Title: Particle filters for large-dimensional problems

Authors: Melanie Ades, Peter Jan van Leeuwen (University of Reading)

With ever increasing model resolution and more complicated observations the data-assimilation problem becomes more and more nonlinear. This calls for fully nonlinear data-assimilation methods, such as particle filters. In particle filters the importance of each particle in estimating the posterior density is dominated by the likelihood of that particle. In high-dimensional systems with a large number of independent observations the likelihood can differ substantially between particles resulting in only a few having statistical significance.

The idea of using the proposal density within the particle filter to provide a continuous guidance towards a future observation has already been discussed in the literature. Using the proposal density the aim is to increase the likelihood of all particles by ensuring they end up significantly close to the observation. However the proposal density is not restricted to continuous guidance but offers a much greater freedom in how we treat the particles. In particular it can be used to ensure that the majority of particles have an approximately equal significance in estimating the posterior density. With the majority of particles being both close to the observations and having statistical significance, the ability to represent a multi-modal posterior density with only a few particles starts to be realised.

We will show how such a particle filter is derived and discuss its application to the barotropic vorticity equations, both in the regime with periodic trajectories and in the chaotic regime, with state dimensions of a few thousands and more.

Title: Adjoint observation impact for a limited area model

Authors: Clark Amerault, James D. Doyle (Naval Research Laboratory, Monterey, CA, USA)

Many global numerical weather prediction systems employ adjoints of a forecast model and data assimilation system to monitor the impact of observations on forecast error. Recently, the adjoint observation impact methodology has been applied to the US Navy's atmospheric limited area forecast model and 3D-Var data assimilation system. This presentation will focus on aspects of the adjoint observation impact system that are unique to a limited area model. Similar to global models, large reductions in forecast error are due to conventional observations from soundings and aircraft. However, the forecast error is much more variable when compared to a global system due to weather features moving into and out of the limited area model's domain. The importance of observations relative to one another is dependent on

the physical location of the model's domain. For example, when compared to radiosondes, cloud winds are more important for a domain centered over the eastern north Pacific Ocean than the continental United States. Furthermore, the area over which the forecast error is calculated must be carefully chosen so that the information is not lost to the lateral boundaries during the adjoint forecast model integration. Results from impact calculations of differing resolutions will also be presented.

Title: The predictability of North American land-falling cyclones

Authors: Brian Ancell (Texas Tech University), Lynn McMurdie (University of Washington), Rolf Langland (Naval Research Laboratory)

Land-falling North Pacific storms often impact the west coast of North America with strong winds, heavy precipitation and large mountain snowfall. These storms are frequently poorly predicted by operational models despite continued improvements in model resolution, model physics and data assimilation. Previous studies have shown that short-term (24-72hr) forecast errors of sea level pressure along the U.S. west coast are statistically related to large-scale upper-level flow regime. We expand on this prior work by using other methods to define predictability and relating this predictability to the specific characteristics of the cyclones themselves. Specifically, we examine 12- to 48-hr forecasts of land-falling North American west coast cyclones over 2 winter seasons. The tools used to assess predictability of the cyclones are 1) ensemble sensitivity produced from a WRF-model ensemble Kalman filter (EnKF), 2) adjoint sensitivity produced with an MM5 adjoint model, and 3) spread growth in the cyclone sea-level pressure field within the same WRF-model EnKF. We use these tools to further distinguish why certain cyclones may exhibit poor predictability, and whether large intrinsic potential for error growth or slower-growing, larger errors present at initialization are more important. Results over the 2-year period are shown, and plans to extend this work beyond this initial examination are discussed. The goal of this project is to determine some of the factors that contribute to predictability of a variety of weather regimes on multiple scales. Determining these relationships could be very important to address the modeling/data assimilation configuration during specific weather regimes associated with poor prediction of cyclones.

Title: Observation Impact on Forecast Error in a Regional Model

Authors: Thomas Auligné, Hongli Wang, Xin Zhang, Xiaoyan Zhang, Qingnong Xiao, Xiang-Yu Huang (National Center for Atmospheric Research)

A new capability to calculate the sensitivity of forecast error to observations, based on adjoint techniques, has been included into the Weather Research and Forecasting (WRF) model and data assimilation suite. This presentation will focus on an effort to understand some of the main sources of uncertainty in the calculation of observation impact. We believe that most of the methodology is applicable to all systems focusing on adjoint sensitivity and observation impact.

