly larger, the errors are smaller at the other forecast times, which indicates that the 4D-Var analysis has a better quality than the routine analysis. The RMSEs of September also have similar improvements over the routine analysis (not shown).

Fig. 3 shows an example of precipitation forecast. In this case, a typhoon (T0115) was approaching and heavy rain areas appeared in the southern coastal region of Japan. The forecast from 4D-Var shows a good agreement with the observation.

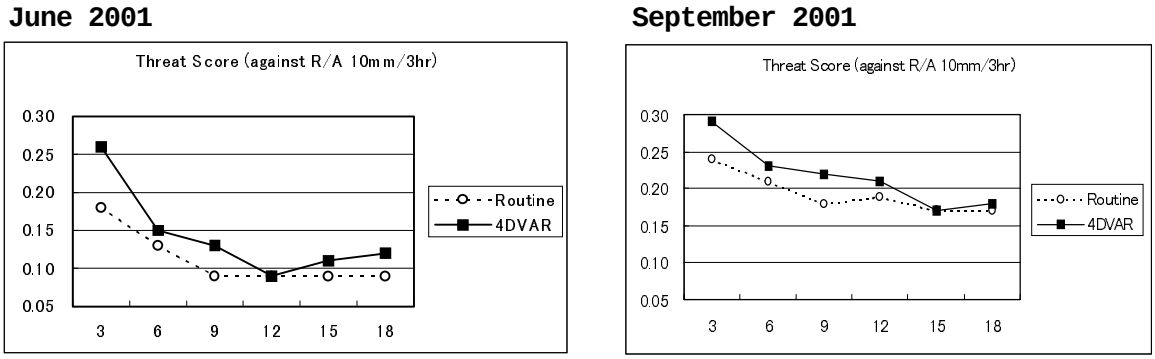


Fig. 1 Threat scores of 10mm/3hour (left: June 2001, right September 2001). Solid lines are scores of 4D-Var and dashed lines are those of routine forecasts.

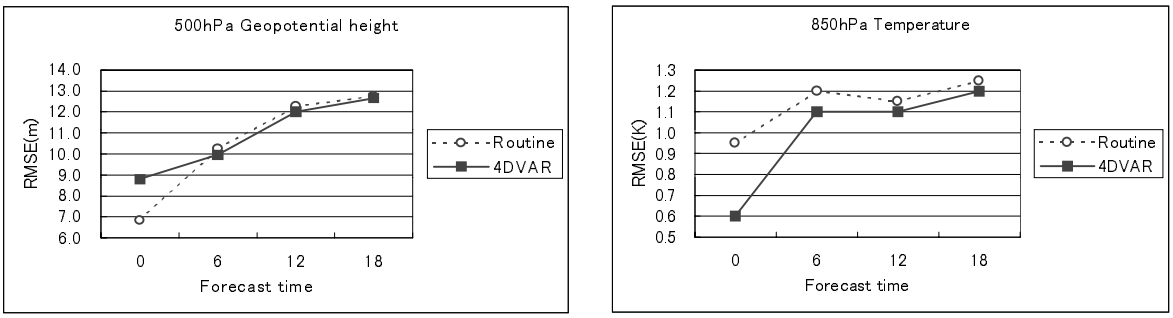


Fig.2 Root mean square errors of 500hPa geopotential height (left) and 850hPa temperature against radiosonde data. Solid lines are scores of 4D-Var and dashed lines are those of routine forecasts.

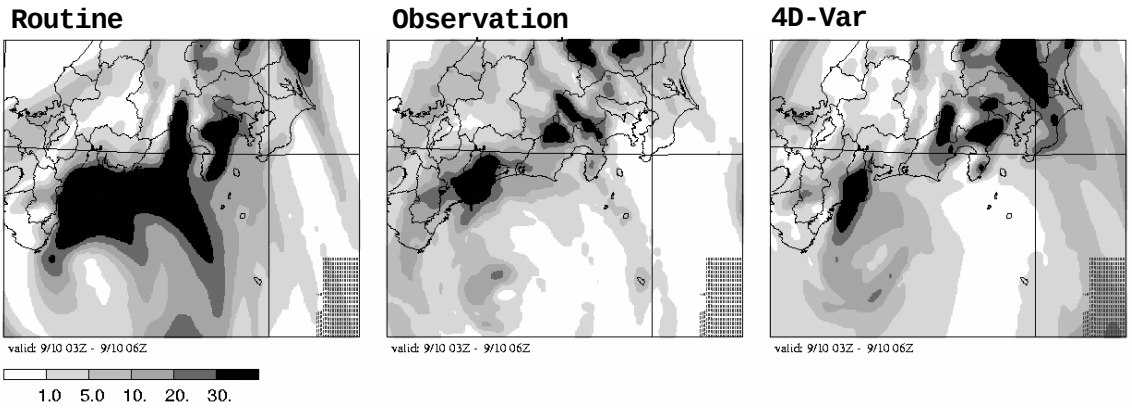


Fig.3 Three-hour precipitation amount in 03-06 UTC 10 September 2001. Initial time of forecasts is 12 UTC 09 September 2001 (FT=15-18).

OZONE DATA ASSIMILATION AT THE MET OFFICE

D.R Jackson and R.W Saunders

Met Office, London Road, Bracknell RG12 2SZ, U.K.
email: david.jackson@metoffice.com

Introduction

It is claimed that ozone assimilation may benefit numerical weather prediction (NWP) via: impact on wind fields through correlations between ozone and potential vorticity (and other dynamic variables); improved assimilation of satellite radiances; improved radiative heating; improved predictions of UV radiation. The impact of including ozone in operational NWP assimilation schemes is currently being investigated at other forecasting centres (see eg Holm et al, 1999; Peuch et al, 2000; Stajner et al, 2001). Here, we describe developments to include ozone within the Met Office's three-dimensional variational assimilation (3D-Var) scheme (Lorenc et al, 2000). Initially, ozone will be analysed using observations from the High Resolution Infrared Radiation Sounder (HIRS) and the Solar Backscatter Ultraviolet spectrometer (SBUV), but it is possible that data from future satellite missions will be incorporated later.

Ozone Assimilation in 3D-Var

Ozone will be assimilated using the Met Office's operational stratosphere - troposphere data assimilation system, which is based on a 40 level version of the Unified Model (UM), with a top level near 0.1 hPa and a horizontal resolution of 2.5^o latitude x 3.75^o longitude. The ozone field is advected by the model forecast winds, but no parametrization of chemical production or loss is included.

The 3D-Var code has been modified to include ozone as a control variable. An important feature of the scheme is the transforms of the control variables that are made in order to eliminate the unwieldy background error covariance from the calculation of the cost function, and to make this calculation much more computationally efficient (Lorenc et al, 2000). To this end, the ozone covariance matrix has been filtered, to allow for different scales in the horizontal, and has been transformed into empirical modes in the vertical, using zonal and seasonal average statistics. For other control variables, a further, physical, transform has been made which eliminates, or greatly reduces, correlations between the variables (see Lorenc et al, 2000). However, for simplicity at this stage it is assumed that ozone is uncorrelated with any of the other analysis variables and thus no physical transform is performed.

To begin with, only ozone data from HIRS (chiefly HIRS channel 9) shall be assimilated. The deep weighting functions of the HIRS channels mean that little information about the vertical ozone structure can be obtained. However, the forward model for HIRS is already in use at the Met Office, and this makes it relatively quick and easy to modify the code to analyse ozone and to test the assimilation system in this basic form, prior to going on to the more sophisticated system detailed below. Furthermore, it has been possible to run tests with ozone added to a 1D-Var version of the HIRS radiance assimilation, in order to assess the most appropriate representation of the ozone background error covariance matrix. Three candidates were tested: 1) background errors provided by ECMWF and calculated using the so-called 'NMC Method' (Parrish and Derber, 1992); 2) a simplified model with no vertical correlations and a variance that is the square of 10 % of the background ozone mixing ratio; 3) a hybrid solution that

uses the vertical correlations from 1) and variances from 2). We examined the ozone increments for these three matrices that resulted when the HIRS channel 9 brightness temperature was reduced by 1 K. The Jacobian indicates the sensitivity of channel 9 to changes in ozone and, encouragingly, all three matrices gave rise to ozone changes near the peak of Jacobian. However, matrices 2) and 3) also produced ozone changes far away from this peak, between 10 and 1 hPa, which is not desirable. For this reason, our initial trials shall use matrix 1). These trials will start shortly.

Future Developments

After completion of the initial trials, the following further refinements to the assimilation system will be added.

- Change the ozone background error covariances to include correlations with other analysis variables. This may lead to improved analyses of both ozone and wind, particularly in the upper troposphere and lower stratosphere.
- Add ozone observations from the SBUV instrument to the assimilation. These observations will provide much more information about the vertical structure of the stratospheric ozone field than HIRS observations can.

The system may eventually be used operationally if it can be demonstrated that it has a positive impact on numerical weather forecasts. Other planned work includes assimilating data from future satellite missions (such as the Atmospheric Infrared Sounder (AIRS) and Envisat) and assessing their impact.

References

Holm, E.V, Untch, A, Simmons, A, Saunders, R, Bouttier, F and Andersson, E (1999) Multivariate ozone assimilation in four-dimensional data assimilation. *SODA Workshop on chemical data assimilation proceedings*, De Bilt, 9-10 Dec. 1998, pp 89-94.

Lorenc, A.C, Ballard, S.P, Bell, R.S, Ingleby, N.B, Andrews, P.L.F, Barker, D.M, Bray, J.R, Clayton, A.M, Dalby, T.D, Li, D, Payne, T.J and Saunders, F.W (2000) The Met. Office global three-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.* **126** 2991-3012.

Parrish, D.F and Derber, J.C (1992) The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Weather Rev.* **120** 1747-1763

Peuch, A, Thepaut, J-N and Pailleux, J (2000) Dynamical impact of total ozone observations in a four-dimensional variational assimilation. *Quart. J. Roy. Meteor. Soc.* **126** 1641-1659

Stajner, I, Riishojgaard, L.P and Rood, R.B (2001) The GEOS ozone data assimilation system: Specification of error statistics. *Quart. J. Roy. Meteor. Soc.* **127** 1069-1094

Assimilation of cloud- and land-affected satellite sounding data at the Data Assimilation Office

Joanna Joiner, Donald Frank, and Arlindo da Silva
NASA Goddard Space Flight Center Data Assimilation Office
(jjoiner@dao.gsfc.nasa.gov)

Satellite data from passive microwave and infrared sounders consistently improve forecasts and analyses in data assimilation systems. However, most numerical weather prediction (NWP) centers use only a small fraction of the data available from these instruments. In particular, NWP centers often exclude data from infrared instruments, which are affected by clouds more than are microwave sensors. Similarly, most NWP centers omit data from land-affected channels.

Sensitive areas for medium-range forecasts are frequently cloudy. Clouds affect ~80% of infrared pixels from the Advanced TIROS Operational Vertical Sounder (ATOVS) flying on NOAA weather satellites. Conservative cloud detection schemes may declare 90% or more pixels as cloud-contaminated.

The next generation of infrared kilo-channel sounders offers more information than the current ATOVS infrared sounding instrument. This next generation includes the Atmospheric Infrared Sounder (AIRS), which will fly on the NASA EOS Aqua satellite. Our ability to use land- and cloud-affected data from these instruments may increase their impact on forecasting capabilities.

Several methods exist for utilizing cloud-affected data in a data assimilation system. These include (1) directly assimilating the cloudy radiances and (2) assimilating cloud-cleared radiances. Direct assimilation of cloudy radiances is very challenging, as it requires reasonably accurate model-generated clouds and a fast and accurate radiative transfer model. At the NASA Data Assimilation Office (DAO), we examined the latter approach.

Assimilating cloud-cleared radiances involves estimating the clear-column radiance that would have been observed in the absence of cloud. We examined the effectiveness of this approach using the DAO¹s next-generation finite-volume Data Assimilation (fvDAS) with a 1D variational radiance assimilation scheme. This system simultaneously performs cloud-clearing and retrieves information about temperature, humidity, ozone, and surface parameters including the surface skin temperature.

The fvDAS experimental setup was at a resolution of 2° latitude × 2.5° longitude for the month of August 1999 with a 2 week spin-up. We conducted a series of experiments using different ATOVS data: 1) DAO CC (includes cloud-cleared data) 2) DAO CLR (clear data only) 3) NESDIS (operational retrievals). One major caveat is that the DAO experiments used the NOAA 15 satellite with the Advanced Microwave Sounding Unit (AMSU) whereas the NESDIS experiment did not.

Figure 1 shows the spatial RMS of the bias and the standard deviation of the radiosonde observed minus 6 hour forecast residuals for heights. Both the DAO CC and DAO CLR have substantially less height bias. The DAO CC has a smaller bias in height in all regions except Asia (NE), where the type of radiosonde used has known temperature bias. Improvements in 5 day forecasts with DAO CC were also achieved. Similar experiments were conducted with and without land-affected channels. A positive but smaller impact was shown on the 6-hour forecast heights.

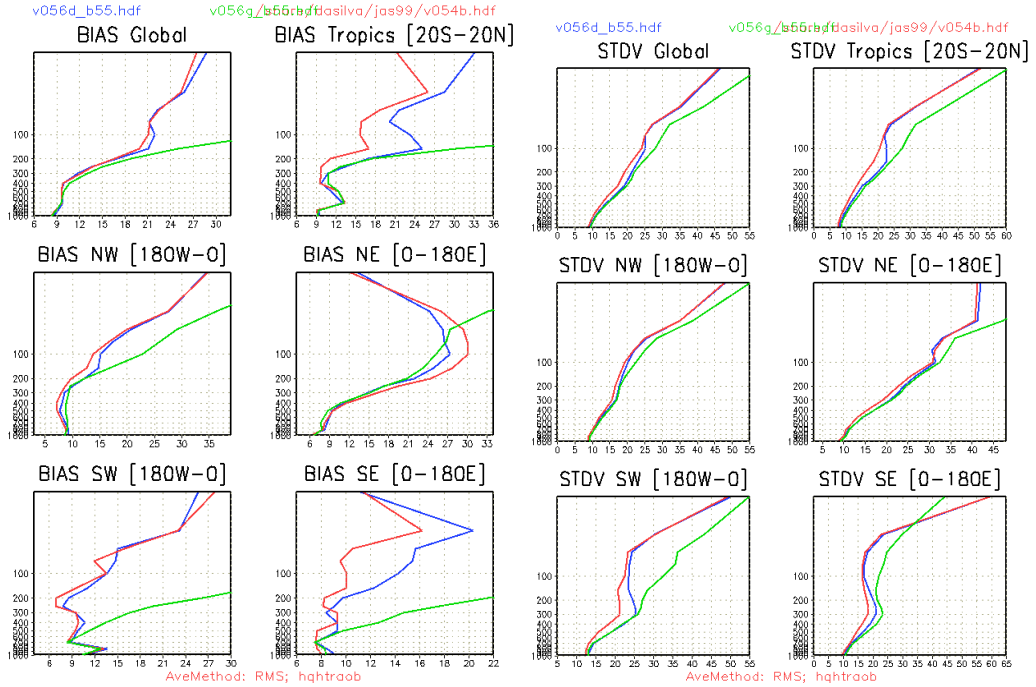


Figure 1: Spatial *RMS of the bias (left) and standard deviation (right) of radiosonde observed heights minus 6 hour forecast residuals averaged over August 1999. Red: DAO CC; Blue: DAO CLR; Green: NESDIS*

BRED VECTORS, LYAPUNOV VECTORS, AND DATA ASSIMILATION

Eugenia Kalnay[1], Matteo Corazza[1,2], and Ming Cai[1]
[1]University of Maryland, College Park, MD 20742, ekalnay@atmos.umd.edu
[2]INFM- DIFI, Universita di Genova, 16146 Genova, Italy

Regional loss of predictability is usually an indication of the presence of a regional instability of the underlying flow, where small errors in the initial conditions (or imperfections in the model) grow to large amplitudes in finite times. The stability properties of evolving flows have been studied using Lyapunov vectors (e.g., Alligood et al, 1996, Ott, 1993, Kalnay, 2002), singular vectors (e.g., Lorenz, 1965, Farrell, 1988, Molteni and Palmer, 1993), and, more recently, with bred vectors (e.g., Szunyogh et al, 1997, Cai et al, 2002). Bred vectors (BVs) are, by construction, closely related to Lyapunov vectors (LVs). In fact, after an infinitely long breeding time, and with the use of infinitesimal amplitudes, bred vectors are identical to leading Lyapunov vectors. In practical applications, however, bred vectors are different from Lyapunov vectors in two important ways: a) bred vectors are never globally orthogonalized and are intrinsically *local in space and time*, and b) they are *finite-amplitude, finite-time vectors*. These two differences are very significant in a dynamical system like the atmosphere whose size is very large. For example, there is "room" in the atmosphere for several synoptic scale instabilities (e.g., storms) to develop independently in different regions (say, North America and Australia), and there are several different possible types of instabilities (such as barotropic, baroclinic, convective, and even Brownian motion), each of them characterized by finite lifetimes. For such a large, complex system, the notion of waiting an infinite time to obtain a single globally dominant Lyapunov vector does not seem to make physical sense.

Bred vectors share some of their properties with leading LVs (Corazza et al, 2001a, 2001b, Toth and Kalnay, 1993, 1997, Cai et al, 2001): 1) Bred vectors are independent of the norm used to define the size of the perturbation. Corazza et al. (2001) showed that bred vectors obtained using a potential enstrophy norm were indistinguishable from bred vectors obtained using a streamfunction squared norm, in striking contrast with singular vectors. 2) Bred vectors are independent of the length of the rescaling period as long as the perturbations remain approximately linear (for example, for atmospheric models the interval for rescaling could be varied between a single time step and 1-2 days without affecting qualitatively the characteristics of the bred vectors).

However, the finite-amplitude, finite-time, and lack of orthogonalization properties of the BVs introduce important differences with LVs:

1) In regions that undergo strong instabilities, the bred vectors tend to be locally dominated by simple, low-dimensional structures. Patil et al (2001) showed that the BV-dim, $\Psi(\sigma_1,\sigma_2,...,\sigma_k)=\left(\sum_{i=1}^{k}\sigma_i\right)^2 \Big/ \sum_{i=1}^{k}\sigma_i^2$ where σ_i are the square roots of the eigenvalues of the local BV covariance matrix, gives a good estimate of the number of dominant directions (shapes) of the local k bred vectors. They showed that the regions with low dimensionality cover about 20% of the atmosphere. They also found that these low-dimensionality regions have a very well defined vertical structure, and a typical lifetime of 3-7 days. The low dimensionality identifies regions where the instability of the basic flow has manifested itself in a low number of preferred directions of perturbation growth.

2) Using a Quasi-Geostrophic simulation system of data assimilation developed by Morss (1999), Corazza et al (2001a, b) found that bred vectors have structures that *closely resemble the errors of the short forecasts used as first guess* (background errors), which in turn dominate the local analysis errors. This is especially true in regions of low dimensionality, which is not surprising if these are unstable regions where errors grow in preferred shapes.