The reference for the calculation of the forecast error is often chosen as the system own analysis after 24 hours, assuming it is uncorrelated with the analysis at initial time. We will assess the sensitivity of the observation impact to the choice of reference. We will then compare results using several norms for the forecast error. In particular, we will introduce a new verification using future observations as reference and we will show its connection with another common metric called the Degree of Freedom for Signal (DFS). The ability of the tangent-linear (resp. adjoint) code to propagate real-size perturbations (resp. sensitivities) will be evaluated and we will study the amount of information lost through the lateral boundary conditions.

Finally we will show an attempt to predict the impact of observations on forecast error. This can be very useful for operational systems where some of the detrimental observations could be discarded a priori.

Title: Optimal linearization trajectories

Authors: Jan Barkmeijer, Roel Stappers (KNMI)

Nowadays linear models play an important role in various components of NWP, such as ensemble forecasting and variational data-assimilation. Nonlinear model behaviour, however, can have severe implications for the usefulness of these linear models. We show that many types of nonlinearity, such as advection, can be taken into account in the tangent linear model by a suitable choice of the linearization trajectory. Using this optimal linearization trajectory we show that the tangent linear model can be used for more than 200 days in a 3-level quasi geostrophic model and more than (the equivalent of) 150 days in the Lorenz 96 model. We introduce an iterative method, purely based on tangent linear integrations, that converges to this optimal linearization trajectory. As such, this iterative method can be used to account for nonlinearity in estimation problems without using the nonlinear model. We will discuss the use of the optimal linearization trajectory in the context of data-assimilation and adjoint observation impact studies.

Title: Variational ensemble data assimilation at Meteo-France for error covariance modelling and ensemble prediction

Author(s): Loik Berre, Gerald Desroziers, Laure Raynaud, Hubert Varella, Laurent Descamps, Carole Labadie

Since July 2008, a variational ensemble data assimilation system (EnVar)has been operational at Meteo-France. It is used to calculate flow-dependent background error covariances for optimizing data assimilation. It also provides perturbed initial states for the Meteo-France ensemble prediction system. Prominent features and developments of this EnVar will be presented.

The ensemble approach has been chosen to be consistent with an error simulation technique of the deterministic 4D-Var data assimilation cycle. Spectral filtering techniques based on objective signal and noise estimates are applied for modelling background error standard-deviations, and wavelet filtering techniques are considered for correlations. Moreover, innovation-based variance estimates are compared to ensemble-based variances to get information about model errors. Acceleration and parallelization opportunities will eventually be discussed.

Title: Errors in ensemble-based error covariance estimates and Hybrid ensemble 4D-VAR

Author(s): Craig H. Bishop (Naval Research Laboratory, Monterey, CA), Elizabeth Satterfield (Naval Research Laboratory, Monterey, CA), David D Kuhl (Naval Research Laboratory, Washington DC),Tom Rosmond (Science Application International Corp., Forks, WA)

Ensemble Kalman filters produce predictions of flow dependent error covariances. Is it possible to empirically measure the accuracy of such predictions? How should such inaccuracy be accounted for in data assimilation? To address these questions, we use the Lorenz 40-variable model to help motivate an even simpler univariate forecasting system in which (i) the truth is a random draw from a climatological distribution of true states (ii) the forecast is the truth plus a random error drawn from a distribution having the true error variance (iii) the true error variance is a random draw from an inverse-gamma distribution, and (iv) the ensemble variance is a random draw from a gamma distribution whose mean is a linear function of the true forecast error variance. Bayes' theorem is used to obtain the distribution of true error variances given an ensemble variance. From this distribution, all aspects of the accuracy of the error variance prediction can be measured. Indeed, the approach allows the accuracy to be measured in terms of "effective ensemble size". It is found that the minimum error variance estimate of the true error variance given an ensemble variance is a weighted linear combination of the climatological variance and the ensemble variance. Hence, the result provides theoretical support for the use of linear combinations of static and ensemble based covariance models in Hybrid 4D-VAR schemes. If the univariate model's assumptions are valid, optimal weights for the ensemble and climatological variances can be recovered from large archives of (innovation, ensemble variance) pairs from any ensemble forecasting system. We make the ansatz that optimal weights for variances will make good weights for the static and ensemble based covariances of Hybrid ensemble 4D-VAR. Low resolution tests using a Hybrid ensemble version of the Navy's NAVDAS-AR data assimilation scheme have supported our ansatz. Limitations of our approach and possible ways of improving upon it will also be discussed.

Title: Quantification of divergence in a mesoscale model

Authors: Vanja Blazica, Nedjeljka Zagar (University of Ljubljana, Ljubljana, Slovenia)

The importance of divergence increases as the scales of atmospheric phenomena decrease. In the mid-latitudes, divergence is an essential ingredient of many mesoscale processes. Its analysis and reliable prediction remains a challenge for mesoscale NWP models. The analysis of divergence in the mesoscale data assimilation systems and associated model errors is a subject of active research. This study aims to quantify the divergent energy on the mesoscale based on the state-of-the-art NWP model ALADIN. The goal is to provide picture of average distribution of divergence versus vorticity as a function of the horizontal scale and the vertical level. Furthermore, it is studied whether the same average percentage of divergence with respect to vorticity, that is present in the model forecast, is characteristic also for the short-range forecast errors.

Title: A hybrid variational-EnKF data assimilation technique for operational global NWP

Authors: Mark Buehner (Meteorological Research Division, Environment Canada)

A new variational data assimilation technique is now being tested to possibly replace the Canadian operational 4D-Var for global NWP. The new approach relies on the four-dimensional ensemble-based background error covariances produced by the Ensemble Kalman Filter (EnKF) and a variational minimization to produce a single deterministic analysis. This approach is well-suited for producing the high-resolution analyses required to initialize global deterministic medium-range forecasts because: 1) it can efficiently assimilate the very large volume of observations currently available, 2) it is likely more easily parallelized than 4D-Var over a very large number of processors, and 3) it is more flexible with respect to modelling the background error covariances than traditional EnKF approaches. Preliminary verification results obtain with the new approach using both pure EnKF covariances and covariances obtained by averaging the EnKF-based and traditional NMC-based covariances will be shown. These results will be compared with those obtained using the 4D-Var and EnKF approaches.

Title: Representing model error in Ensemble Data Assimilation

Authors: Carla Cardinali, Roberto Buizza (ECMWF), Gabor Radnoti (ECMWF), Nedjeljka Zagar(Lubjana University)

The paper investigates a method to represent model error in an ensemble assimilation (EDA) system. The ECMWF operational EDA simulates the effect of both data and model uncertainties. The data errors are represented by adding perturbations with statistical characteristics implied by the observation error covariance matrix whilst the model uncertainties are represented by adding stochastic perturbations to the physical tendencies to simulate the effect of random errors in the physical parameterizations (ST-method). In this work an alternative method (XB-method) is proposed to simulate model uncertainties, based on adding to the model background field, perturbations with statistical characteristics defined by the model background error covariance matrix. EDAs with similar data uncertainties but different model error representations are compared and the proposed XB-method is found to lead to the largest spread both in amplitude and spatial scale. Normal-mode diagnosis not only has confirmed that XB-EDA methodology produces larger spread at all scales but also that the spread projects more than the other onto the unbalanced part of the motion. The different EDAs have been employed to define the background error covariance matrix to be used in a higher-resolution deterministic assimilation system. Specific diagnostics have been applied to estimate the quality of the modelled background error covariance matrix from the different ensembles presented. Results have shown that XB method produces a background error covariance matrix that allows the assimilated observations to be more influential. Consequently, despite the fact that all the other ensemble methodologies apply an artificial global inflation whereby background error variances are inflated by a certain factor, all the corresponding deterministic assimilation systems result having a considerably smaller degree of freedom for signal. XB-EDA based background error statistics can therefore be used in data-assimilation without any artificial inflation that is currently used in the operational ensemble analysis.

Title: How to optimally treat large scale information in limited area ensemble-based data assimilation?

Author: Jean-François Caron (Met Office, Reading, UK)

The Met Office is developing a convective scale version of the Met Office ensemble prediction system, MOGREPS, where the analysis perturbations are produced using an Ensemble Transform Kalman Filter (ETKF). The primary goals of this work are 1) to conduct examinations of 1-hour forecast error covariances for the benefit of a NWP-based nowcasting system currently under development at the Met Office and 2) to perform predictability studies of localized weather at very short time scale. This convective scale ETKF system uses a model with a grid resolution of 1.5 km covering southern United Kingdom and has an hourly cycle. The lateral boundaries are provided by the 18 km and 12-hourly cycling regional component of MOGREPS which cover the North Atlantic and Europe.