3) The number of bred vectors needed to represent the unstable subspace in the QG system is small (about 6-10). This was shown by computing the local BV-dim as a function of the number of independent bred vectors. Convergence in the local dimension starts to occur at about 6 BVs, and is essentially complete when the number of vectors is about 10-15 (Corazza et al, 2001a). This should be contrasted with the results of Snyder and Joly (1998) and Palmer et al (1998) who showed that hundreds of Lyapunov vectors with positive Lyapunov exponents are needed to represent the attractor of the system in quasi-geostrophic models.

4) Since only a few bred vectors are needed, and background errors project strongly in the subspace of bred vectors, Corazza et al (2001b) were able to develop cost-efficient methods to improve the 3D-Var data assimilation by adding to the background error covariance terms proportional to the outer

product of the bred vectors, thus representing the "errors of the day". This approach led to a reduction of analysis error variance of about 40% at very low cost.

5) The finite amplitude of the BVs provides a natural filter of fast but irrelevant instabilities due to nonlinear stauration. As shown by Lorenz (1996) Lyapunov vectors (and singular vectors) of models including these physical phenomena would be dominated by the fast but small amplitude instabilities, unless they are explicitly excluded from the linearized models.

6) Every bred vector is qualitatively similar to the locally *leading* LV. LVs beyond the leading LV are obtained by orthogonalization after each time step with respect to the previous LVs subspace. The orthogonalization requires the introduction of a norm. Using an enstrophy norm, the successive LVs have larger and larger horizontal scales, and a choice of a stream function norm would lead to successively smaller scales in the LVs. Beyond the first few LVs, there is little qualitative similarity between the background errors and the LVs.

In summary, in a system like the atmosphere with enough physical space for several independent local instabilities, BVs and LVs share some properties but they also have significant differences. BV are finite-amplitude, finite-time, and because they are not globally orthogonalized, they have local properties in space. Bred vectors are akin to the leading LV, but bred vectors derived from different arbitrary initial perturbations remain distinct from each other, instead of collapsing into a single leading vector, presumably because the nonlinear terms and physical parameterizations introduce sufficient stochastic forcing to avoid such convergence. As a result, there is no need for global orthogonalization, and the number of bred vectors required to describe the natural instabilities in an atmospheric system (from a local point of view) is much smaller than the number of Lyapunov vectors with positive Lyapunov exponents. The BVs are independent of the norm, whereas the LVs beyond the first one do depend on the choice of norm: for example, they become larger in scale with a vorticity norm, and smaller with a stream function norm.

These properties of BVs result in significant advantages for data assimilation and ensemble forecasting for the atmosphere. Errors in the analysis have structures very similar to bred vectors, and it is found that they project very strongly on the subspace of a few bred vectors. This is not true for either Lyapunov vectors beyond the leading LVs, or for singular vectors unless they are constructed with a norm based on the analysis error covariance matrix (or a bred vector covariance). The similarity between bred vectors and analysis errors leads to the ability to include "errors of the day" in the background error covariance and a significant improvement of the analysis beyond 3D-Var at a very low cost (Corazza et al, 2001b).

References

Alligood K. T., T. D. Sauer and J. A. Yorke, 1996: Chaos: an introduction to dynamical systems. Springer-Verlag, New York.

Buizza R. et al, 1993: Computation of optimal unstable structures for numerical weather prediction models. Tellus, 45A, 388-407.

Cai, M., E. Kalnay and Z. Toth, 2002: Potential impact of bred vectors on ensemble forecasting and data assimilation in the Zebiak-Cane model. Accepted in J of Climate.

Corazza, M., E. Kalnay, D. J. Patil, R. Morss, M. Cai, I. Szunyogh, B. R. Hunt, E. Ott and J. Yorke, 2001: Use of the breeding technique to determine the structure of the "errors of the day". Submitted to Nonlinear Processes in Geophysics.

Corazza, M., et al, 2001: Use of the breeding technique in the estimation of the background error covariance matrix for a quasigeostrophic model. AMS Symp. on Observations, Orlando, FA, 14-17 Jan 2002.

Farrell, B., 1988: Small error dynamics and the predictability of atmospheric flow, J. Atmos. Sciences, 45, 163-172.

Kalnay, E 2002: Atmospheric modeling, data assimilation and predictability. Chapter 6. Cambridge University Press, UK. In press.

Kalnay E and Z Toth 1994: Removing growing errors in the analysis. Preprints, Tenth Conference on Numerical Weather Prediction, pp 212-215. Amer. Meteor. Soc., July 18-22, 1994.

Lorenz, E.N., 1965: A study of the predictability of a 28-variable atmospheric model. Tellus, 21, 289-307.

Lorenz, E.N., 1996: Predictability- A problem partly solved. Proc. of the ECMWF Seminar on Predictab., Reading, England, Vol. 1 1-18

Molteni F. and TN Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. Q. J. Roy. Meteorol. Soc. 119, 269-298.

Morss, R.E.: 1999: Adaptive observations: Idealized sampling strategies for improving numerical weather prediction. Ph.D. Thesis, Massachussetts Institute of Technology, 225pp.

Ott, E., 1993: Chaos in Dynamical Systems. Cambridge University Press. New York.

Palmer, TN, R. Gelaro, J. Barkmeijer and R. Buizza, 1998: Singular vectors, metrics and adaptive observations. J. Atmos Sciences, 55, 633-653.

Patil, DJ, BR Hunt, E Kalnay, J. Yorke, and E. Ott, 2001: Local low dimensionality of atmospheric dynamics. Phys. Rev. Lett., 86, 5878.

Patil, DJ, I. Szunyogh, BR Hunt, E Kalnay, E Ott, and J. Yorke, 2001: Using large member ensembles to isolate local low dimensionality of atmospheric dynamics. AMS Symposium on Observations, Data Assimilation and Predictability, Preprints volume, Orlando, FA, 14-17 January 2002.

Snyder, C. and A. Joly, 1998: Development of perturbations within growing baroclinic waves. Q. J. Roy. Meteor. Soc., 124, pp 1961.

Szunyogh, I, E. Kalnay and Z. Toth, 1997: A comparison of Lyapunov and Singular vectors in a low resolution GCM. Tellus, 49A, 200-227.

Toth, Z and E Kalnay 1993: Ensemble forecasting at NMC - the generation of perturbations. Bull. Amer. Meteorol. Soc., 74, 2317-2330.

Toth, Z and E Kalnay 1997: Ensemble forecasting at NCEP and the breeding method. Mon Wea Rev, 125, 3297-3319.

A model to calculate the covariances of homogeneous isotropic stochastic fields of forecast errors

Dr. Ekaterina G. Klimova

Institute of Computational Technologies,
Siberian Branch, Russian Academy of Sciences,
Ac. Lavrentjev Avenue, 6
Novosibirsk 630090 Russia

klimova@ict.nsc.ru

Application of the Kalman filter theory to the problem of meteorological data assimilation for the present-day prognostic models is a task whose realization is difficult because of a high order of covariance matrices occurring during this process. At the same time, the atmospheric dynamics is described by particular prognostic equations whose characteristics are well investigated. Besides in the turbulent theory the known analytical equations for structural functions of fields of temperature and wind are received.

A simplified model for calculating the covariance matrices of homogeneous isotropic stochastic fields of forecast errors is proposed in [4]. For a local covariance of forecast errors between two specified points an analytical equation is obtained that is made by analogy with the derivation of equations for the structural functions in the turbulent theory.

Let's name as an error of the forecast a deviation of forecast fields from "true", thus we shall consider, that the errors of the forecast depend only on errors in the initial data. Let "true" state of an atmosphere is described by the baroclinic adiabatic model of atmosphere for a region based on the primitive equations, in (x, y, p) coordinate system [7].

The important feature of numerical algorithm of realization of the given model is the application of idea of G.I. Marchuk of splitting of the dynamic operator of mathematical model on physical processes. In this case at the first stage the system of the equations of advection of mass and temperature along trajectories of motion and on second - system of the equations of adaptation of a wind and geopotential fields are solved.

Let's consider system of the prognostic equations on a time interval (t_k, t_{k+1}). Let $\tilde{u} = u(t_k), \tilde{v} = v(t_k)$. Let's designate through $\delta u, \delta v, \delta\tau, \delta z$ errors of the forecast of wind field (u, v, τ) and geopotetial z, accordingly. The equations for covariances of the forecast errors of a field δu between two points with coordinates $a_1 = (x_1, y_1, p_1)$ and $a_2 = (x_2, y_2, p_2)$ are obtained by analogy with the derivation given in [6]:

$$\frac{\partial \overline{\delta u_1 \delta u_2}}{\partial t} + \tilde{u}_1 \frac{\partial \overline{\delta u_1 \delta u_2}}{\partial x_1} + \tilde{v}_1 \frac{\partial \overline{\delta u_1 \delta u_2}}{\partial y_1} - f\overline{\delta v_1 \delta u_2} + \tilde{u}_2 \frac{\partial \overline{\delta u_1 \delta u_2}}{\partial x_2} +$$

$$+\tilde{v}_2 \frac{\partial \overline{\delta u_1 \delta u_2}}{\partial y_2} - f\overline{\delta v_2 \delta u_1} + g\frac{\partial \overline{\delta z_1 \delta u_2}}{\partial x_1} + g\frac{\partial \overline{\delta z_2 \delta u_1}}{\partial x_2} = 0.$$

The equations for covariances of other fields are similar.

From the turbulent theory it is known, that the homogeneous isotropice field of wind velocity in an incompressible fluid does not correlate with any scalar field [6].

Hence, if we assume that stochastic fields of forecast errors are homogeneous and isotropic, the terms with a gradient of geopotential in the equations for covariances are rejected.
If to follow idea of splitting of the dynamic operator of mathematical model on a step of advection on trajectories and step of adaptation of the meteorological fields, at a stage of adaptation in the appropriate equations the terms with a gradient of geopotential thus should be rejected, and so in system of the equations of adaptation there are only terms with the Coriolis force.
The account of Coriolis force means turn of a vector of horizontal wind velocity $U = (u, v)$, and, as in homogeneous isotropic stochastic field all density of distribution by definition do not depend on any turn, those terms with Coriolis force in system of the equations of adaptation should be rejected too. Thus, dynamics of covariances of errors of the forecast, in case they are homogeneous and isotropic, on a small time interval (t_k, t_{k+1}) is described by model of advection of substation on trajectories of particles.
The model developed is used to calculate the covariance of homogeneous isotropic stochastic fields of forecast errors.
Such simplification does not lower essentially order of considered matrices, so the simplified model was obtained under the following additional assumptions:

- the calculation of the forecast error covariances is carred out for the vertical normal modes of the prognistic model;
- the calculation of the forecast error covariances is based on the assumption that the errors of vertical normal modes do not correlate;
- the background fields of velocity components in the advection operator do not depend on vertical coordinate p (i.e., the background flow is close to barotropic).

The main results are published in [1-5]. This work is supported by the Grant of the Government Support for the Leading Scientific Schools N 00-15-98543 and the Integration Grant of SB RAS N 64.

References

1. E.G.Klimova. Algorithm of the data assimilation of meteorological observations based on the extended suboptimal Kalman filter. - Russian Meteorology and Hydrology, 1997, N 11.
2. E.G.Klimova. Asimptotic behavior of the data assimilation scheme, based on algorithm of a Kalman filter. - Russian Meteorology and Hydrology, 1999, N 8.
3. E.G.Klimova. Simplified models for calculation of covariance matrices in the Kalman filter algorithm. - Meteorologiya i Hydrologiya, 2000, N 6. (Translated on English: Russian Meteorology and Hydrology).
4. E.G.Klimova. A model to calculate the covariances of homogeneous isotropic stochastic fields of forecast errors. - Meteorologiya i Hydrologiya, 2001, N 10.
5. E.G.Klimova. Model for calculation the forecast error covariances in the Kalman filter algorithm based on the full quations. - Meteorologiya i Hydrologiya, 2001, N 11.
6. Monin A.S., Yaglom A.M. Statistical Fluid Mechanics, vol.1, Cambridge (Mass., USA) : MIT Press, 1971.
7. G.S.Rivin. Numerical modelling of background atmospheric processes and the problem of aerosol transport in Siberian region.- Atmospheric and Oceanic Optics, 1996, V. 9, N. 6.

DATA ASSIMILATION AND WEATHER REGIMES IN A THREE-LEVEL QUASI-GEOSTROPHIC MODEL

D. Kondrashov*, M. Ghil, K. Ide, University of California, Los Angeles,
and R. Todling, NASA, Goddard Space Flight Center.

1. INTRODUCTION

Extended-range weather prediction depends in a crucial way on skill at forecasting the duration of a blocking event or other persistent anomaly that is under way at initial forecast time. The ability to forecast the subsequent onset of another persistent anomaly—after the break of the current one—has proven even more elusive.

To address this crucial problem, we study the application of advanced data assimilation methods on predicting the transitions between atmospheric weather regimes. Marshall and Molteni's (1993) three-level quasi-geostrophic (QG) model in spherical geometry has been shown to have a fairly realistic climatology and exhibit multiple regimes that bear some resemblance to those found in observations. Using this model, we study the transition mechanism between such regimes.

2. CLUSTERING ANALYSIS

The dataset for analysis was obtained from an 18,000-day perpetual-winter simulation of our QG model — whose three levels are at 200 mb, 500 mb and 800 mb — on a T21 (64 x 34) grid. In order to examine the phase-space structure of atmospheric dynamics in such a high-dimensional system, it is necessary to reduce the dataset's dimensionality. For this purpose, we apply empirical orthogonal function (EOF) analysis to the unfiltered 500-mb level streamfunction anomalies in the Northern Hemisphere (NH), where the gridded data points are weighted by the cosine of their latitudes. The leading 10 EOFs are responsible for 47% of the variance of the dataset, the first mode capturing 11%, and the second 6%.

In order to objectively identify weather regimes in the QG model simulation, we apply two independent clustering techniques and compare the results (see Table 1 of Ghil and Robertson 2001). One technique is the *k*-means algorithm used by Michelangeli et al. (1995) and the other is the Gaussian mixture model used by Smyth et al. (1999) for the classification of NH weather regimes in observed geopotential height fields.

For a given number *d* of leading EOFs, both techniques provide the number of clusters *k* and the cluster centroids in a *d*-dimensional subspace of the model's phase space. We want each cluster to correspond to a weather regime of the QG model's physical space. Therefore it is critical for our study to optimize the classification into clusters over various subspaces.
The classifiability index of the *k*-means algorithm measures the stability of the cluster solutions as a function of *k*, across different initial (random) seeds of the algorithm, based on the correlation between the cluster centroids. The classifiability index of the QG model simulation is very high for both $k = 4$ and $k = 3$. We thus conclude that the *k*-means algorithm alone cannot identify the optimal clusters of the QG model

The Gaussian mixture model uses a linear combination of *k* Gaussian density functions. Unlike the *k*-means algorithm, each data point in the *d*-dimensional space can have a degree of membership in several clusters, depending on its position with respect to the centroid and the weight of a cluster (Smyth et al. 1999).

According to the cross-validated log-likelihood criterion the mixture model consistently gives $k = 6$, which is higher than the values $k = 3$ or 4 obtained by the *k*-means algorithm. Hannachi and O'Neill (2001) found that the Gaussian mixture model tends to overfit the clusters when the distribution of the data is not Gaussian. This is the case here, too. In fact, when using either half of the entire dataset, the cross-validated log-likelihood suggests a higher probability for $k = 4$ and 5 than for $k = 6$

Next, we compare the anomaly maps of the centroids produced by the two methods (see, for instance, Table 2 in Robertson and Ghil 1999).

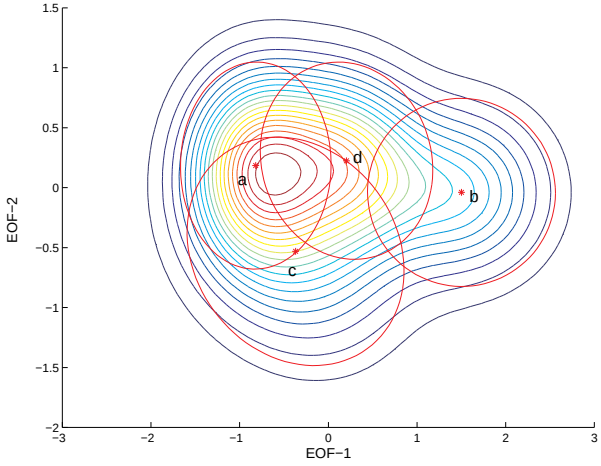


Fig.1: Probability density function of the QG model's 500-mb stream function field, as estimated by the mixture model for $k = 4$ and $d = 5$ and projected onto the plane spanned by EOF-1 and EOF-2.

To do so, we compute the pattern correlation coefficients of the cluster centroids in physical space for pairs of visually similar streamfunction anomaly maps produced by the two clustering techniques and compare the results for different values of *k*. We obtain the maps that correspond to the cluster centroids in the *d*-dimensional subspace by computing the EOF expansion of the 500-mb streamfunction field, *i.e.* the QG model's second level, truncated at $d = 10$.
The best agreement was found between the two methods for

k = 4 for all values of d. We conclude, therefore, that k = 4 yields the optimal set of clusters for our QG model. The 500-mb streamfunction anomaly maps obtained by the k-means method for the cluster centroids shown in Figure 1 are plotted in Figure 2.

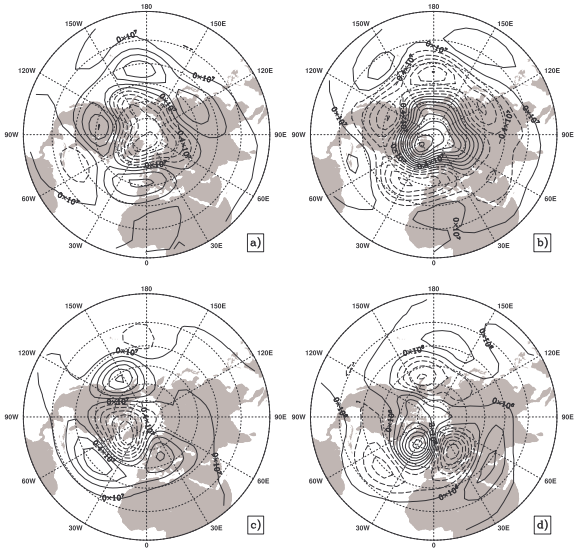


Fig. 2: Streamfunction anomaly maps of cluster centroids obtained by the k-means algorithm, for d = 5 and k = 4. Land masses are shaded.

Each of the regimes in Figure 2 represents one of the opposite phases of two spatial patterns. Clusters **c** and **d** capture the two extreme phases of the North-Atlantic Oscillation (NAO), while their patterns outside the Atlantic sector complete a NH wavenumber-three pattern. Clusters **a** and **b** have a central feature that extends over the whole Arctic and is rather zonally symmetric, with a substantial wavenumber-four component. It thus has certain features in common with the Arctic Oscillation (Thompson and Wallace 1998) and with Mo and Ghil's (1988) North-South seesaw. We denote these four regimes by AO^+ (panel *a*), AO^- (panel *b*), NAO^+ (panel *c*) and NAO^- (panel *d*).

3. PREFERRED TRANSITIONS

Using the clustering results for k = 4, the Markov chain of transitions between the four regimes is obtained. In the d-dimensional space, each weather regime is defined by the ellipsoid of covariance around the centroid, whose semi-axes equal the corresponding eigenvalues, as shown in Figure 1. A data point is assigned to a weather regime if it lies within the corresponding ellipsoid. If a data point belongs to several ellipsoids, we assign it according to the maximum probability value.

The preferred transition paths between the four regimes are shown in the Table 1.

	AO^+	AO^-	NAO^+	NAO^-
AO^+	**0.24**	0.01	**0.34**	**0.40**
AO^-	0.02	**0.51**	0.07	**0.39**
NAO^+	**0.49**	0.06	0.25	**0.21**
NAO^-	**0.40**	**0.17**	**0.22**	**0.21**

Table 1: Transition probabilities estimated using mixture models regimes for d=5; transitions that are significant at 95% are in bold (Vautard et al. 1990).

Three of the four regimes have highly statistically significant reinjection rates. Aside from these bold diagonal entries in the table, we note that a strong preferential path leads from both the zonal sectorial regime NAO^+ and the blocked NAO^- to the high-index hemispheric regime AO^+, as well as from the NAO^- to the low-index AO^-. The opposite transitions from AO^+ to NAO^+ and NAO^-, and from AO^- to NAO^- are also highly significant. A preferential cycle connects, moreover, the sectorially blocked and zonal regimes NAO^- to NAO^+ and back. The hemispheric regimes AO^- and AO^+, however, are not directly connected to each other. This is because AO^- and AO^+ are separated by NAO^- and NAO^+ in the probability density function (Fig. 1).

4. DATA ASSIMILATION USING PSAS

We use NASA Goddard's Physical-Space Statistical System (PSAS; Cohn et al. 1998) data assimilation framework to carry out identical-twin experiments with our QG model. The purpose of these experiments is to clarify the physical mechanisms of the regime transitions captured in Table 1.

Synthetic observations are simulated to correspond to both conventional and satellite networks. Their effects on pinpointing the transitions between regimes and capturing their causal mechanisms are evaluated. Implications of observing system design on extended-range prediction in the model are discussed.

More complete version of this communication can be found at http://www.atmos.ucla.edu/tcd/MG/mg_ref_preprints.html .

5. REFERENCES

Cohn, S. E., da Silva, A., Guo, J., Sienkiewicz M., and Lamich D., 1998: Assessing the effects of data selection with the DAO Physical-space Statistical Analysis System. *Mon. Wea. Rev.*, **126**, 2913-2926.

Ghil, M., and Robertson, A. W., 2001: "Waves" vs. "particles" in the atmosphere's phase space: A pathway to long-range forecasting? *Proc. Natl. Acad. Sci.*, accepted.

Hannachi, A., and O'Neill, A., 2001: Atmospheric multiple equilibria and non-Gaussian behaviour in model simulations. *Q. J. R. Meteorol. Soc.*, **127**, 939-958.

Marshall, J., and Molteni, F., 1993: Towards a dynamical understanding of planetary-scale flow regimes. *J. Atmos. Sci.*, **50**, 1792-1818.

Michelangeli, P.A., Vautard, R., and Legras, B., 1995: Weather regimes: recurrence and quasi-stationarity. *J. Atmos. Sci.*, **52**, 1237-1256.

Mo, K., and Ghil, M., 1988: Cluster analysis of multiple planetary flow regimes, *J. Geophys. Res.*, **93D**, 10927-10952.

Robertson, A. W., and Ghil, M., 1999, Large-scale weather regimes and local climate over the western United States. *J. Climate*, **12**, 1796-1813.

Smyth, P., Ide, K. and Ghil, M., 1999: Multiple regimes in Northern Hemisphere height fields via mixture model clustering. *J. Atmos. Sci.*, **56**, 3704-3723.

Thompson, D. W., and Wallace, J. M., 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.,* **25,** 1297-1300.

Vautard, R., Mo, K. C., and Ghil, M., 1990: Statistical significance test for transition matrices of atmospheric Markov chains, *J. Atmos. Sci.*, **47**, 1926-1931.

Case Studies of the Impact of Scatterometer Winds on the Analysis and Forecasting of Tropical Cyclones using a Non-hydrostatic Model

C.C. Lam and Edwin S.T. Lai
Hong Kong Observatory, Hong Kong, China
nwp@hko.gov.hk

Experiments to study the impact of scatterometer winds on forecasting tropical cyclones were carried out using the Advanced Regional Prediction System (ARPS), a non-hydrostatic modelling system originally developed by the Center for Analysis and Prediction of Storms in the United States (Xue *et al*., 2000). Data were assimilated using the Bratseth method of successive correction type (Brewster, 1996). Besides conventional and QuikSCAT data, local observations from radar, automatic weather stations and wind profilers were also ingested. The model was run at 6-km resolution in the inner domain and was one-way nested to the outer 30-km model; both were set up with 40 vertical levels.

Tropical cyclones Utor (0104) and Yutu (0107) over the South China Sea in 2001 were used as test cases. The radii of strong, gale or storm force winds in different quadrants were objectively derived from the model's analyzed and predicted surface wind fields. In the case of Utor, the inclusion of QuikSCAT data produced a much more realistic distribution of high winds in the model analysis, particularly in the depiction of gale force winds over the northwestern quadrant of the cyclone circulation (Fig. 1). The onset of northwesterly gales in Hong Kong later that evening was consistent with the analysis containing QuikSCAT information. In the case of Yutu, unreliable QuikSCAT data in the proximity of land could have led to the displacement of the analyzed cyclone centre from its best track position prior to Yutu's landfall. This shows that rigorous quality control of QuikSCAT data is essential, especially when cyclones approach the coastal region.

Effects of tropical cyclone bogus data on the initial vortex representation in the presence of QuikSCAT data were also examined. The resultant analyses were quite sensitive to the inclusion of bogus data in some cases; e.g. the location and the strength of maximum winds showed significant differences in the analysis of Utor on the morning of 5 July 2001 (Fig. 2). It poses the question of whether or not bogus data should be dispensed with as more frequent and wider coverage of scatterometer wind data become available in the future.

Although improvement in surface wind distribution was found in the model analysis with the ingestion of QuikSCAT wind data, no significant difference was observed in the model forecasts more than a few hours ahead. For a more effective application of scatterometer winds in the model prediction of tropical cyclones, incorporation of retrieved sea surface pressure data into the model analysis will be explored.

References

Brewster, K., 1996: Implementation of a Bratseth analysis scheme including Doppler radar. Preprints, *15th Conf. on Weather Analysis and Forecasting*, AMS, Boston, 92-95.

Xue, M., K.K. Droegemeier and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) – A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. *Meteor. Atmos. Phy.*, **75**, 161-193.

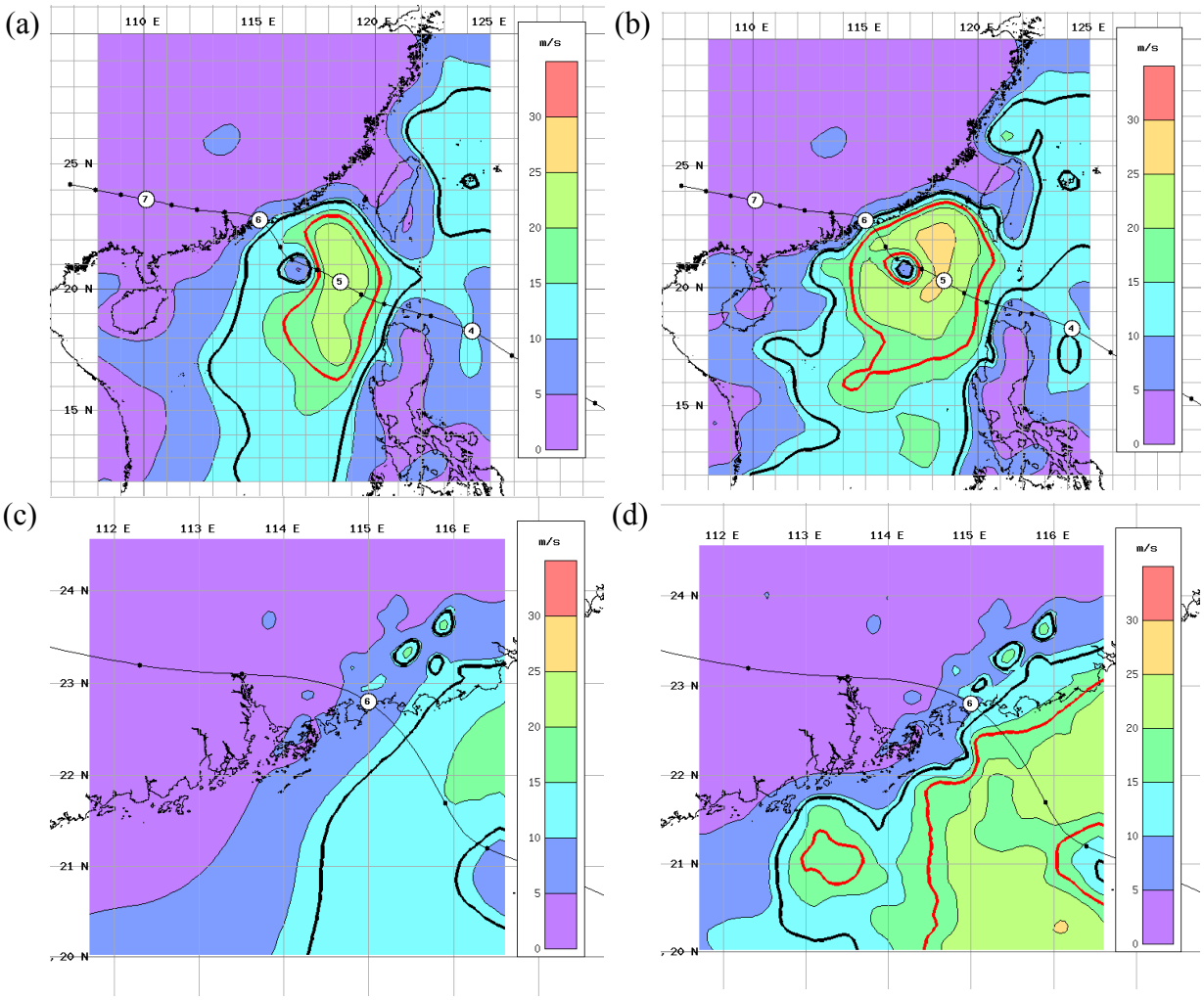


Fig.1 ARPS surface wind analyses for 12 UTC 5 July 2001: without QuikSCAT (left panels) and with QuikSCAT (right panels). Model resolution is 30 km in (a), (b) and 6 km in (c), (d). Isotach contours of 11.5 m/s (strong wind) and 17.5 m/s (gale) are highlighted in bold black and red respectively. Six-hourly best track positions of Utor analyzed by the Hong Kong Observatory are dotted along the solid line, with dated circles marking the daily positions at 00 UTC.

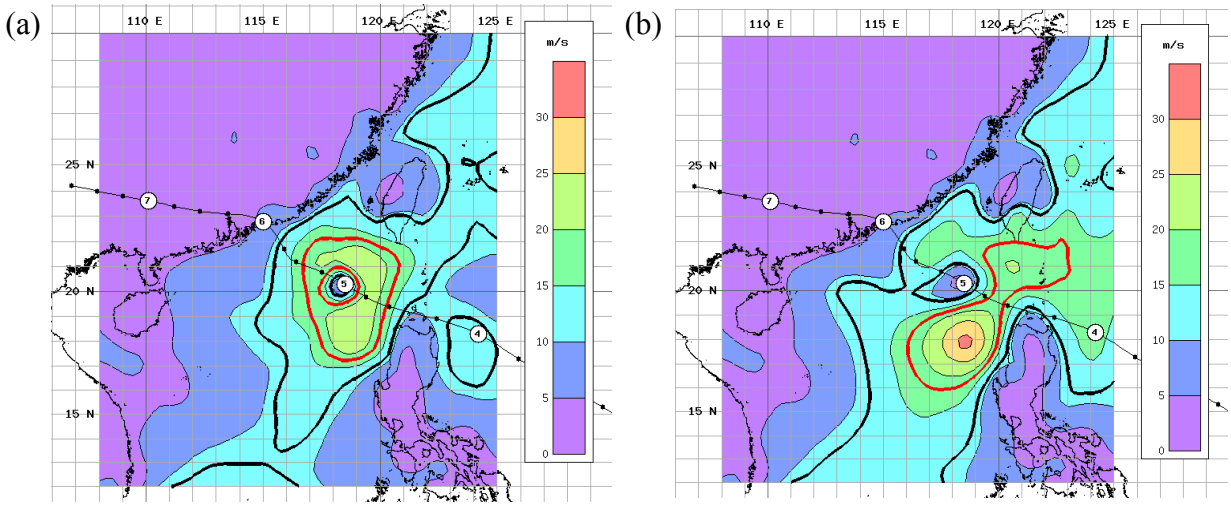


Fig.2 ARPS surface wind analyses for 00 UTC 5 July 2001: (a) with TC bogus and (b) without TC bogus. QuikSCAT data were used in both analyses.

Developments in the Met Office 3DVar Scheme for the UK Mesoscale Model

by
B. Macpherson, S. A. Harcourt, N. B. Ingleby, R. J. Renshaw,
A. J. Maycock, B. V. Chalcraft, S. Anderson, M. Sharpe, C. A. Parrett
NWP Division, Met Office, Bracknell, UK. (email: bruce.macpherson@metoffice.com)

Since its operational implementation in October 1999, described by Andrews et al. (2000), the Met Office Mesoscale 3DVar scheme (Mes 3DVar) has undergone two main upgrades. The first was implemented in June 2000, the second in February 2001.

Of the numerous ingredients in the *June 2000 package*, the most significant was an improved representation of forecast error covariances. The covariance statistics are based on summer and winter data, instead of just winter data, as previously. They use 'Rotated Kinetic Energy Vertical Modes', which replaced unrotated streamfunction and velocity potential vertical modes. This resulted in sharper vertical wind correlations in better agreement with observation. An error in the horizontal spectrum was corrected, which gave unrealistically short horizontal scales, especially in the upper-air. Horizontal scales for wind and temperature were increased further, based on analysis of forecast difference statistics. The streamfunction and velocity potential variances were scaled to give a less divergent wind increment, which should be retained better by the model. Other ingredients included improved observation processing of surface pressure, as already implemented in the global model, and restored use of hourly synops, taking advantage of the availability of time interpolated model background values.

A trial of this package showed a systematic slight improvement in precipitation forecasts at each threshold and forecast time up to t+24 (Figure 1).

The *February 2001 package* had several components: introduction of a soil temperature increment, assimilation of selected wind profiler data and assimilation of higher density Meteosat wind data. The soil temperature increment is equal to the atmospheric temperature increment at the lowest model level. It is added to the surface temperature, T_* and the top deep soil temperature level. Objective verification showed a systematic improvement in screen temperature forecasts up to t+24 (Figure 2). Screen relative humidity errors were also reduced.

Wind profiler data is thinned in time to one report per hour. Monitoring statistics have guided the data selection. The Aberystwyth profiler (03501) is used above 700hPa. The 4 UK boundary layer profilers at Aber (03500), Camborne (03807), Dunkeswell (03840) and Wattisham (03591) are used below 500 hPa. The Swedish weather radar VAD profiler at Leksand (02430) is used below 500 hPa, along with the profilers at Cabauw (06348), Basel (06601) and La Ferte Vidame (07112, above 700 hPa). For quality control, the initial probability of gross error, PGE = 0.05 generally, but is higher (0.1) for La Ferte Vidame and Basel.

Higher density Meteosat wind data have replaced low resolution 6-hourly NESDIS SATOB satellite winds. The EUMETSAT data are available every 90 minutes. Coverage after quality control is improved from ~20 to 100+ data per run, including more in the Biscay area. Most of the extra data are in the 100-400hPa band.

Reference

Andrews, P.L.F., Chalcraft, B.V., Harcourt, S.A., Macpherson, B., Maycock, A., 2000: The UK Met Office Mesoscale 3D-Var Data Assimilation Scheme. CAS/JSC WGNE Research Activities in Atmospheric and Oceanic Modelling, Report No 30, pp 1.3-1.4

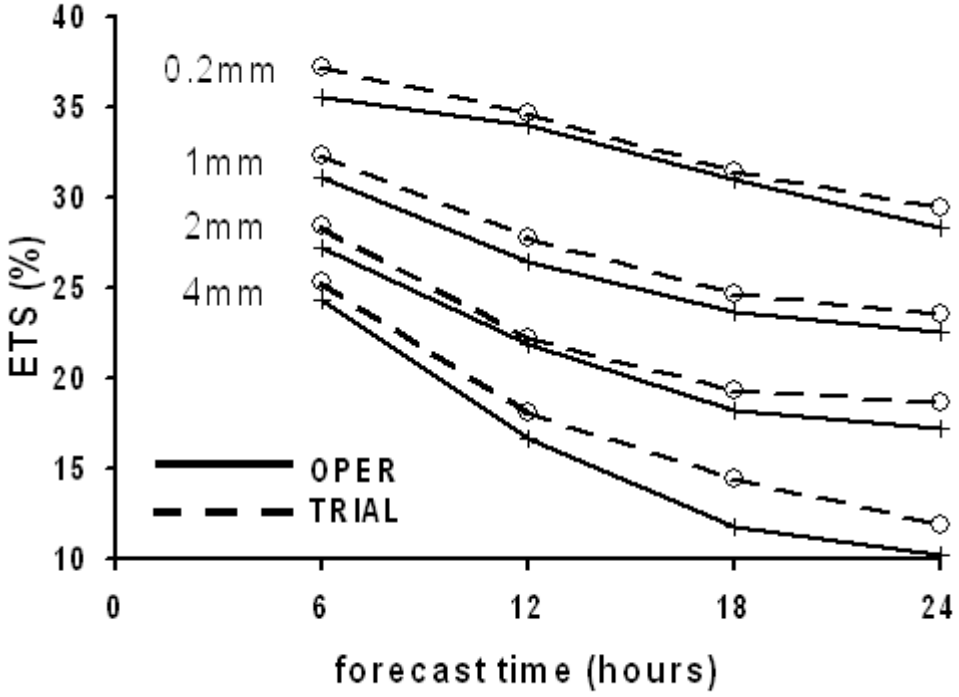


Figure 1: Verification of 6-hourly precipitation accumulations v radar data (at 4 different thresholds) for the 'June 2000 Mes 3DVar package', labelled 'TRIAL', against operational forecasts, labelled 'OPER'. Equitable Threat Score (ETS) as function of forecast time, averaged over 35 forecasts from April 2000.