In this poster we will present the issues we faced due to the discontinuities introduced by the current ETKF method between the analysis perturbations and the perturbations at the lateral boundaries. To reduce theses discontinuities, we developed and tested a so-called scale-selective ETKF approach where the ETKF transform matrix is applied only to the small scale component of the 1.5km forecast perturbations while the large scale component of the analysis perturbations is taken directly from parent EPS. The changes implied by this new approach in terms of both ensemble forecast performance and forecast error covariances will be presented. Finally, we will discuss on the implication of our results for limited area ensemble-based data assimilation.

Title: Challenges of assimilating observations for producing high-resolution sea-ice analyses.

Author(s): Alain Caya, Mark Buehner, Michael Ross, Tom Carrieres (Environment Canada)

Assimilation of observations with different resolutions for producing high-resolution sea-ice analyses offers some interesting problems. Satellite borne passive microwave sensors are the primary source of information and complementary to higher resolution data obtained from other sensors (e.g. Synthetic Aperture Radar). Our analysis grid has a higher resolution than the passive microwave data and therefore an appropriate observation operator is required to avoid contaminating high resolution analysis details. A spatial averaging footprint operator is used to map fields from the analysis grid to the passive microwave observation locations. Another interesting aspect of sea-ice analysis is that the ice concentration analysis variable is bounded between zero and one. This violates the assumption of Gaussian distributed errors. Moreover, the sea-ice concentration field is discontinuous and often undergoes abrupt transitions between zero and one at the ice edge. This makes the background-error horizontal correlation particularly difficult to model. The problem is demonstrated using a simple Gaussian correlation function. A technique using displacement error as a control variable is proposed to partially overcome this challenge.

Title: Adapting the UKMO linear model for NWP-based nowcasting

Authors: C. L. Charlton-Perez, S. P. Ballard, Z. Li , D. Simonin, H. Butter, N. Gaussiat, L. Hawkness-Smith

The UK Met Office is developing an NWP-based nowcasting system which will operate at high resolution in space (1.5 km) and time (0-6 hourforecast). The main aim is to forecast convective storms and then use these forecasts to predict floods. We will present results from testing the ability of the linear model used for 4D-Var to represent the dynamics and microphysics of the nonlinear model at convection-permitting scales.

Our team's work is focussed on the assimilation of high resolution, novel observation types to enhance this nowcasting system. These observation types include Doppler radar winds, radar reflectivity, radar-derived rain rates and in the future may include radar refractivity. Initial trials have used a data assimilation system run with hourly cycling either in 3D-Var or 4D-Var mode along with latent heat nudging of radar-derived precipitation rates and humidity nudging based on 3D cloud cover analyses. For the trials the model domain covers southern England and Wales with boundary conditions provided by a 4km resolution forecast with a domain covering the whole of the UK.

Unlike the Unified Model at coarser resolutions, both the linear and nonlinear models in the nowcasting system operate without a convective parameterization. Therefore, we subject the linear model to a series of linearization tests to confirm that the linear model is simulating the nonlinear model well at the nowcasting scales. Several areas fori mprovement have been identified using measures of linearization error. In the linear model, parameters for both the precipitation efficiency and the boundary layer scheme for vertical diffusion of moisture and temperature are being re-tuned. The treatment of moisture in the linear model microphysical scheme is being re-assessed in conjunction with the assimilation of radar reflectivity which demands a different model operator for reflectivity of ice vs. liquid water.

Title: Image assimilation with the weighted ensemble Kalman filter

Authors: Anne Cuzol (University of South Brittany, INRIA), Rennes Bretagne (Atlantique Etienne Mémin, INRIA), Rennes Bretagne (Atlantique)

We present a sequential data assimilation method based on the combination of two assimilation techniques: the ensemble Kalman filter (EnKF) and the particle filter. Both techniques are based on Monte Carlo sampling allowing approximate solving of non linear stochastic filtering problems. However, while the EnKF is still based on a Gaussian assumption, the particle filter does not rely on such an assumption but is known to be less efficient when the number of available samples is small. In practice, both techniques are combined in the sense that the sampling step of the particle filter is based on the EnKF technique, followed by a weighting of samples using observations. The association of these two approaches is a step toward an efficient application of ensemble techniques to high-dimensional and non linear / non Gaussian problems, such as those encountered in meteorology or oceanography. We show the

performances of this new approach on high-dimensional problems where the goal is to filter turbulent velocity fields from image observations. The assimilation technique associates a non linear stochastic dynamical model to linear observations extracted from the image sequences, or directly to the image data through a non linear observation operator. The method has been validated on a synthetic sequence, and applied to real oceanographic satellite image sequences of SST (sea surface temperature). This work corresponds to an extension of the preliminary study published in Tellus (N. Papadakis, E. Memin, A. Cuzol, N. Gengembre. Data assimilation with the Weighted Ensemble Kalman Filter. Tellus A, vol.62(5), p. 673-697, 2010). In a second part, we present a way to improve the assimilation when the time step between observations (images) is very long. We make use of a conditional simulation technique in order to reduce dynamical discontinuities produced in that case by the sequential techniques.

Title: Accelerating and parallelizing minimizations in ensemble and deterministic variational assimilation

Author(s): Gerald Desroziers, Loik Berre (CNRM/GAME, Meteo-France)

EnKF can be transposed to variational ensemble assimilation, where a set of perturbed variational analyses are performed. In this case, however, there is an evident important additional cost associated to the use of multiple minimizations. The aim of the presentation is to investigate different techniques to reduce the cost of the multiple minimizations that have to be performed.

The use of a preconditioning technique, based on Ritz eigenpairs resulting from a first minimization performed by a combined Lanczos / conjugate-gradient, algorithm is in particular investigated. The possibility to construct more explicitly a new perturbed solution, with Lanczos vectors issued from a single prior unperturbed or perturbed minimization, is also studied. This appears to provide a first significant reduction in the cost of the new minimization.

Finally, a new approach is proposed to generalize the previous idea to the use of a multiple set of Lanczos vectors issued from an ensemble of perturbed assimilations. The application of this procedure to a simplified analysis problem shows encouraging results, as it appears to be a possible way for reducing the global cost of an ensemble variational assimilation. Moreover, this seems to provide an efficient strategy for parallelizing such an ensemble variational assimilation but also the deterministic variational assimilation itself.

An application of the previous approach to the 4D-Var assimilation ensemble associated with the French global prediction system will also be shown.

Title: Multi-scale sensitivity and predictability aspects of a severe European winter storm

Authors: James D. Doyle, Clark Amerault, Carolyn A. Reynolds, P. Alex Reineceke (Naval Research Laboratory, Monterey, CA, USA)

A severe winter storm, referred to as Xynthia, crossed Western Europe on 26-28 February 2010 and was the most intense in this region in more than a decade. The violent storm claimed the lives of more than 50 people with many of the deaths in France related to strong winds and a storm surge. The storm produced heavy rains, and hurricane-force winds, which caused widespread power failures and severely impacted the transportation system. The insured losses from the storm were estimated to be $2-4 billion.

In this study, the adjoint and tangent linear models for the atmospheric portion of the nonhydrostatic Coupled Atmosphere/Ocean Mesoscale Prediction System (COAMPS) are used to explore the mesoscale sensitivity and predictability characteristics associated with Xynthia. Unique aspects of the adjoint modeling system include a full adjoint to the microphysics and a nested grid capability that allows for multi-scale sensitivity calculations. The adjoint is applied using the nesting option with 45 and 15 km meshes. Results indicate that 36-h forecasts of intensification of the extratropical cyclone in Western Europe are very sensitive to the initial state moisture along a moist plume. The adjoint model indicates sloped maxima in the wind, thermal, and microphysical sensitivity fields oriented along a mid-tropospheric front, which ultimately influence the cyclone structure and intensification rate. Adjoint optimal perturbations that are scaled to the size of typical observational errors (1 m/s, 1 K) exhibit rapid growth rates in the near-surface horizontal velocity of 30 m/s and substantial structural changes to the low-level velocity in the region of surge. Implications of the sensitivity fields for the predictability of mesoscale aspects of Xynthia will be addressed, as well as the generality of the results. Challenges for applying adjoints at high resolution (e.g., nested grids, microphysics) will be highlighted.

Title: Using a Bi-Conjugate Gradient minimization algorithm for variational data assimilation.