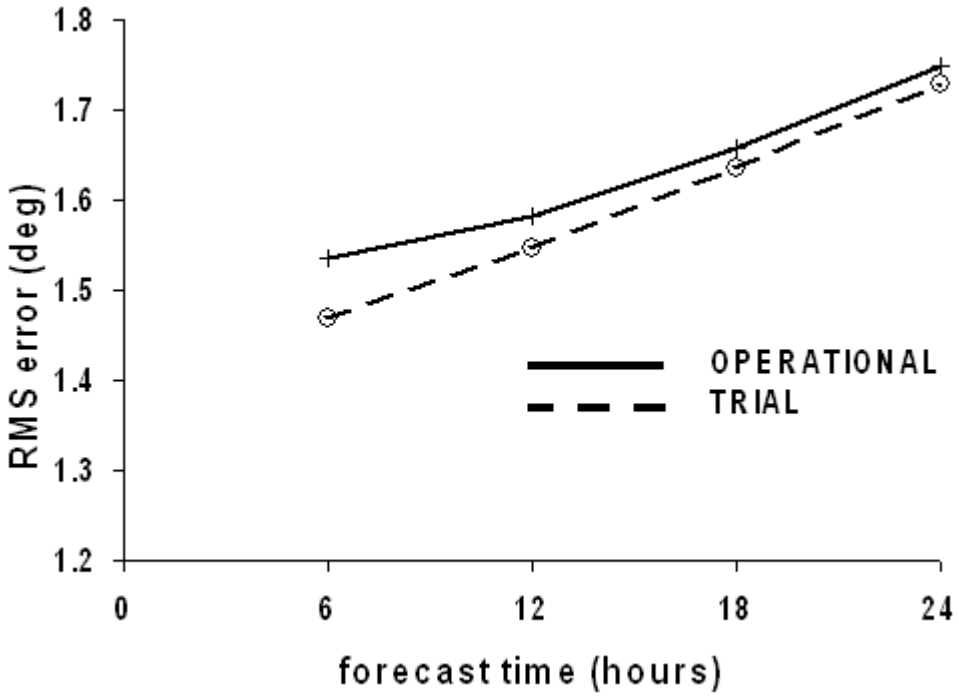


Figure 2: Root-mean-square error in screen temperature as a function of forecast time for the 'February 2001 Mes 3DVar package', labelled 'TRIAL', against OPERATIONAL forecasts. The results are averaged over 80 forecasts in a 3 week period in December 2000 and January 2001. Verification at all UK stations.

Incremental Digital-Filtering Initialization of GME in Vertical Mode Space

D. Majewski and W. Wergen, Deutscher Wetterdienst, 63067 Offenbach, Germany
detlev.majewski@dwd.de and werner.wergen@dwd.de

Since October 4, 1999 an incremental digital filtering initialization (IDFI, Lynch, 1997) is applied during the data assimilation of the operational icosahedral-hexagonal grid point model GME (Majewski et al., 2002). The IDFI should only work on the analysis increments, thus the balanced first guess of the model should not be altered at all but IDFI should modify the analysed fields only in those regions where observations have disrupted this balance.

The IDFI consists of the following three steps

- A DFI (Digital Filtering Initialization) is performed on the 6-hour first guess fields → $(FG)_{DFI}$. The DFI consists of a 3-h adiabatic backward integration of GME, followed by a 3-h diabatic forward one.
- A DFI is performed on the analysed fields → $(ANA)_{DFI} = (FG)_{DFI} + (INCR)_{DFI}$, where INCR is the analysis increment. In regions without any observations the analysis increment vanishes.
- The incrementally initialized analysis is defined as $(ANA)_{IDFI} = FG + (ANA)_{DFI} - (FG)_{DFI} = FG + (INCR)_{DFI}$.

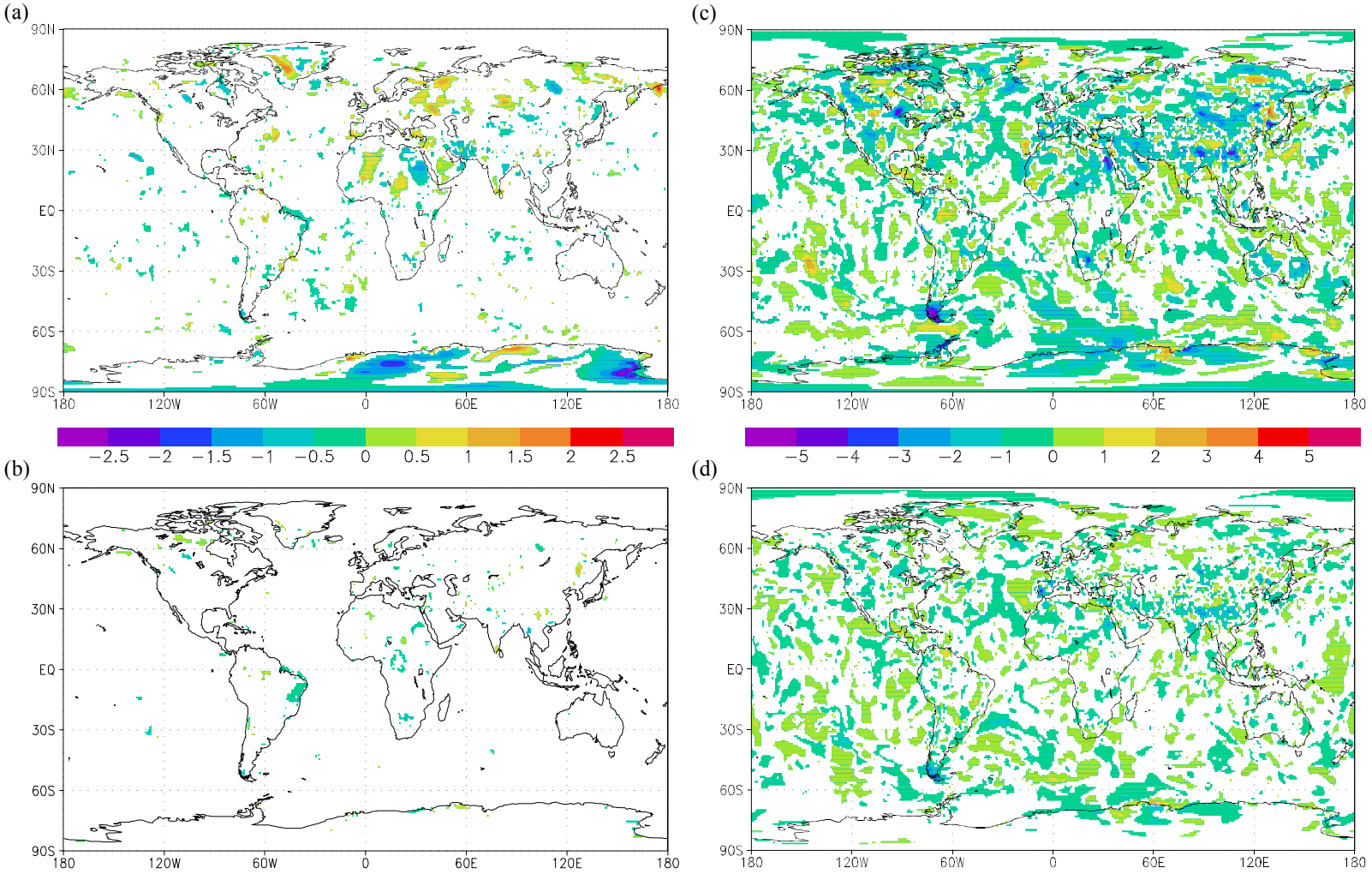
Thus in regions without observations the balanced FG of GME should not be altered.

But even though IDFI is quite effective in removing the initial noise from the forecast, the impact of the initialization on the analysed fields seems to be unreasonably large in some data sparse areas like the oceanic boundary layer or the free atmosphere in convectively unstable situations. The modification of the analysed fields by the IDFI in these regions are mostly due to physical processes like turbulence or convection in the 3-h diabatic forward step of the filtering procedure. If data lead to a modified vertical stability compared to the first guess fields, convective and turbulent processes may be different between the steps $(FG)_{DFI}$ and $(ANA)_{DFI}$. Thus $(INCR)_{DFI} = (ANA)_{DFI} - (FG)_{DFI}$ will not only reflect the changes of the analysis which are necessary to control noise in the forecast but also changes due to the adaptation of the model atmosphere to the modified stratification.

To further restrict the impact of IDFI on the analysed fields to noise control alone, the filtering, i.e. the steps $(ANA)_{DFI}$ and $(FG)_{DFI}$ is performed in vertical mode space. Only the external mode plus the first nine internal ones are filtered, all higher internal modes which represent e.g. boundary layer processes are taken unchanged from the analysed values. Moreover, relative humidity and cloud liquid water are reset to their analysed values, too, to reduce the impact of IDFI on the hydrological cycle. This modified IDFI scheme has the same properties regarding noise control as the original one but the impact on fields like temperature at 850 hPa and low level winds (Fig. 1 a, b, c and d) is smaller. Since June 12, 2001 the IDFI in vertical mode space is used operationally.

Lynch, P., 1997: The Dolph-Chebyshev window: A simple optimal filter. Mon. Wea. Rev., 125, 655-660.

Majewski, D., D. Liermann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, W. Wergen and J. Baumgardner, 2002: The operational global icosahedral-hexagonal grid point model GME: Description and high resolution tests. Mon. Wea. Rev., to appear.



Difference between initialized and uninitialized analyses at 12 UTC on 30 April 01.

(a) Temperature (K) 850 hPa, full IDFI.

(b) Same as a), but IDFI in vertical mode space, first 10 modes filtered

(c) Wind speed (m/s) at 1000 hPa, full IDFI.

(d) Same as c), but IDFI in vertical mode space, first 10 modes filtered

Evaluating data interpolation in moving sparse noisy data to a uniform grid

Soumo Mukherjee*†, Daniel Caya†, René Laprise†

1. The Problem

The Canadian Regional Climate Model (CRCM) (Caya and Laprise, 1999) is routinely used to provide regional climate change projections. In order to assess the quality of CRCM, simulations run in the past must be compared with observational data.

However, observations contain error, and the observational network is distributed inhomogeneously. In Canada the observational network is densest towards the South (near population centres), whereas model output is homogeneously distributed on a 45km X 45km grid, leaving open the question, how can a fair comparison be made?

We propose to use a multi-variate, noisy-data interpolator to grid the observational network. However before doing so, the performance of the interpolator itself must be appropriately understood. Thus in the present experiment, we take CRCM screen temperature over the Quebec Region and choose noisy data subsets (simulating observational networks), trying to reproduce the original field using our interpolator.

2. Methodology

a. The model

ANUSPLIN, developed at the Australian National University makes use of thin-plate smoothing splines to minimize noise, thus creating smooth fields (for a more complete description, see Hutchinson, 1997). Clearly, maintaining smooth fields comes at the cost of preserving data-fidelity. Through minimization of the appropriate penalty function, ANUSPLIN finds the optimal balance between exact data interpolation (keeping loyal to the data, leaving rough fields) and regression (producing a smooth, less loyal field), objectively.

b. The data

Presented here (Figure 1) is screen temperature (ST) produced over the Quebec region from top left (70°W, 66°N) to bottom right (70°W, 40°N), on July 1st, 1978 using the CRCM. This dataset was thinned by leaving only every Nth row and column (termed NxN). Another subset (STN) corresponding to actual observational stations present in Canada at the time, was used. The full (1x1) dataset was also sullied with random (uncorrelated) noise at the 2, 5, 10, and 20% (of the maximum field range ~25°C) levels (corresponding to ~±0.5, 1, 2.5, 5°C). At these noise levels, the thinning procedure was then used to obtain noisy data subsets.

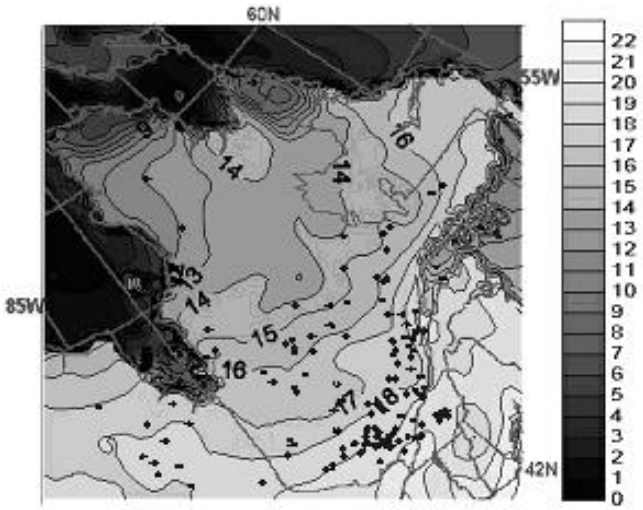


Figure 1. Summer Screen Temperature over Quebec with superimposed STN mask (black dots)

The experiment was then to interpolate from the NxN sparse fields to the original (dense) grid in order to determine the effect of selection (sparseness) as the fields deteriorate. Similarly the performance of ANUSPLIN was ascertained in the face of noise. Finally the noisy sparse sets were used to simulate an observational network where both effects were present.

c. Diagnostics

Principal diagnostics include difference maps depicting the difference between the fitted field and the original set, with the lighter colours depicting over-estimation and the darker, under-estimation.

The variance estimate, σ^2 (not shown), is the sum-of-squares of the fitted residuals, and represents the common data error. The model standard error, σ_m^2 (also not shown), represents the distributed Bayesian error-of-fit estimate. The prediction standard error, σ_p^2, represents the total error, as contributed to by both of these errors:

$$\sigma_p=(\sigma_m 2 +\sigma 2)1/2 \qquad (1).$$

Hence this is the distributed error we could expect to calculate from our model given a certain data set.

3. Results

Once the original field is removed from the interpolated field with relatively high level of noise (10% or ±2.5°C) and sampled with few data points (4% of the original set), we can see that the interpolator performs well (Figure 2) with 60% of the difference over land within ±1°C (if we include water, this drops to 50% over the whole domain). Approaching the coastline, values are underestimated, as over-smoothing

***Corresponding author address:* Soumo Mukherjee, Sc. Terre-Atmosph., UQAM Salle PK-6435, B.P. 8888, Succ. "A", Montreal QC, Canada H3C 3P8.**
e-mail: soumo@sca.uqam.ca
†*Université du Québec à Montréal*

occurs in order to lessen the gradient. Reaching the water the opposite is true for the same reason. The effect of the land-sea interface is reduced in the St. Lawrence as it is interior to the domain. In the Hudson, Ungava, and James Bays, as well as Hudson Strait, and the Labrador Sea, there is less data on at least one side (domain border), so over-estimation occurs in accord with the rest of the domain (higher temperature).

The prediction standard error (equation 1) rises to ±3°C at noise level 10% and 5x5 data subset, but

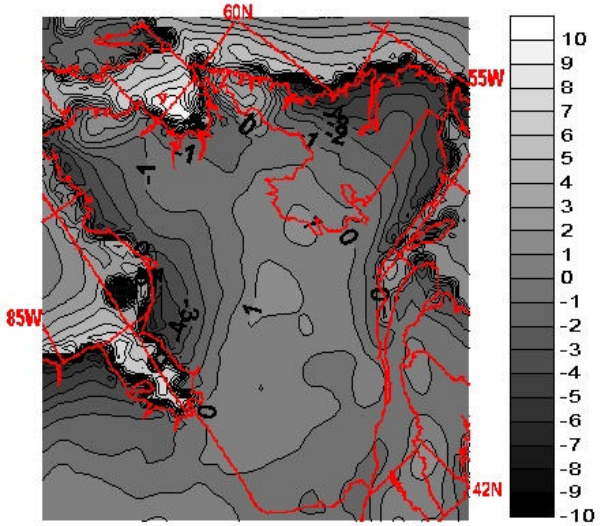


Figure 2. Difference of interpolated field (10% noise, 5x5 dataset) from original field (fig.1)

remains as low as just over ±1°C for no noise, 2x2 subsets, or using all data and noise levels less than 5% (not shown). Shown (Figure 3), are the σ_p for all cases with NxN subsets along the horizontal, and noise increasing vertically).

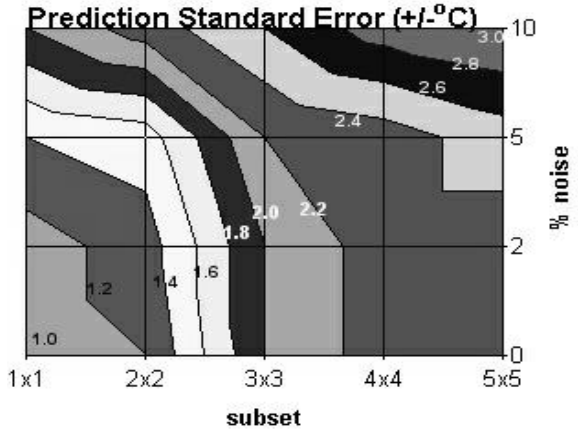


Figure 3. Prediction standard error composites

Interpolation from the realistic, scattered stations (see Figure 1 for distribution) is poor over water but good on land. The difference (interpolate less original field) map (Figure 4) shows a range that is greater by 2°C than that of the difference field of the 5x5 subset with 10% noise (Figure 2). The range is also shifted upwards (indicating overestimation) because the stations are mostly located on land, so values over water are under-represented. This is seen in over-estimated values upwards of 15°C in Hudson Bay. However, 57% of the total area is within ±1°C, due to better interpolation over land. Indeed, this is comparable to fields somewhere in the range of 2-5% noise for the 5x5 subset, or 5-10% noise for the denser 3x3 and 4x4 subsets.

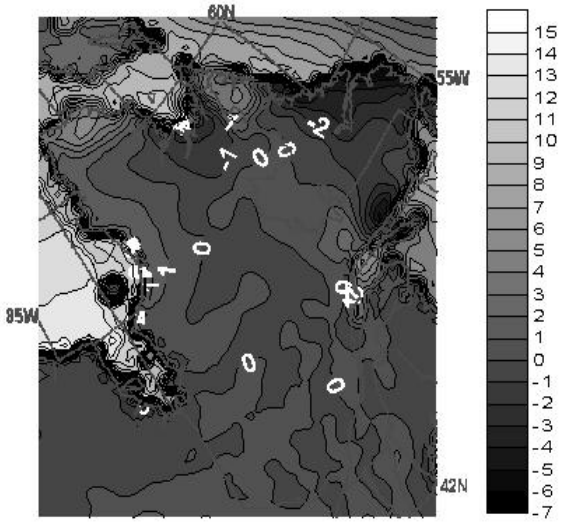


Figure 4. Difference of interpolated field (STN) from original

4. Conclusions

Problems occur mainly in regions of sharp change between land and water near the boundaries where the gradient can be as high as 8°C/100km. However, in the St. Lawrence for example, these steep (10°C/100km) gradients do not pose a problem as they are sufficiently interior (hence more anisotropically surrounded by data). Irregularly distributed data leads to exceedingly poor sampling in data-sparse regions, with better results over land.

Further work is being done to ascertain the effect of inclusion of the boundary, extra stations, as well as field continuity (precipitation, seasons, climatologies), and spatial distribution (non-homogeneous sets).

References

Caya, D. and Laprise, R. 1999: A semi-implicit semi-Lagrangian regional climate model: The Canadian RCM. *Mon. Wea. Rev.*, **127**, pp. 341-362.

Craven P. and Wahba G. 1979: Smoothing noisy data with spline functions. *Numerische Mathematik* **31**, pp. 377-403.