Authors: Amal EL Akkraoui, Ricardo Todling, and Yannick Tremolet.

Presently, a preferred minimization for strong constraint 4dvar uses a Lanczos-based CG algorithm. This requires the availability of the square-root of the background error covariance matrix. In the context of weak constraint 4dvar, this requirement might be too restrictive for the formulations of the model error term. It might therefore be desirable to avoid a square-root decomposition of the augmented background term. In this respect, an appealing minimization scheme is the double CG minimization employed, for example, in the Grid-point Statistical Interpolation analysis (GSI). However, this procedure cannot easily take advantage of the preconditioning using the eigenvectors of the Hessian, readily available in the Lanczos-based CG. Alternatively, this could be exploited by considering a bi-conjugate gradient (Bi-CG) algorithm, which by construction is suitable to non-symmetric matrices.

The present work introduces the mathematical formulation of the Bi-CG and a Lanczos-biorthogonalization procedure by which the left- and right-Hessian eigenvectors are calculated for use in the preconditioning. Implementation of the scheme is done in the context of GSI, and preliminary studies are presented for the 3dvar. Also, three minimization strategies are compared in GSI: the double CG (with an added re-orthogonalization step for improved performance), the square-root(B) Lanczos-based CG, and the new Bi-CG algorithm. Results show that the three algorithms converge with the same rate and to the same solution. Despite the additional computational cost, the importance of the re-orthogonalization step is shown to be fundamental, especially for the BiCG scheme. Furthermore, when using the eigenvectors for preconditioning, the BiCG behavior is comparable to that of the Lanczos-CG algorithm. Both schemes construct the same approximation of the Hessian with the same number of eigenvectors, and benefit in the same way of the reduction of the condition number. The efficiency, computational cost,and stability of the three algorithms are discussed.

Title: Data assimilation into groundwater contaminant models using an ensemble variational approach.

Authors: M. EL GHARAMTI (King Abdullah University of Science and Technology, KAUST, Thuwal, KSA), U. ALTAF (King Abdullah University of Science and Technology, KAUST, Thuwal, KSA), I. HOTEIT (King Abdullah University of Science and Technology, KAUST, Thuwal, KSA), A. W. HEEMINK (Delft Institute of Applied Mathematics, Delft University of Technology, Delft, Netherlands).

The adjoint method has been used very often for variational data assimilation. Using the available data, the control variables in the model are identified by minimizing a certain cost function that measures the differences between the model outputs and the data. In order to obtain a computationally efficient procedure, the minimization is performed with a gradient-based algorithm where the gradient is determined by solving the adjoint problem. The computational cost to run the adjoint model often exceeds several forward model runs which in return is computationally inefficient. The second drawback of the adjoint method is the programming effort required for the implementation of the adjoint model code.

We propose a new variational data assimilation approach based on model reduction using Proper Orthogonal Decomposition (POD), which avoids the implementation of the adjoint of the tangent linear approximation of the original nonlinear model. An ensemble of the forward model simulations is used to determine the approximation of the covariance matrix and only the dominant eigenvectors of this matrix are used to define a model subspace. By projecting the original model onto this subspace, an approximate linear model is obtained. The adjoint of the tangent linear model is replaced by the adjoint of this linear reduced model. Thus the adjoint model is run in reduced space with negligible computational cost. Once the gradient is obtained in reduced space it is projected back in full space and the minimization process is carried in full space.

In the presentation we will introduce the ensemble approach to variational data assimilation. The characteristics and performance of the method will be illustrated with a number of real life data assimilation applications with ground water contaminant models. For such models, we will develop a procedure to estimate some important model parameters such as the permeability of the underground geologic rocks in addition to various subsurface contaminant state estimates.

Title: A high-level abstraction for developing adjoint models

Authors: P. E. Farrell, S. W. Funke (Imperial College London)

Adjoint models are extremely useful for many scientific challenges, but the development of adjoint model is typically very difficult. Algorithmic differentiation offers one strategy for simplifying this task: it takes the abstraction that a model is a sequence of primitive instructions, each of which may be differentiated in turn. While extremely successful, this low-level abstraction runs into difficulties in cases of practical interest (differentiating through linear solves, model I/O, calls to external libraries, language features that are unsupported by the AD tool, etc.). In this work, we suggest an alternative, complementary, higher-level abstraction: that a model is a sequence of linear solves. By annotating the model source code with library calls that build a tape of the operators involved and their dependencies, and supplying callbacks to compute the action of these operators, the task of correctly deriving the adjoint model is made significantly easier. We demonstrate this with an application to Fluidity, an extremely complex multiphysics package under development at Imperial College London.