Hutchinson, M. F. 1997: ANUSPLIN VERSION 3.2, http://cres.anu.edu.au/outputs/anusplin.html.

CONSTRUCTION OF MULTIDIMENSIONAL BASIS VECTORS FOR ANALYSIS AND PROGNOSIS

V. Penenko and E. Tsvetova
Institute of Computational Mathematics and Mathematical Geophysics SD RAS, Novosibirsk
e-mail Penenko@sscc.ru, tsvet@OMMFAO.sscc.ru

A new approach to construction of atmospheric circulation scenarios for making ecological predictions and planning is proposed. It is based on a combined use of hydrodynamic models and archived data on climate. The problem of long-term prediction of hydrodynamic background for ecological studies and planning has no unique solution until now. Its specificity is in the fact that the characteristic lifetime of the already existing and designed objects as the sources of anthropogenic impact on the climate system are, as a rule, much longer than the characteristic intervals of predictability of the current hydrodynamic models. This means that it is preferable to make use of the scenario approach in solving such problems. In construction of scenarios, both mathematical models and measurement data are generally used. It is clear that scenarios for ecological prediction should reflect actual situations on the climatic scale.

In the proposed technique, the set of scenarios is constructed. The technique is based on the combination of the methods of four system levels, namely, the method of factor analysis, methods of studying the sensitivity of models and functionals, methods of direct and inverse modeling with assimilation of actual data. The basic vectors and the corresponding phase subspaces are the key elements of the constructions.

Identification of the main factors that govern the behavior of the climate system fills a highly important place in the methodology of formation of scenarios and analysis of the results obtained by modeling. In terms of the main factors, it is possible to identify the manifestation of the climate system response to anthropogenic impact.

Following the ideas of factor analysis [1,2], the calculation of the orthogonal basis functions is made by the formulas

$$\vec{F}_p = \sum_{\beta=1}^{n} (a_{\beta p} / \lambda_p)\vec{\varphi}_\beta, \quad p = \overline{1, m \le n}. \tag{1}$$

Here $\{\vec{\varphi}_\beta, \beta = \overline{1,n}\}$ is the initial set of normalized vectors of the state functions. The components of the latter are centered about their mean values; λ_p are the eigenvalues ordered in diminution, and $\vec{a}_p = \{a_{\beta p};\ p, \beta = \overline{1,n}\}$ are the corresponding eigenvectors of $n \times n$ Gram's matrix that is built on the initial set $\{\vec{\varphi}_\beta\}$. The vectors $\{\vec{F}_p\}$ and $\{\vec{\varphi}_\beta\}$ are of the same block structure of the type:

$$\{\vec{\varphi}_\beta\} \equiv \{\vec{\varphi}_{i\beta}((\vec{x},t)_k),\ k = \overline{1,K},\ i = \overline{1,N},\ (N \ge 1),\ \vec{x} = (x_1, x_2, x_3)\},\ (\vec{x},t)_k \in D_t^h. \tag{2}$$

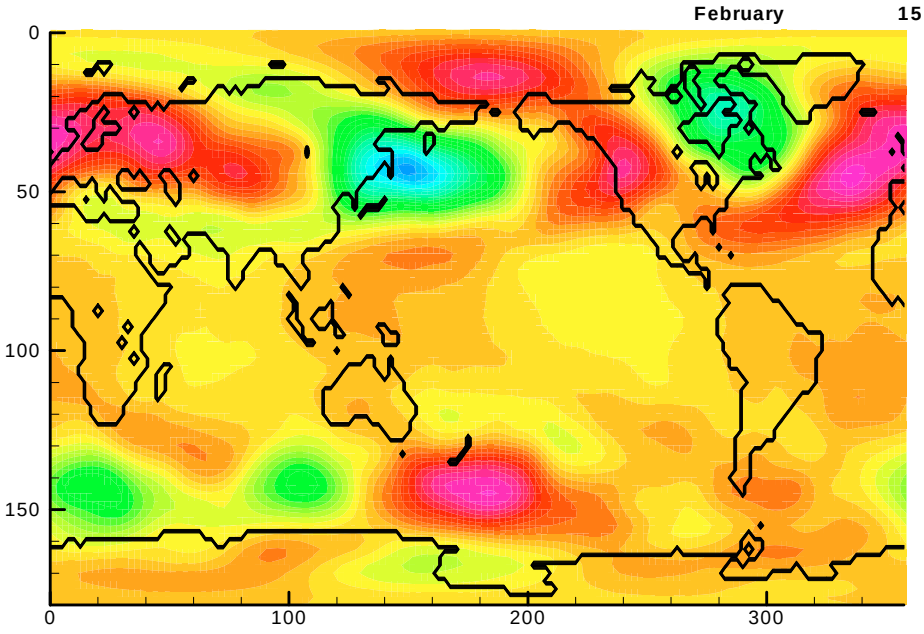
Here N is the number of different, in physical sense, components; D_t^h is the discrete analog of the space-time domain D_t. The structure of the grid in D_t^h and K, N are prescribed as input parameters. The values of K and N may be as great as it is required. The state functions of the climatic system, such as geopotential, temperature, the components of the velocity vector and others, the adjoint functions that correspond to them, sensitivity functions of the model may be included into the number of the vector's components. The union of the divers information and the choice of the inner product for the construction of the Gram's matrix are made with the help of variational principle and energetic functional of integral identity for the basic model of the processes.

The eigenvalue problem for Gram's matrix is solved under conditions

$$\lambda_p = \sum_{\beta=1}^{n} a_{\beta p}^2 = \max_{\langle a_{\beta p} \rangle}, \quad p = \overline{1,n}. \tag{3}$$

The maximums are sought on the set of values of the Fourier coefficients of the initial vectors $\{\vec{\varphi}_\beta\}$ decomposition with respect to the vectors of the sought basis $\{\vec{F}_p\}$. It is done by successive exhaustion process of subspaces beginning with the dimension n to 1.

The basis (1) and phase subspaces are derived to study the variability of the climatic system dynamics [3]. That is why the parameter n is taken equal to the number of years in Reanalysis data base [4]. The time interval in D_t^h is accepted as long as 1 month with daily behavior in 12 hours. The resolution in horizontal directions (x_1, x_2) of the global system, is taken as $2{,}5^\circ \times 2{,}5^\circ$ in spherical coordinates. The number of vertical levels is prescribed as a parameter, from 1 to 20. The parameter N is taken in dependence on the goal of the study as well. The month interval is a compromise between the informative quality and the amount of calculations. Thus, the set of 12 monthly basis vectors is produced from 40-year Reanalysis data (1960-1999). It is used to discover the main factors and analyze the multi-year atmospheric dynamics and quality of the atmosphere.



In Fig. the fragment of the first month main factor (FMMF) of the geopotential at 500 mb corresponding February 15, 00.00 is demonstrated. The space-time behavior of the FMMF shows not many almost stationary extrema areas that can be interpreted as energy active zones.

Taking into account the constructions mentioned above, a new type of geophysical hydrodynamics models has been introduced. It can be called the model with "guide". In such models, long time-space observations and the bases of the type (1) are used to form the special phase spaces, the so-called "guides". The models of hydrodynamics and pollutant transport are used here as interpolants which assimilate the elements of the guiding phase space under the given optimal criteria. The fields of application of such models are: reconstruction of the state functions from the current monitoring data in real time, diagnosis of the processes, formation of scenarios for ecological forecast and design, etc. The principle advantage of the approach is the fact that predictability is no longer crucial for it because the method is based on the use of the model and observed data.

This work is supported by the European Commission (ICA2-CT-2000-10024), the RFBR (00-15-98543, 01-05-65313) and SD RAS (00-56, 00-64, 00-73).

References

1. Harman, H.H. Modern Factor Analysis. 1976. 3rd Ed. Univ. of Chicago Press, Chicago.
2. Preisendorfer, R.W. Principle Component Analysis in Meteorology and Oceanography. Elsevier, 1988.
3. Penenko V.V., Tsvetova E.A. Formation of atmospheric circulation scenarios for climate and ecological studies. Atmos. Oceanic Opt. Vol.14. N 6-7, 444-448.
4. E. Kalnay, M. Kanamitsu, R. Kistler, et al., The NCEP/NCAR 40-year reanalysis project. Bull. Amer. Meteor. Soc. **77**, 437–471 (1996).

VARIATIONAL FAST DATA ASSIMILATION ALGORITHMS

Vladimir Penenko and Elena Tsvetova
Institute of Computational Mathematics and Mathematical Geophysics SD RAS
Novosibirsk, Russia
e-mail penenko@sscc.ru, tsvet@OMMFAO.sscc.ru

The fast data assimilation (FDA) algorithms are proposed to study the atmospheric, oceanic and environmental problems. They are based on a variational principle. The numerical schemes are obtained from the **local** optimum conditions for objective functionals in the framework of splitting technique that is used for the construction of the discrete form of the model of processes.

Let's take a mathematical model and a set of measured data

$$B\frac{\partial \vec{\varphi}}{\partial t} + G(\vec{\varphi}, \vec{Y}) - \vec{f} - \bar{r} = 0, \tag{1}$$

$$\vec{\Psi}_m = [H(\vec{\varphi})]_m + \vec{\eta}, \tag{2}$$

where $\vec{\varphi} \in \Im(D_t)$ is the state function, $\vec{Y} \in R(D_t)$ is the vector of model parameters, $G(\vec{\varphi}, \vec{Y})$ is a space operator of the model, $\vec{f}$ is a source term, $\vec{r}, \vec{\eta}$ are the terms describing uncertainties and errors of the model and data, $\vec{\Psi}_m$ is a set of measured data, $H(\vec{\varphi})$ is a model describing association between the state function and measured quantities, operation $[\,]_m$ denotes transfer to the measurement locations, $\Im(D_t)$ is the space of the state functions, $R(D_t)$ is the space of admissible parameter values. The model (1) is supposed to be defined on the space-time domain $D_t = D \times [0, \bar{t}]$. D is the domain with space coordinates $\vec{x}$; $[0, \bar{t}]$ is the time interval. The set of measurements is given on the subdomain $D_t^m \in D_t$.

The problem considered can be solved with the help of variational approach and splitting technique. The extended quality functional is introduced in the form

$$\tilde{\Phi}^h(\vec{\varphi}) = 0.5\left\{(1-\alpha)(\vec{\eta}^T M_1 \vec{\eta})_{D_t^m} + \alpha(\vec{r}^T M_2 \vec{r})_{D_t^h}\right\}^h + \left[I^h(\vec{\varphi}, \vec{Y}, \vec{\varphi}^*)\right]_{D_t^h} \tag{4}$$

where $M_i, (i = \overline{1,2})$ are weight matrices, $0 < \alpha \le 1$ is a weight coefficient, indices T, h denote transposition and discrete analog,

$$I(\vec{\varphi}, \vec{Y}, \vec{\varphi}^*) = \int_{D_t}\left((B\frac{\partial \vec{\varphi}}{\partial t} + G(\vec{\varphi}, \vec{Y}) - \vec{f} - \bar{r}), \vec{\varphi}*\right)dDdt = 0 \tag{5}$$

is the variational form of the model (1), $\vec{\varphi}* \in \Im^*(D_t)$ is the co-state function, that is generalized Lagrange multiplier. In this case, the model (1) plays the role of restriction to the state function and expresses the connection between model parameters and the state functions. The inner product in (5) is defined from the form of the functional of the energy balance of the system. The first bracket in (4) is the discrepancy functionals for the model of measurements and model of processes.

Three basic methods of data assimilation can be derived from the variational principle for (1)-(4). They are the assimilation with adjoint functions, the procedure of the Kalman-Bucy filtering and FDA. They are equivalent in accuracy but different in realization and hence in their efficiencies.

It is reasonable to design FDA using the properties of splitting technique and the conditions of local minimum of goal functionals (4). In fact, to estimate the state function, one can originate from the conditions of successive minimization of the functional that expresses the amount of uncertainties of the model and data. In other words, if uncertainties of the model within the time interval $[t_{j-1}, t_j]$ are con-

centrated at the last stage of splitting scheme, giving the solution for the moment t_j, data assimilation can be successively combined with realization of this particular stage, leaving the rest stages without changes.

The computational scheme of such approach can be obtained with the help of stationary conditions for the goal functional (4) with respect to the grid components of the state function just on the integer time steps t_j. In short, this is the idea of fast data assimilation. It should be mentioned that the whole family of fast data assimilation procedures can be derived in the frames of the approach. The form of the particular algorithm depends on the version of splitting scheme and approximations of the functionals (4)-(5). The case when the grid D_t^h is built in such a way that elements of the set D_t^m, which is the measurements' support, coincide with the grid points of D_t^h, admits the simplest realization. Thus, numerical model with successive data assimilation is realized by the splitting scheme, the last stage of which being modified.

The discrete analog of (4) can be written in the form

$$\Phi^h(\vec{\varphi}) = 0.5\left\{\sum_{j=1}^{J} \left((1-\alpha)(\vec{\eta}^T M_1 \vec{\eta})^j_{D_m} + \alpha(\vec{r}^T M_2 \vec{r})^j_{D_t^h}\right)\Delta t\right\} +$$

$$\sum_{j=1}^{J}\left\{\sum_{k=1}^{n}\left\{\left(\left(\vec{\varphi}^{\,i+\frac{k-n}{n}} - \vec{\varphi}^{\,i+\frac{k-n-1}{n}}\right) + \Delta t\left(A_k\left(\vec{\varphi}^{\,i+\frac{k-n}{n}}\right) - (\vec{f}+\vec{r})^j \delta_{j,j+\frac{k-n}{n}}\right)\right)\vec{\varphi}*^{\,i+\frac{k-n}{n}}\right\}\right\} = 0 \qquad (6)$$

where k is the number of the stage, n is the amount of fractional stages, A_k is implicit linearized approximation of split part of the model operator $G_k(\vec{\varphi},\vec{Y})$, $G\ (\vec{\varphi},\vec{Y}) \equiv \sum_{k=1}^{n} G_k(\vec{\varphi},\vec{Y})$, $\delta_{\alpha,\beta}$ is Kronecker-delta. The inner products in (6) are taken over space domains. The second line in (6) is the approximation of integral identity (5) for the model (1) in terms of splitting technique.

The fast assimilation algorithm can be derived now from the minimum conditions for the modified functional (6) with respect to the components of the state functions $\vec{\varphi}^{\,j}$ in the node points of the grid domain D_t^h. The modification is to exclude the functions $\vec{\eta}, \vec{r}, \vec{\varphi}^*$ from (6) with the help of the discrete analog of (2) and the splitting scheme for the model (1). The splitting scheme is obtained by means of the stationary conditions for the functional (6) with respect to the variations of grid components of $\vec{\varphi}^*$.

Omitting intermediate transformations, let us write the system of equations of the last split stage for calculation of the state functions $\vec{\varphi}^{\,j}$ with the use of observed data $\vec{\Psi}_m^j$

$$\alpha\left(E + \Delta t\, A_n\right)^T M_2\left(\left(E + \Delta t\, A_n\right)\vec{\varphi}^{\,j} - \Delta t\ \vec{f}^{\,j}\right) + (1-\alpha)\,\Delta t \tilde{H}^T M_1\left(\left(\tilde{H}\vec{\varphi}\right)^j - \vec{\Psi}_m^j\right) = 0\,. \qquad (7)$$

Here $\tilde{H}$ is the linearized operator of measurement model (2).

As the starting point of our considerations is (4) , we have a possibility to design (6) in such a way that all operators in the splitting schemes, including (7), are realized with the help of simple and efficient direct (non-iterative) numerical algorithms. Parameter α is used to control the assimilation procedure. If $\alpha = 1$, the model ignores measured data, and if $\alpha \to \varepsilon$, the data are predominate in the calculation of the state functions. Here ε is of a small positive value. The contribution of each element is defined in dependence on the degree of reliance to this component. The observed data are involved in modeling process as soon as new information becomes available. The accuracy of the algorithm is defined by that of the functionals' approximation in (4) - (6).

This work is supported by the European Commission (ICA2-CT-2000-10024), the RFBR (00-15-98543, 01-05-65313) and SD RAS (00-56, 00-64, 00-73).

Simplified Initialization

Yury A. Pichugin
Voeikov Main Geophysical Observatory, St. Petersburg, Russia
pichugin@JP4974.spb.edu

The procedure of initialization usually used in hydrodynamic forecasts apparently is excessively complicated. Obviously, this procedure can be simplified sparing a time of calculations and ensuring a necessary result.

The initializated field $\boldsymbol{X}^*$ can be obtained using the formula $\boldsymbol{X}^* = \sum_{i=-N}^{N} h_i \boldsymbol{X}_i$, where the values h_i ($i= -N,...,N$) are coefficients of a digital filter, $\boldsymbol{X}_i = \boldsymbol{X}(t_0+i\Delta t)$ ($i= -N,\ldots,N$), t_0 is an initial instant, Δt is a time step. The time series $\boldsymbol{X}_i$ ($i= -N,...,-1$) is obtained by adiabatically integration backward from $\boldsymbol{X}_0$ or by diabatically integration forward from $\boldsymbol{X}_{-N}$. (Huang and Lynch, 1993).

Using a symmetric digital filter, for example $h_i = h_{-i} = (0.5+0.5\cos(\pi i/(N+1))(N+1)^{-1}$ ($i=1,\ldots,N$), $h_0=(N+1)^{-1}$ (Jenkins and Watts, 1968), we can obtain the initializated field by the formula $\boldsymbol{X}^* = h_0\boldsymbol{X}_0 + \sum_{i=1}^{N} 2h_i\boldsymbol{X}_i$, where the time series $\boldsymbol{X}_i$ ($i=1,\ldots,N$) is the result of diabatically integration forward from $\boldsymbol{X}_0$. Obviously, the result of this filtration is determined only by magnitudes of the time interval $N\Delta t$ and the time step of integration Δt.

Note also that there are no essential arguments to use recursive filters (Lynch, 1993) invented for the purposes of radio engineering, where the filtration of a signal is carried out simultaneously with reception.

References

Huang X.-Y. and P. Lynch (1993), Diabatic Digital-Filtering Initialization: Application to the HIRLAM Model. *Mon. Wea. Rev.*, Vol. 121**,** pp. 589-603.

Jenkins G.M. and D.G. Watts (1968), Spectral analysis, San Francisco, Holden-Day.

Lynch P. (1993), Digital Filters for Numerical Weather Prediction (1993), HIRLAM, Technical Report No. 10.

ASSIMILATION EXPERIMENTS WITH GPS/MET REFRACTIVITY

Paul Poli* and Joanna Joiner
NASA Goddard Data Assimilation Office, * also at Météo France
ppoli@dao.gsfc.nasa.gov, jjoiner@dao.gsfc.nasa.gov

The Global Positioning System (GPS) represents a free source of stable electromagnetic *L*-band signals available for radio occultation purposes around our planet [*Kursinski et al.,* 1997]. The proof-of-concept GPS/MET experiment has collected observations of atmospheric refractivity in 1995-1997. These observations have been combined with a 6-hour forecast background in a one dimensional variational analysis (1DVAR) to yield profiles of atmospheric temperature and humidity [*Poli et al.,* 2001]. Comparisons with close radiosondes have shown that the 1DVAR retrieved profiles of temperature present smaller bias and standard deviation than the background information.

In order to perform an impact study of the GPS/MET refractivity data on a global circulation model, we use the Data Assimilation Office (DAO) state-of-the-art Finite Volume Data Assimilation System (FVDAS). The FVDAS is used to assimilate GPS/MET observations *via* the 1DVAR analysis.

The setup of the assimilation experiment is as follows. We use GPS/MET refractivity data from June 21st to July 4th, 1995 to perform GPS 1DVAR analyses using the latest available 6-hour forecast background. The GPS temperature retrievals are then assimilated like the other observation types (radiosondes, TOVS interactive retrievals) in the FVDAS using the Physical-space Statistical Analysis System (PSAS). The aim is to improve the analyzed state of the atmosphere before issuing the next 6-hour forecast. The run of the FVDAS for the two-week time period with this setup constitutes the assimilation experiment 'GPS'. As a comparison, we also ran one experiment without assimilating the GPS data ('CONTROL').

We then compare the results of the two runs 'CONTROL' and 'GPS' for the two-week time period. One way to assess the impact of the GPS observations is to look at differences between 6-hour forecasts and conventional observations. Smaller differences indicate that the GPS/MET observations help the forecast fit with independent observations. Figure 1 shows the standard deviation of such differences for the meridional wind measured by radiosondes in the Southern hemisphere. The 'GPS' run presents smaller differences than the 'CONTROL' run in the upper troposphere-lower stratosphere.

Another way to evaluate the impact of the GPS observations is to issue five-day forecasts from two different initial states. Looking at several days we find mixed results. Clearly, we need more GPS observations to fully assess the impact of the GPS refractivity on 5-day forecasts.

The assimilation experiments presented here used a limited number of observations (total of 797 profiles for a two-week time period). Current GPS missions (Champ, SAC-C) collect larger amounts of observations and we look forward to performing similar studies with these datasets.

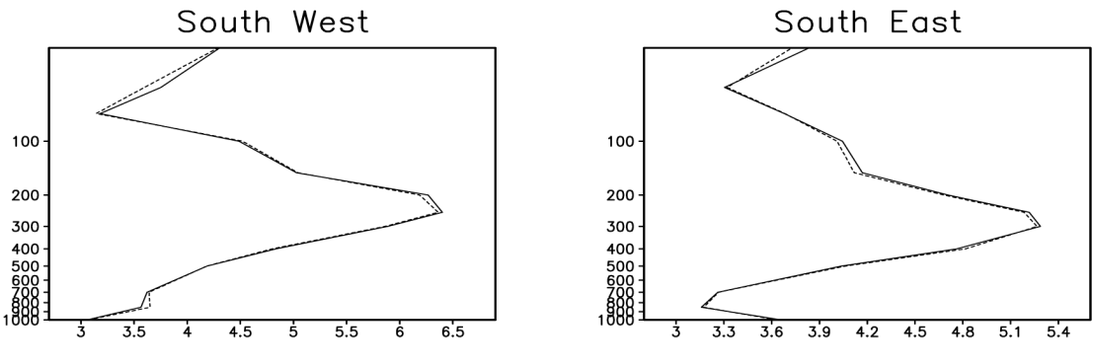


Figure 1: Standard deviation of differences between 6-hour forecast meridional wind and radiosonde observations, Southern hemisphere. Solid curve is 'CONTROL', line with dashes is 'GPS'.

Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy, Observing Earth's atmosphere with radio occultation measurements using the global positioning system, *J. Geophys. Res. 102,* 23,429-23,465, 1997.

Poli, P., J. Joiner, and E. R. Kursinski, 1DVAR analysis of temperature and humidity using GPS radio occultation refractivity data, *Submitted to J. Geophys. Res.*, 2001.

Development of an MM5-based Four-Dimensional Variational Analysis System for Distributed Memory Multiprocessor Computers

Frank H. Ruggiero
Air Force Research Laboratory, AFRL.VSBL
29 Randolph Road, Hanscom AFB, MA 01731-3010
Email: frank.ruggiero@hanscom.af.mil

George Modica, Thomas Nehrkorn, Mark Cerniglia
Atmospheric and Environmental Research, Inc.
131 Hartwell Avenue, Lexington, MA 02421-3126

John G. Michalakes
National Center for Atmospheric Research, NCAR/MMM
3450 Mitchell Lane, Boulder CO 80301

Xiaolei Zou
Florida State University
404 Love Building, Tallahassee, FL 32306

The Air Force Research Laboratory is leading a multi-year effort to develop an updated version of the National Center for Atmospheric Research/Penn State University Fifth Generation Mesoscale Model (MM5) four-dimensional variational analysis (4DVAR) system. The main objective of this update is to optimize the code to run on distributed memory computers with the goal of achieving substantial speedup in wall-clock time processing and allowing mesoscale 4DVAR to enter the realm of operational forecasting. The previous version of the MM5 4DVAR system (Zou et al. 1997 [4]) is coded for single processor computer architectures and its non-linear, tangent-linear, and adjoint components are based on version 1 of MM5. To achieve the project goal, two major steps need to be accomplished. First, a version of an MM5 4DVAR based on version 3 of the model has to be created. The reason for this is to be able to take advantage of the existing parallelization mechanisms (Michalakes 2000 [2]) already in place for the version 3 non-linear model. The second step is to then install the parallel code architecture into the 4DVAR system's tangent-linear and adjoint components.

The first step has been completed. A complete 4DVAR system based on MM5v3.4 has been developed and tested. The tangent-linear and adjoint components were developed using the Tangent Linear and Adjoint Model Compiler (TAMC) automatic adjoint code generator. TAMC (Giering and Kaminski 1998 [1]) is a source-to-source translator that generates FORTRAN code for the tangent-linear model or adjoint from the FORTRAN code of the non-linear model. It is possible to incorporate the TAMC as part of the model compilation process, requiring the maintenance of just the non-linear model code. However in this project, TAMC is used as a development tool only. This results in separately maintained tangent-linear model and adjoint versions of the code. This approach makes it possible to minimize changes to the MM5 non-linear

model code as supported by the National Center for Atmospheric Research but does require a mixture of manual and automatic code generation. Modules of the tangent-linear model and adjoint were created and tested individually. Unit testing was done by finite differencing comparisons with the non-linear model for the tangent-linear model and direct comparisons with the tangent-linear model for the adjoint. When all modules were assembled, a final unit integration test was carried out for both the tangent-linear model and adjoint. Additional testing was carried out by comparing the results from the v3 MM5 4DVAR with those from the v1 MM5 4DVAR.

The new MM5 4DVAR system contains most of the limited set of physics packages that are in MM5v1 based 4DVAR system. In particular, there are tangent-linear model and adjoint routines for bulk boundary layer processes, convection, and explicit, grid-scale precipitation. Additional physics options planned will be added as well as the tropical cyclone bogussing scheme of Zou and Xiao (2000 [3]).

Work on parallelizing the 4DVAR code for distributed memory computers is on going. The non-linear model, which is MM5v3.4, is already structured to run on distributed memory computers. Parallel code architecture for the tangent-linear model has been installed and tested. Using a problem with 84 x 75 x 27 grid points, running on an IBM SP P3, and using 24 processors, the tangent-linear model ran over 20 times faster in terms of wall-clock time than when running on one processor. Work to improve the wall-clock speed up is continuing, with the focus being on trying to eliminate the I/O bottleneck that occurs from reading in the base state information at every time step. A parallelized version of the adjoint is currently under development and is expected to be complete by mid-2002. Release of a beta version of the fully parallelized MM5v3 based 4DVAR is scheduled for late 2002.

References

[1] R. Giering and T. Kaminski. Recipes for adjoint code construction. *ACM Trans. Math. Software,* 24(4):437-474, 1998.

[2] J. Michalakes. The same source parallel MM5. *Sci. Progr.,* 8:5-12, 2000.

[3] X. Zou and Q Xiao. Studies on the intialization and simulation of a mature hurricane using a variational bogus data assimilation scheme. *J. Atmos. Sci.* 57:836-860, 2000.

[4] X. Zou, W. Huang, and Q. Xiao. A user's guide to the MM5 adjoint modeling system. Technical Note 437+IA, NCAR, Boulder, CO, 1997.

Benford's Law and background field errors in data assimilation

R S Seaman
Bureau of Meteorology Research Centre
Melbourne, Australia

Benford's Law is a logarithmic distribution of first significant digits that is observed in many naturally occurring non-meteorological data sets, such as catchment areas of rivers, stock market prices and tax returns. In such data sets, the frequency of occurrence (f) of first digit (d) is given by

$$f = \log (1 + 1/d), d = 1,2 .. 9$$

where the logarithm is base 10. That is, the first significant digit 1 occurs with a frequency of about 30%, while 9 occurs with a frequency of less than 5%. Benford's Law has been used as a tool for the detection of accounting and taxation fraud.

The present application is motivated by a recent proof by Hill (1996) that Benford's Law can be considered as the asymptotic result of combining random data from random distributions. The goodness of fit to Benford's Law may therefore indicate the extent of mixing in a data set.

Background field errors in most data assimilation systems are assumed to be normally distributed. Ensembles of background field errors from the Australian Bureau of Meteorology's data assimilation system have been used to calculate (i) the goodness of fit of the errors to a Gaussian distribution, and (ii) the goodness of fit of the first significant digits of the errors to Benford's Law. A quasi-inverse relation was found between these two goodness of fit measures. This tends to support the suggestion (Ingleby and Lorenc 1992, Devenyi and Schlatter 1994) that background field errors in general may be considered as mixtures of Gaussian distributions. Confirmation is provided by statistical simulation, and by studies of individual ensembles.

References

Devenyi D and Schlatter T W. 1994. Statistical properties of three hour prediction "errors" derived from the mesoscale prediction system. Mon. Weath. Rev., 122, 1263-80.

Hill T P. 1996. A statistical derivation of the significant digit law. Statistical Science, 10, 354-63.

Ingleby N B and Lorenc A C. 1992. Forecast errors and buddy checks. Preprints. 12th conference in probability and statistics in the atmospheric sciences. Toronto, Amer. Met. Soc., 207-13.

Seaman R S. 2002. The relevance of Benford's Law to background field errors in data assimilation. Aust. Met. Mag., 51, 25-33.

BMRC Data Assimilation Progress Report

P. Steinle, B. Harris, R. Seaman, J. Kepert, , J. Paevere T. Glowacki,
W. Bourke, C. Tingwell, J. Le Marshall, T. Le

Bureau of Meteorology Research Centre
Melbourne, Australia

Development and Implementation of BMRC 3D-VAR Assimilation System

A new three dimensional variational (3D-VAR) analysis scheme using an observation space formulation is currently being experimentally evaluated in BMRC. This system can use either standard functional representations of background error correlations (as were commonly used within OI schemes) or empirical correlations as obtained from ensembles of analysis-forecast cycles (as implemented by Fisher and Buizza), or statistics of scaled differences between 48 hour and 24 hour forecasts (the so-called NMC method). This 3D-VAR scheme can be used for both global and regional systems. The analysis procedure is based on a series of local sub-volume analyses, with data selection extending significantly beyond the individual analysis volumes. Scattering from observation analysis increments back to the analysis grid-points is performed once all sub-volumes of the required domain have been analysed; this is an important part of ensuring that the analysis is independent of the geometry of the localized solution domains. The inclusion of the direct (3D) assimilation of radiances and the use of alternative coordinates for specifying background error correlations is in progress.

The currently used statistics based on the NMC method of scaled differences are being compared to the statistics based upon ensembles of differences between 6-hour forecast background fields, generated by parallel analysis-forecast cycles with perturbed observations following Fisher and Buizza (ECMWF). The main differences between the statistics generated by the two approaches appear to be (i) the ensemble approach implies shorter length scales, and (ii) the ensemble approach implies more variance in the divergent component of the wind, especially outside the tropics. Parallel assimilation and forecast trials utilising the two alternative structures are in progress.

Global 1D -VAR Assimilation of ATOVS.

The present operational global assimilation system utilises a 1D-VAR retrieval of both NOAA15 and NOAA16 AMSU-A and HIRS radiances, utilising 1D radiances as available from NESDIS. An extended version of this global system has been experimentally implemented and tested with 50 vertical levels (with the top level at 0.1 hPa). This version of the system allows the full forward calculation of ATOVS radiance first guess values in the 1D-VAR retrieval scheme still using level 1D ATOVS radiances. This obviates the need to interpolate NESDIS temperature retrievals above this level, as is the case in the current 29 level operational system, which has a top level of 10 hPa. Extensive global assimilation experiments have been conducted during year 2000 (both T79/L50 and TL239/L50) and medium-range prediction performance in the stratosphere has been substantially improved .

In the immediate future 1C radiances, will be utilised and each instrument (HIRS or AMSU-A or B) in the ATOVS instrument package will be treated as a separate observation. Additionally direct readout/processing of radiances at BoM ground-stations delivers 1C radiances which are desirable in support of early cut-off regional assimilation. Initial experiments with the use of 1C AMSU-A radiances alone have indicated a deficiency arising from the absence of the infra-red HIRS; incorporating these HIRS radiances and and appropraite cloud clearing strategy is the focus of present work.

Regional 1D -VAR Assimilation of ATOVS

The online 1D-VAR ATOVS radiance retrieval scheme, implemented operationally within the global system (GASP), has also been integrated with the Bureau's Limited area Assimilation and Prediction System (LAPS), as part of the effort to unify the data assimilation component of the local and global forecasting systems. The current operational version of LAPS has 29 vertical levels up to 50 hPa and NESDIS retrievals are used to extend the first guess profiles above this. The 1D-VAR retrievals are used over the sea and below 100 hPa.

Extensive near real-time assimilation and forecasting experiments have been conducted to test the performance of the LAPS 1D-VAR system. Data assimilation was performed by means of the cold-start strategy employed in the Bureau's operational LAPS system in which three six-hourly assimilation cycles leading to the base-time of the forecast are built on a GASP first-guess, with lateral boundary conditions also provided by the GASP system. For the second and third cycles, in which LAPS forecasts provide the first guess, radiance bias corrections calculated and updated continuously by the LAPS 1D-VAR system were used. Only the same observational data as available to the operational LAPS system, which employs a two hour cut-off at the forecast base-time, were presented to the experimental 1D-VAR analysis system; these included NESDIS 1D ATOVS radiance data.. (NESDIS temperature and moisture retrievals only are available to the operational system). S1 skill scores calculated from the 1D-VAR forecasts and averaged over a four month period show a significant improvement over those obtained from the operational system: the MSLP forecasts are improved by 1.1 S1 points at +24 hours and 0.8 S1 points at +48 hours. Forecasts of 850 hPa geopotential height show the most marked gains, with an improvement of 1.9 points at +24 hours. The LAPS 1D-VAR system is to be implemented operationally in March 2002.

Work is underway to test the 1D-VAR system in an extended version of LAPS with an increased number of vertical levels and the model top raised to 0.1 hPa, following similar extensions to GASP. This eliminates the need for NESDIS retrievals and will facilitate the use of locally received and processed ATOVS radiances whose timeliness will improve the amount of data available to the operational LAPS system. Early results indicate similar performance to the 29 level system at the surface and further gains to forecast skill above 500 hPa.

Global Assimilation of Scatterometer Data

Scatterometer data (Quickscat) is now being assimilated on an experimental basis within the global GASP system, with modest positive impact on medium range prediction , chiefly in the Southern Hemisphere. Quality control procedures have been supplemented with background checks of wind direction, to remove incorrectly dealiased data.. It is planned to experiment with a more physically based quality control system. There is some evidence of contention with other observation types, chiefly cloud drift winds, which seems to be due to a lack of boundary layer structure in the background covariance. The scatterometer data is expected to be included into the operational global system as part of the next major upgrade.

Additionally intercomparisons (several months) of Quickscat wind vectors with the current operational GASP and LAPS assimilation and prediction (where the Quickscat data has not been assimilated), have shown that the model analyses and short term predictions (18 and 24 hr) have no marked systematic biases relative to the scatterometer . This has been conducted in the context of clarifying/verifying the behaviour of the marine boundary layer in the models..

Generation and Quality Control of GMS-5 Atmospheric Motion Vectors

High spatial and temporal resolution atmospheric motion vectors from GMS-5 are generated routinely at the Bureau of Meteorology (BoM) for operational and research applications. Motion vectors are determined by tracking features in infrared, water vapour and high resolution visible imagery using all available Stretched VISSR images arriving at the BoM. This currently results in wind vectors being generated four times per day over the full disc from images separated by half an hour and every hour where imagery is available hourly. Recently, hourly winds have only been available over the Northern Hemisphere because of the restricted observation cycle of GMS-5.

The benefit of these high resolution winds, generated from both images separated by half an hour and one hour in assimilation in the LAPS NWP system for the Australian Region has recently been quantified in a regional impact study. A key element in utilising the winds has been quality control of the vectors both in the Northern and Southern Hemispheres. Initially, a locally developed methodology to estimate expected errors was based on image correlations, internal consistency of the vectors and differences from the model first guess. Currently, use of the QI approach, developed at EUMETSAT, is being tested to optimise the selection and use of vectors for analysis.

Validation of GPS based Total Precipitable Water estimates

Following activities at other major meteorological centres, a study of GPS based retrievals has been undertaken by Curtin University (Dr. N. Penna) and the Bureau of Meteorology Research Centre (BoM). It aims to investigate potential improvement to meteorological prediction by assimilating GPS estimated Total Precipitable Water (TPW). Derivation of realistic quality indicators for GPS estimated atmospheric water vapour is the first step in this direction. Seventeen GPS stations are included in the study, stretching from Cocos Island in the tropics, through the Australian Continent to Macquarie Island and Antarctica (3 stations). Using pressure and temperature estimates from collocated or nearby surface stations, and GPS data, TPW estimates were computed over the year 2000 for all these stations. For validation purposes, these were subsequently compared with Integrated Precipitable Water (TPW) computed from radiosonde reports, co-located or close to GPS locations. The comparison with 14 available radiosondes is for the period Jannuary - December 2000. The comparison statistics indicate the largest biases for Antarctica. The standard deviations of (GPS - radiosonde TPW) were largest for the sites with highest atmospheric moisture content or where the distance between the GPS station and the radiosonde / surface station is high. Comparison of GPS TPW with estimates from operational analyses (BoM and ECMWF) for the year 2000 is now being undertaken.

Data Management Developments

The Meteorological Archival and Retrieval System (MARS) is a software package developed at the European Centre for Medium Range Weather Forecast (ECMWF). This software has been made available to the Bureau of Meteorology.

BMRC examined and implemented a MARS prototype system in late 1998 to address a number of data management issues at BoM, where the current operational implementation data archive based on Neons/ORACLE database is considered inadequate to support research and climate activities. In late 2001, following successful prototyping of this system, MARS has now been implemented at BoM, on a 4-node 12-processor IBM SP2 system that couples directly to the Bureau's 8-drive StorageTek SILO tape system and SAN disk storage farm. High-speed HIPPI links connect the SP2 to the Bureau's NEC SX5 supercomputers as well as a number of key research and operational platforms.