Title: Assimilation of MODIS and AMSR-E Snow Parameter Observations into a Physical Snow Model

Authors: Steven J. Fletcher[1], Glen E. Liston[1], Christopher A. Hiemstra[2] and Steven D. Miller[1].

Affiliation: [1]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado, USA.

[2] Cold Regions Research and EngineeringLaboratory (CRREL), US Army Corps of Engineers, Fort Wainwright, Alaska, USA.

The assimilation of binary (yes/no) type observations of non-state variables, as well as lower resolution observations of state variables, into any data assimilation scheme poses a significant problem to the scheme. However, these data may still be of use to the assimilation scheme. In the problem of modeling the evolution of snow there are high spatial resolution snow cover observations from the Moderate Resolution Imaging Spectroradiometer (MODIS), which is not a state variable of the numerical model but is a binary function of the state variable, Snow Water Equivalent (SWE). An addition problem with these data is that this binary function is non-unique, as a positive snow cover signal will correspond to any value of SWE potentially. Direct observations of SWE are available from the Advanced Microwave Scanning Radiometer - for the Earth Observing System (AMSR-E) but are at a coarse spatial and temporal resolution, introducing another non-unique function between SWE and the a set of observations. In this

paper we present some data assimilation schemes to assimilate MODIS snow cover and AMSR-E SWE observations; firstly on their own into a high resolution physical snow model and then through different combinations of the two. The results from these experiments help demonstrate that we can assimilate binary type observations if we can constrain the non-uniqueness and that they have a positive impact on the modeled snow evolution and to have better agreement with future MODIS snow cover observations.

Title: Measures of observation impact in non-Gaussian data assimilation

Author(s): Alison Fowler, Peter Jan Van Leeuwen (Department of Meteorology, University of Reading, UK)

The impact of new observations in a data assimilation system is particularly important in the geosciences, where observations can be expensive and limited. Many measures of the impact of observations on both the analysis and the forecast currently exist. However, in applying these measures operationally assumptions about the linearity of the model and the observation operator are often necessary. These assumptions may have limited value with current and future highly non-linear models and non-linear data assimilation methods. Consequently our understanding of the potential of new observations may also be affected. The aim of this work is to highlight such cases when the non-Gaussianity of the prior becomes large enough such that impact measures based on linearizations are no longer useful. Four measures of the observations impact on the analysis will be studied; the sensitivity of the analysis to the observations, the degrees of freedom for signal, mutual information and relative entropy. In a Gaussian framework the first three of these measures can be shown to be calculable from the ratio of the background and observation error variances, for relative entropy the exact values of the background and observations and their error variances must also be known. Idealised experiments will show that representing the prior beyond the first and second moments can have a large impact on these measures, in some cases making them complicated functions of the characteristics of the prior and likelihood distributions. These results will then be put into context through the use of particle filter techniques.

Title: Preconditioning of conjugate-gradients in observation space with an application to 4D-Var data assimilation

Author(s): Serge Gratton (ENSEEIHT, Toulouse, France), Selime Gurol (CERFACS, Toulouse, France), Philippe Toint (University of Namur, Belgium)

We investigate the trucated Gauss-Newton (TGN) method for the solution of 4D-Var data assimilation problem. TGN relies on the approximate solution of the quadratic approximations of the nonlinear least-squares cost function. The well-known method to solve the quadratic approximations which can be formulated as linear least squares problems is via conjugate-

gradients. This technique is known as incremental 4D-Var data assimilation among the community.

This study considers an alternative conjugate-gradient-like method called Restricted Preconditioned Conjugate Gradient (RPCG). RPCG performs the quadratic minimization in observation space and produces mathematically equivalent iterates to that of incremental 4D-Var approach in exact arithmetic. This algorithm is computationally interesting whenever the dimension of the observation space is significantly smaller than that of the state space since it reduces the computational cost. The relation between RPCG and the other well-known observation space solver called PSAS whose iterates does not provide monotonic decrease of the cost function in state space is explained.