MARS is currently used to archive selected global model and global ensemble system output, in addition to research experimental data.. It is expected that MARS will gradually replace the Bureau operational real time Neons/ORACLE database as the repository of all archived NWP model and observational data.

The operational 3D-Var assimilation system of JMA for the Global Spectrum Model and the Typhoon Model

Yoshiaki Takeuchi and Tadashi Tsuyuki
Numerical Prediction Division, Japan Meteorological Agency
email: ytakeuchi@met.kishou.go.jp

1. Introduction

The JMA 3D-Var operational assimilation system (Takeuchi, 2002) has been used for the global assimilation cycle four times a day (00,06,12,18UTC), the global early analysis for the global spectral model (GSM) forecast (00,12UTC), and the typhoon analysis for the typhoon model (TYM) forecast (06,18UTC) since 25 September 2001. The scheme has the following benefits over the former optimal interpolation (3D-OI) scheme 1) applying more general balance conditions among analysis variables (i.e. wind, temperature, surface pressure, and specific humidity) such as the geostrophic balance and surface frictions, 2) analyzing all observational data simultaneously, 3) capability of satellite data assimilation in a straight-forward manner. Typhoon bogus data are embedded in first guess fields of surface pressure, temperature, and wind.

The 3D-Var uses 6-hour forecast from GSM(T213L40) as a first guess (background). All data within 3 hours from analysis time are regarded as observational data taken at the analysis time. An incremental method is adopted in the 3D-Var to save computer resources. The method calculates an analysis increment at a lower resolution (T106L40) and then adds the increment to the high-resolution first guess. The background error was calculated based on one year statistics by using the NMC method.

2. Improvements in analysis

In numerical weather prediction, an initialization procedure is applied to eliminate noises caused by dynamically unbalanced analysis fields. The analysis field by the 3D-Var is well dynamically balanced, therefore, the analysis is not much modified by the initialization. Fig.1 shows the differences between initial fields and analysis fields of vertical p-velocity at 850hPa at 12UTC 17 September 2001. The shaded area denotes modifications by the initialization is more than 2hPa/hour. Though the 3D-OI analysis is modified everywhere, the 3D-Var analysis is not modified except for high terrain areas and around typhoons near Taiwan island and the south sea of Japan.

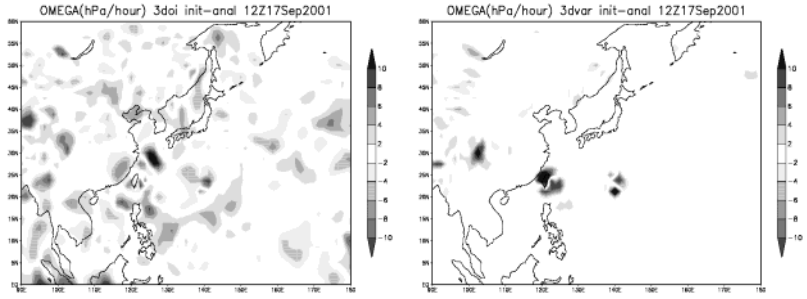


Fig.1 Differences between initial and analysis fields of vertical p-velocity (hPa/hour) at 850hPa at 12UTC 17 September 2001. Left panel for 3D-OI and right for 3D-Var.

3. Improvements in forecast

In this section, it is shown how much the 3D-Var improves short range forecasts. Examples shown are the results from forecast experiments by GSM and TYM for July 2000, March 2001 and July 2001. All experiments use the same observations and the same model. Therefore, we can find impacts of the change in the analysis scheme from 3D-OI to 3D-Var.

3.1 Performance of GSM forecast

Fig.2 shows the RMSE score of the experiment for March 2001. We carried out cycle analysis during 1 to 31 March and 216hour forecasts are performed from the initial time at 12UTC from 8 to 22 March. In Fig.2 improvements in sea level pressure, 500hPa geopotential height, 850hPa wind are remarkable even in the Northern hemisphere. The advantage of 3D-Var is more significant in the Southern hemisphere. For experiments for July 2000, the score of the Northern Hemisphere sea level pressure and 500hPa geopotential height is similar to those for the 3D-OI and the score of 850hPa wind is improved. In the Southern hemisphere, all verification scores are also improved.

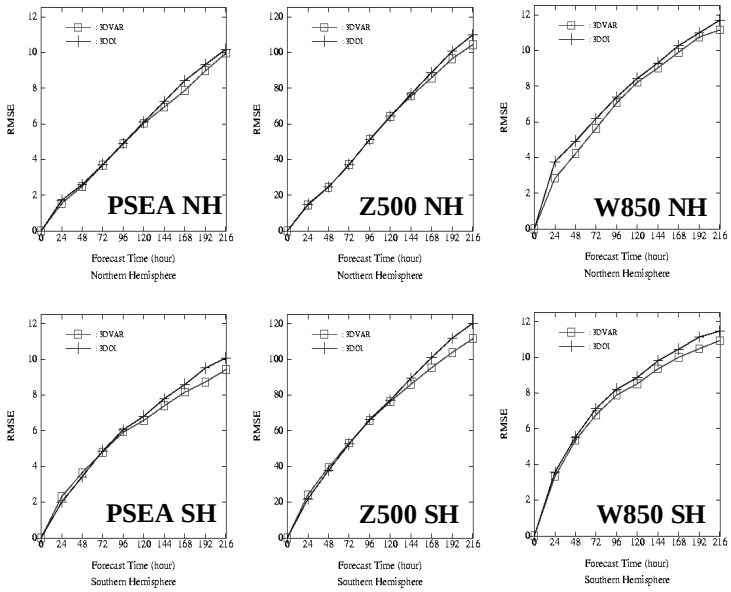


Fig.2 RMSE of forecast score in March 2001. Upper panel are for the Northern hemisphere and lower for the Southern hemisphere. Square marks denote 3D-Var and cross marks denote 3D-OI.

3.2 Impacts on typhoon forecasts

This section describes a few examples on the improvements of forecast of typhoon center position. Fig.3 is the comparison of typhoon (T0104) forecasts by GSM, the initial fields of which are given by the 3D-Var and the 3D-OI, with best tracks given as a truth. The initial time is 12UTC 9 July 2000, forecast time is 90hour. The character B shown in Fig.3 denotes best track, V denotes the 3D-Var, and O denotes the 3D-OI. The forecasted track by the 3D-OI crosses mid-Korean peninsula and that by the 3D-Var goes through northern part of Korean peninsula which is close to the best track.

Fig.4 shows the result for T0108 by TYM from the initial time at 00UTC 27 July 2001. Though the 3D-OI forecast turn right toward the Okinawa islands, the 3D-Var forecast goes straight to Taiwan island similar to the best track. Though the number of the case studies is insufficient to carry out statistical verification, the forecasts of typhoon tracks are improved in most cases.

4. Summary

The 3D-Var assimilation scheme for GSM and TYM was developed and implemented as the operational analysis system on 25 September 2001. Some experiments prior to the implementation show improvements in short- to medium-range forecasts and typhoon forecasts. The improvements are achieved without adding any new observation data and without change in forecast models.

We hope the improvement in forecast models and utilization of satellite and airplane data will contribute to more accurate initial fields for GSM and TYM. In a few years a 4D-Var scheme for those models will be implemented in operation and satellite data, the observation time of which is not regular, is treated more accurately and the forecast score will be improved significantly.

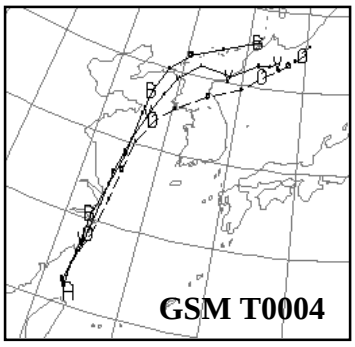


Fig.3 Typhoon forecast by GSM. The initial time is 12UTC 09 July 2000. B denotes best track, O for 3D-OI, and V for 3D-Var. Each character is positions of 24, 48, and 72 hours forecast

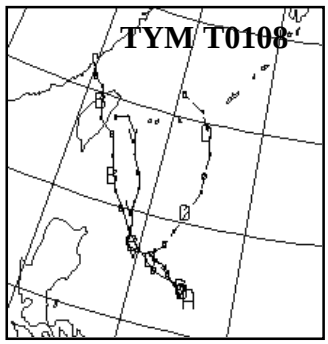


Fig.4 Same as Fig.3 except for the forecast by TYM. The initial time is 00UTC 27 July 2001.

References

Takeuchi, Y., 2002: Global analysis. *Outline of the operational numerical weather prediction at the Japan meteorological agency.* JMA, 16-20.

Imposing penalty both on gravity and Rossby modes in the Variational Initialization

Kyuichiro Tamiya
Meteorological Research Institute, 1-1 Nagamine,Tsukuba,Ibaraki,305-0052,Japan
e-mail:ktamiya@mri-jma.go.jp

1 introduction

How does the variational initialzation represent the data void regions such as the mountainous land in the meso-scale context. We have conducted an (time-1, space-2 dimensional) data assimilation experiment using a flux type shallow water equation. The case in our interest is such as the mountainous region with no reliable observation around. We may take the aproximately uniform flow near the mountain, but such flow should not be appropriate for the initial field because they show an apparent transient character due to the mountain forcing. But we show in this short note that we can control such transient character and get the spin-up-free initial fields by imposing penalty upon tendencies not only of gravity modes but also of Rossby modes.

2 numerical experiment

Cost function $I(\psi)$ consists of three parts, the first two of which is the background term $J_b(\psi)$ and the observation term $J_o(\psi)$ which constrains $\psi = (U, V, H)$ near background (U_b, V_b, H_b) and obserbations (U_o, V_o, H_o) each:

$$J_b(\psi) = (U - U_b)^2 + (V - V_b)^2 + g\bar{H}(H - H_b)^2 \tag{1}$$

$$J_o(\psi) = (U - U_o)^2 + (V - V_o)^2 + g\bar{H}(H - H_o)^2 \tag{2}$$

and the last is the penalty term $J_B(\psi)$:

$$J_B(\psi) = \left(\frac{\partial U}{\partial t}\right)^2 + \left(\frac{\partial V}{\partial t}\right)^2 + g\bar{H}\left(\frac{\partial H}{\partial t}\right)^2 \tag{3}$$

adding up them with weights $w(\vec{x})$ and an adjustable constant c,

$$I(\psi) = \int dt d\vec{x}(w(\vec{x})J_o(\psi) + c\frac{a^2}{g\bar{H}}J_B(\psi)) \tag{4}$$

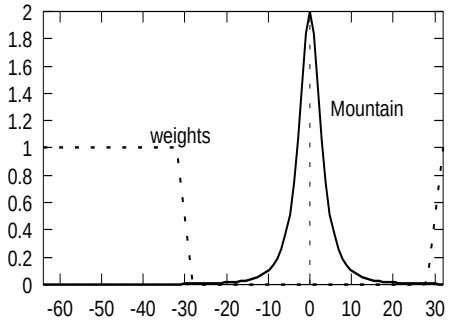
where $w(\vec{x})$ goes to zero near the mountain (centerd at $\vec{x} = \vec{x}_0$! "with raious a)(Fig. 1), representing to be no available data there.

A time integration from an uniform flow was conducted and the short assimilation period $[0, T](T \leq a/\sqrt{gH})$ was selected after fully vortex sheddings started(Fig. 2). We assumed the background field to be an uniform flow and instead of taking $J_b(\psi)$ as a part of the cost function, we searched after the minimum starting from this uniform background.

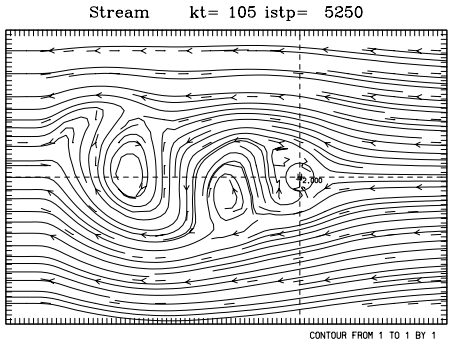
We have studied the effect of c on variational calculations. In case of $c \ll 1$, we have got an uniform flow near the mountain not going around the mountain nor over the mountain.(Fig 3). On the other hand when $c \gg 1$,we have got a pair of steady vortex quite different from observations. In the medium range of c, there are some vortices produced by the mountain shown in Fig 4.

The weak constraint by the term $J_B(\psi)$ is an extension of the balance condition, usually applied for gravity mode control but here also for Rossby modes. This term results in vortex creation behind the mountain. In figure 4 there are two vortices(a pair of vortex) just behind the mountain. This inconsistency with the observation is caused by the ambiguous uniform background field not by the inapropriateness of the applied penalty condition.

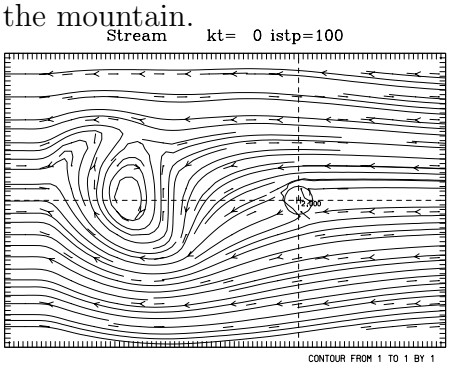
The forecast from this initial field(Fig. 4) was close to the original one and did not show a transient feature as the vortex shedding started at the begining of the forecast, while from the figure 3 initial, The vortex shedding followed a time-consuming creation of vortex pair from the uniform flow.



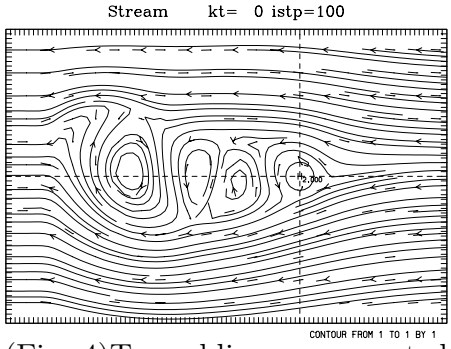
(FIg. 1)Lateral view: the mountain height is double of the water depth. There's uniform flow from the right. $w(x) = 0$ near the mountain.



(Fig. 2)Pseudo-observation produced by the time integration. We can see some eddies emitted in the wake of the mountain.



(Fig. 3)Uniform flow is seen near the nountain in case of $c \ll 1$



(Fig. 4)Two eddies are generated just behind the mountain when $c = 1$, one of them is excess compared with Fig. 2 above but that is effective for quick startup.

ASSIMILATION OF INTEGRATED WATER VAPOUR FROM GROUND-BASED GPS OBSERVATIONS AT DWD

Maria Tomassini [12] and Christoph Schraff [1]

[1] Deutscher Wetterdienst, Kaiserleistr. 42,
63067 Offenbach, Germany
Tel. +49.69.80622711 - maria.tomassini@dwd.de

[2] GeoForschungsZentrum Potsdam

APPROACH

Integrated Water Vapour (IWV) derived from measurements by ground-based Global Positioning System (GPS) stations have been used in assimilation experiments at the Deutscher Wetterdienst (DWD). The GPS network used for the experiments consists of 76 stations located in Germany and in the Netherlands. Half-hourly GPS data are processed in near real time by the GeoForschungsZentrum (GFZ) and assimilated into the non-hydrostatic limited area model of DWD, namely the Lokal Modell (LM). The horizontal and vertical resolution of the LM used in this study is 7 km and 35 levels. The data assimilation scheme of LM is based on nudging towards observations. At present the LM analysis is produced with a continuous assimilation cycle using observations from synoptic stations, radiosondes and aircraft.
The method chosen for the assimilation of GPS data consists in relaxing the model IWV values towards the observed ones. A “pseudo-observed” profile of specific humidity based on the observed IWV and the vertical structure of the model humidity field is derived and then nudged at each single vertical level of the model.

DATA

Two assimilation experiments GPS and CNT, with and without GPS data respectively, were run during the period from 17 to 25 August 2001. Two 24 hour forecasts were produced every day, starting from the 00 UTC and 12 UTC analysis of the experiments. Unstable, stormy weather conditions, with a strong south-westerly circulation were prevailing in Central Europe until 19 August, and this was followed by a more stable period with a weak anticyclonic circulation. The period was selected because the LM did not perform well in forecasting precipitation.

RESULTS

The assimilation experiments confirm that the model IWV is relaxed towards the GPS IWV successfully during the assimilation cycle. For instance, the rms of the difference observed minus analyzed IWV at 00 UTC is reduced from the CNT value of 2.6 kg m^{-2} to 1.1 kg m^{-2} in the GPS. The forecasts of the experiments have been verified against the GPS IWV observations themselves and against upper-level observations from radiosondes. Both comparisons show that the assimilation of the additional GPS data has an effect within a forecast range of up to 15 hours. The rms error of the 12 hour forecasts (started from 00 UTC) against GPS IWV observations is 2.2 kg m^{-2} for the GPS experiment, 0.3 kg m^{-2} smaller than the CNT forecast.
The verification against radiosondes observations indicates positive impact of GPS in the 12 hour forecast of humidity, temperature, and wind, neutral impact on the geopotential, and neutral impact on all variables after 24 hours. Figure 1 shows the results from the upper-air verification of relative humidity and temperature for the 12 hour forecasts (15 cases). The improvement in the humidity rms error is mostly concentrated between 800 and 600 hPa and it is in the order of 2% relative humidity (10% improvement). A minor improvement is also detectable in the forecast of temperature and wind velocity. The mean error of the relative humidity is slightly larger for the GPS experiment. This can be related to a slight positive bias of the GPS data (0.6 kg m^{-2} for August 2001 with respect to the LM analysis). It is interesting to mention that

most of the improvement occurs in the period after 19 August. The absence of GPS stations upwind, i.e. south-west of Germany can be one cause of the little impact during the first spell.
An evaluation of the results has also been made comparing analyses and forecasts of precipitation with surface observations and radar images. The signal in the precipitation analysis fields is mixed, with the GPS data improving some bad cases but also tending to deteriorate the analysis in some areas without precipitation. The overall impact on the precipitation forecast (6 to 18 hour range) is neutral. In most of the cases the assimilation of GPS data does not considerably change the performance of the model. However, in few cases a small impact was found, not always positive.

Further work has to be dedicated to the tuning of GPS IWV nudging, especially to understand and correct cases of negative impact. For example, information on cloudiness could be used to improve the vertical distribution of the influence of the integrated value. Investigations on a possible bias correction of the GPS data are also required.

This work was supported under the grant of the German Federal Ministry of Education and Research (BMBF) No. 01SF9922/2.

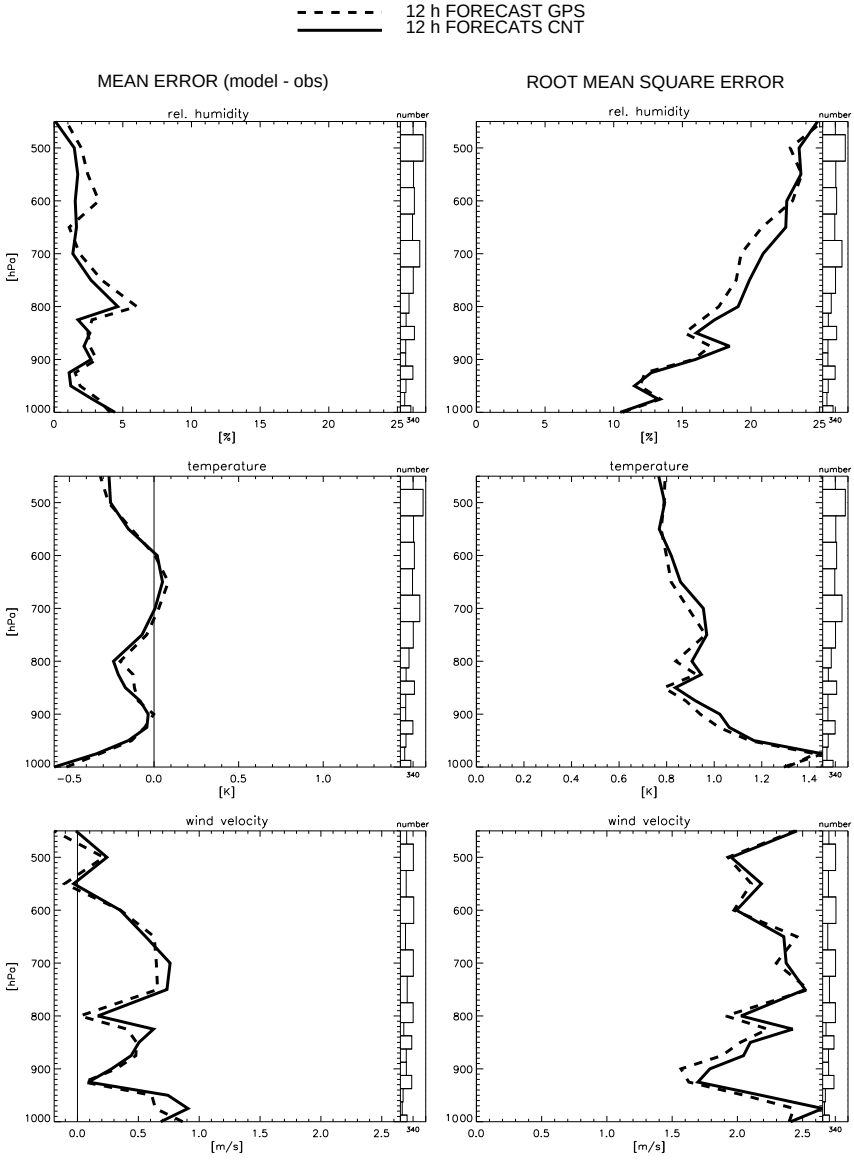


Figure 1. Impact of the assimilation of GPS data on the 12 hour forecast fit to observations. The mean error (left) and the root mean square (right) of the experiment with GPS data (dashed curves) and of the control (solid curves) are showed for relative humidity (top), temperature (middle) and wind velocity (bottom). The statistics have been computed for 15 cases using data from radiosondes in Germany and surroundings.

Using of the Chebyshev-Laguerre Polynomials for Representing the Vertical Structure of the Atmosphere

Tsvetkov V.I., Snopova O.V., and Rozinkina I.A.
(E-mail: tsvetkov@rhmc.mecom.ru, snopova@rhmc.mecom.ru, inna@rhmc.mecom.ru)
Hydrometeorological Center of Russia, 9-13 Bolshoy Predtechenskii per., 123242, Moscow, Russia

The use of analytical orthogonal functions for describing the horizontal structure of variables in NWP models has been widely used in numerical weather predictions. The vertical structure in such models is usually approximated by parameters specified at discrete levels that are irregular in height. The analytical representation of the vertical structure has not received much attention. An attempt to express the vertical structure in terms of finite sums of analytic orthogonal functions was made by Francis [1]. However, applications of his method were faced with certain difficulties because of the bulky formulas and repeated summations. This paper outlines the possible use of Chebyshev-Laguerre polynomials for representing the vertical atmospheric structure in numerical models.

The Chebyshev-Laguerre polynomials are orthogonal functions defined on the positive half-line and having the weight

$$h(\xi) = \xi^{\beta} e^{-\xi}, \tag{1}$$

where $\beta > -1,\ \xi \in (0, \infty)$.

These polynomials can be represented in terms of the Γ-function as

$$L_n(\xi, \beta) = \sum_{k=0}^{n} \frac{n!\,\Gamma(n+\beta+1)(-\xi)^{n+k}}{\Gamma(k+\beta+1)k!\,(n-k)!}, n \geq 0 \tag{2}$$

and satisfy the orthogonality condition

$$\int_0^{\infty} \xi^{\beta} e^{-\xi} L_q(x, \beta) L_r(\xi, \beta) d\xi = \begin{cases} 0, q \neq r \\ d_q = q!\,\Gamma(q+\beta+1), q = r \end{cases}. \tag{3}$$

Quadrature formulas for Fourier series expansions in terms of the Chebyshev-Laguerre polynomials can be found in [2, 3].

Note two features of the Chebyshev-Laguerre polynomials that can be used in meteorology. The first is that the domain of definition of these polynomials is the interval $(0, \infty)$, which can be interpreted as the thickness of the atmosphere. The second feature is associated with the recurrence relation for the derivative,

$$\frac{\partial L_n(\xi, \beta)}{\partial \xi} = n L_{n-1}(\xi, \beta + 1), \tag{4}$$

which can be used in numerical weather forecasting as follows. The geopotential Φ is represented as the Fourier series expansion in terms of the Chebyshev-Laguerre polynomials with parameter β. Then using the equation

$$\frac{\partial}{\partial \xi} \Phi + RT = 0, \tag{5}$$

where $\xi = -ln(p/p_0), p_0 = 1000 hPa,$ we can easily calculate the temperature *T*, because it has the same Fourier series but with the parameter $(\beta + 1)$.

Fourier series expansions of meteorological elements have been applied to the preparation of data for the Russian Hydrometeorological Center spectral atmospheric model. Numerical experiments on the transfer of the geopotential height Φ from *p*-system (17 vertical levels) to σ-system (31 vertical levels) and back were carried out. Thus, the root mean square error was found to decrease down to 10^{-8}hPa. Moreover, the quality of the model

outputs was improved, especially at the upper-tropospheric levels and higher. Figure 1 shows the anomaly correlation coefficients for the geopotential height at 100 hPa for the Northern Hemisphere. Curve *1* corresponds to the standard (spline) interpolation method, and curve *2* corresponds to Fourier series expansions in terms of the Chebyshev-Laguerre polynomials.

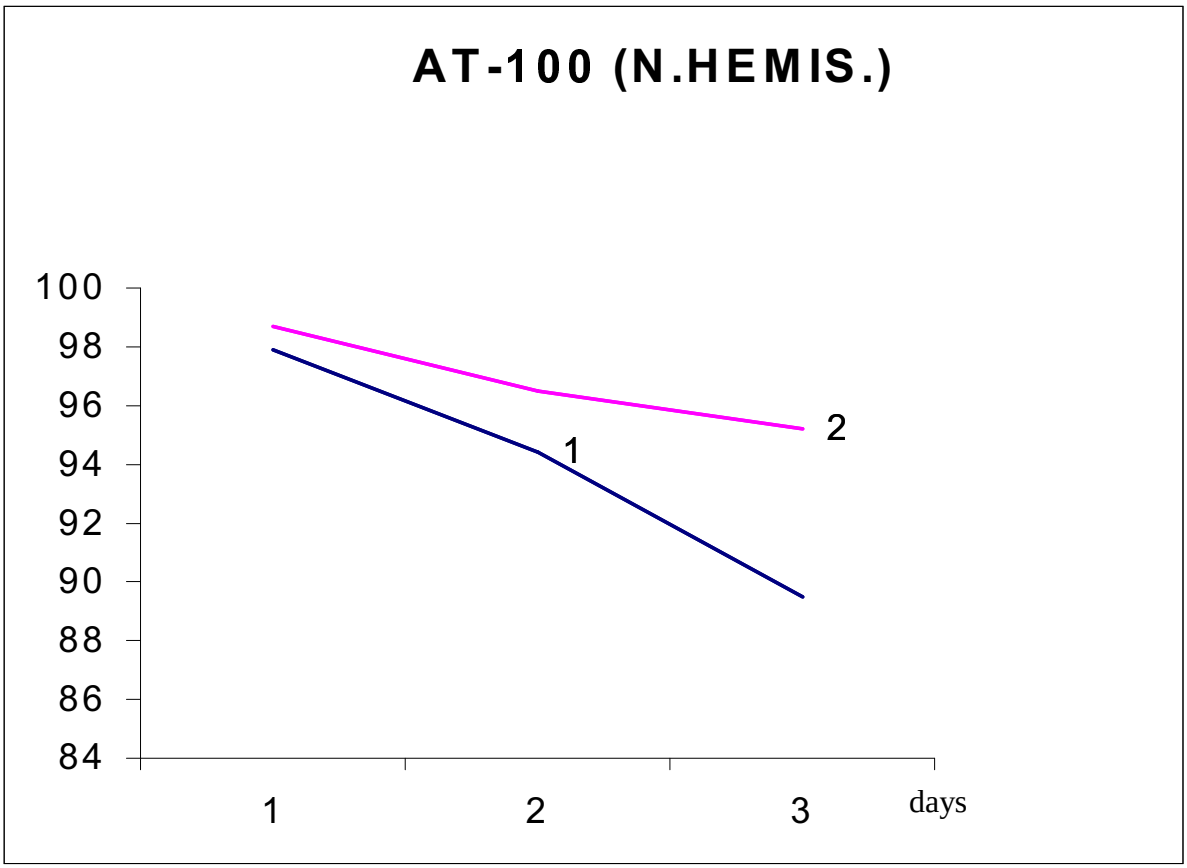


Fig. 1. Anomaly correlation coefficients for the geopotential height at 100 hPa in the case of input data preparation using (1) spline interpolations and (2) Fourier series expansions.

This study was financially supported by the Russian Foundation for Basic Research, grant nos. 00-05-64803, 01-05-65400, and 01-05-65493.

References

1. Francis, P.E., 1972, The Possible Use of Laguerre Polynomials for Representing the Vertical Structure of Numerical Models of the Atmosphere, *Quart. J. Roy. Met. Soc.*, vol. 98, no. 417, pp. 662-667.
2. Tsvetkov, V.I., 1990, The Roots of Orthogonal Polynomials in Numerical Integration Algorithms, *Acad. Sci. BSSR, Math. Inst.*, no. 1(401).
3. Tsvetkov, V.I., 1990, The Quadrature Formulas Based on Classical Orthogonal Polynomials, *Acad. Sci. BSSR, Math. Inst.*, no. 2(402).

Development of a variable-resolution global data assimilation system

M.D.Tsyroulnikov, M.A.Tolstykh[1], A.N.Bagrov, and R.B.Zaripov

Russian Hydrometeorological Research Centre, Moscow

[1] *Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow*

Corresponding author's email: *cyrulnic@rhmc.mecom.ru*

The data assimilation system is based on sequential "analysis-forecast" approach. At forecast step, the global finite-difference semi-Lagrangian NWP model [3] is used. It is designed in spherical geometry with a pole rotated to the region of interest. It is supposed to supplement the model in future with the variable resolution in (pseudo) latitude. The model includes the parameterizations of subgrid-scale processes developed in Météo-France for French operational model ARPEGE/IFS.

At analysis step, a 3-D multivariate Optimum Interpolation scheme is currently implemented. The analysis scheme is based on the sequential approach. That is, at each analysis cycle, the observations are divided into several batches with each batch being assimilated sequentially so that the "first-guess" at step $k+1$ is the analysis at step k. For this scheme to be optimal, the "first-guess" statistics should be recalculated at each step. Following the approach proposed in [4], in each batch but the last one, we include observation types with quasi-uniform and relatively dense coverage. As a consequence, we are allowed to approximately recalculate the "first-guess" statistics off-line [4]. Actually, we use the three batches: first, all SYNOP observations, second, the polar-orbiting temperature and humidity profiles (SATEM retrievals are used currently), and third, all other observations (radiosondes, aircraft winds, and satellite cloud-track winds).

The upper-air observations are assimilated by means of the incremental approach (e.g., [1]). In addition, sea-surface temperature is analyzed (using SHIP observations) following [2]. Soil temperature and soil moisture forecasts are corrected using the respective lowest-model-level analysis increments. The snow cover forecast is corrected using the "optimally" interpolated snow depth observations from SYNOP reports.

Although our data assimilation system is at the early stage of development, the first results are encouraging. The assimilation experiments are performed without pole rotation and variable resolution. The horizontal resolution is 1.5°, 28 vertical levels, the forecast-model time step is 36 min. We performed assimilation during the test period in February, 2000. Thirty consecutive forecasts started every 12 hours from the assimilated analyses are examined. The RMS errors of the 72 hours forecasts are shown in Figs. 1-2. Verification is against assimilated analyses, the area is Northern Hemisphere north of 20°. The similar results are obtained for verification against radiosondes (not shown). From Figs. 1-2, we see that the problems in our data assimilation system are primarily concentrated in the lower troposphere whereas in the upper troposphere and lower stratosphere, the results are more satisfactory.

This work is supported by Russian RFBR Grant 01-05-64582.

References

[1] Courtier P., Andersson E., Heckley W., Pailleux J., Vasiljević D., Hamrud M., Hollingsworth A., Rabier F. and Fisher M. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart. J. Roy. Met. Soc.* **124** (1998), 1783-1807.

[2] Reynolds R.W., and Smith T.M. A high resolution global sea surface temperature climatology. *J. Climate* **8** (1995), 1571-1583.

[3] Tolstykh M.A. Semi-Lagrangian high resolution model of the atmosphere for numerical weather prediction, *Russian Meteorology and Hydrology* 2001, N4, 1-9.

[4] Tsyroulnikov M.D. Three-dimensional assimilation of observations with uniform coverage using optimal spatial filtering. *Res. Activities in Atm. and Oceanic Mod.*, WMO, 1999, Rep. N 27, 1.40-1.41.

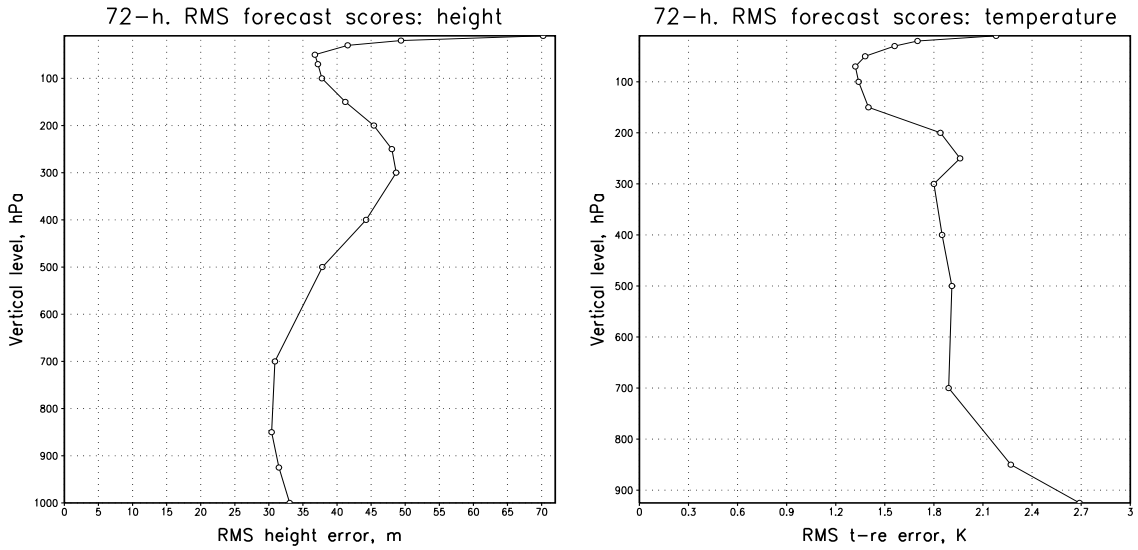


Figure 1:

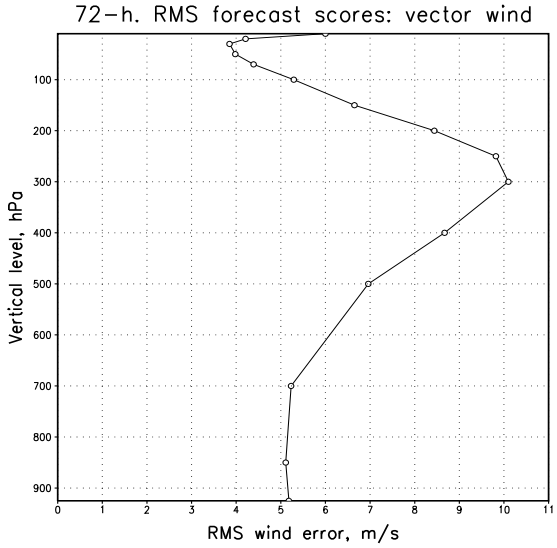


Figure 2:

SPARC Data Assimilation Working Group

William Lahoz[1], Ivanka Stajner[2], and Richard Swinbank[3]

[1]Data Assimilation Research Centre, University of Reading, Reading RG6 6BB, UK
[2]NASA/Goddard Space Flight Center, Data Assimilation Office, Code 910.3, Greenbelt, MD 20771, USA
[3]NWP Division, Met Office, Bracknell RG12 2SZ, UK
Corresponding author e-mail: wal@met.reading.ac.uk

To complement the WGNE activities in data assimilation, SPARC (Stratospheric Processes and their Role in Climate) has recently set up a working group on Data Assimilation. The emphasis of this group will be on stratospheric data assimilation, for application to numerical weather prediction, and monitoring of climate variations.

At the initial working group meeting, which was held in December 2001, it was decided to hold a SPARC Data Assimilation workshop in June 2002. The workshop will be hosted by the NASA/Goddard Data Assimilation Office through the Goddard Earth Sciences and Technology Center at the University of Maryland Baltimore County (UMBC), during 10-12 June. This workshop has a goal of bringing together scientists from different international centers (including the met agencies, space agencies and academia) and initiating comparisons between the assimilated products, so participants are invited from all groups working in this field.

All stratospheric products are of interest, but the main emphasis of this workshop will be on the lower stratosphere. Initial comparisons will encompass meteorology, chemistry and transport, with a particular focus on the assimilation of trace gases: ozone and water vapor. Presentations about ongoing work and future plans, such as assimilation of data from Envisat or the EOS satellites, are also welcome. Further details can be found on the web page http://dao.gsfc.nasa.gov/DAO_people/ivanka/sparcDA/. The local organiser of the workshop is Dr Ivanka Stajner, NASA/Goddard Space Flight Center, Data Assimilation Office, Code 910.3, Greenbelt, MD 20771, USA, email: ivanka@dao.gsfc.nasa.gov.

For further details about the group, contact Dr William Lahoz, Data Assimilation Research Centre, University of Reading, Reading RG6 6BB, UK, email: wal@met.reading.ac.uk, or Dr Richard Swinbank, NWP Division, Met Office, Bracknell RG12 2SZ, UK, email: richard.swinbank@metoffice.com.