The techniques used to have a better convergence like subspace preconditioning or re-orthogonalization considered to be expensive when used by incremental approach becomes possible to implement when using RPCG while keeping the convergence properties in state space. In this study, we focus on warm start Quasi-Newton preconditioners which are generated using by-products of RPCG. We derived the refined convergence bounds which results in a practical solution method. We illustrate the results using a toy-problem based on heat equations.

Title: Sensitivity patterns of sub-ice shelf melt rates to ocean circulation under Pine Island Glacier from an adjoint ocean general circulation model

Authors: Patrick Heimbach(1) and Martin Losch(2)

(1): MIT/EAPS, Cambridge, MA, USA (2): AWI, Bremerhaven, Germany

Melt rates underneath floating ice shelves are difficult to observe directly. Estimates rely mostly on indirect methods and are subject to large uncertainties. These uncertainties are a serious gap in the connection between observed oceanic changes and ice shelf thinning. The present study is a first step toward addressing this issue in the context of inverse modeling. We examine to which extent ocean hydrographic observations away from the ice-ocean boundary can be used to constrain sub-ice shelf melt rates. To this end, we derive comprehensive sensitivity patterns of sub-ice shelf melt rates to changes in ocean circulation. Our study region is the Pine Island Ice-Shelf, a well-contained region in the Amundsen Sea Embayment, and well-suited for this feasibility study. The sensitivity patterns are computed with an adjoint model of the MITgcm, a full-fledged ocean general circulation model that resolves the sub-ice shelf circulation and includes a thermodynamic melt rate parameterization. The adjoint state can be used to identify dominant water mass pathways and time scales that affect melt rates, provide guidance for oceanographic field campaigns for deploying limited measurement assets in an optimal manner, and establish the feasibility of connecting hydrographic observations to constrain melt rates in formal estimation approaches such as those undertaken within the Estimating the Circulation and Climate of the Ocean (ECCO) consortium.

Title: Jacobians of the GEOS5 Relaxed Arakawa-Schubert convection scheme.

Authors: Dan Holdaway, Ron Errico (GESTAR/GMAO, Goddard Space Flight Center.)

In order to solve the minimization problem that is inherent in 4D variational data assimilation the adjoint of the numerical forecast model is required. Moist convective processes in the atmosphere can exhibit highly nonlinear behavior, meaning the tangent linear model, and thus the adjoint, would not always provide a suitable approximation for use in the minimization problem. As well as the inherent nonlinear interactions that occur in convective plumes the numerical representation of convection generally consists of a number of nonlinear discrete switches. Due to the high degree of nonlinearity tangent linearizations of the complete convection schemes, as used in modern global forecast models, are generally deemed inappropriate; for the adjoint calculation a much simpler scheme is usually considered. The aim of these simpler schemes is to capture the residing behavior but in a way that can be accurately linearized. In order to quantify the nonlinearity in the RAS convection scheme that is used in NASA's GEOS5 model, and thus proceed onto constructing a suitable version of the parameterization for use in the assimilation, the Jacobians are examined. The Jacobians are obtained by comparing the difference between the output of the RAS scheme when perturbing the model state and the output when using a control model state. The Jacobian structure, eigenvalues and magnitudes are examined to gauge sources of sensitivity and nonlinearity in the scheme and provide insight into the ways in which the scheme may be suitably simplified for use in the 4D variational data assimilation.

Title: Dynamical sensitivity analysis of tropical cyclogenesis: a barotropic mode in the eastern Pacific

Authors: Brett Hoover, Michael Morgan (University of Wisconsin)

Tropical cyclones (TCs) in the eastern Pacific often form on the cyclonic shear-side of a persistent zonal jet extending from the surface to the lower troposphere; the existence of this jet implies the possibility for barotropic energy conversion as a mechanism for the formation of the ambient low-level vorticity necessary for tropical cyclogenesis. Such a scenario provides for the possibility a 'barotropic mode' in eastern Pacific tropical cyclogenesis that is distinct from a more traditional cyclogenesis process wherein convection and diabatic heating play a more central role. One may be prompted to ask - is there any evidence of a barotropic mode in eastern Pacific tropical cyclogenesis?

An adjoint model is employed to compute sensitivity of tropical cyclone intensity to perturbations of the model initial conditions. Of 24 cases of tropical cyclogenesis in the eastern Pacific between 2008-2009, at least half display structures in sensitivity to lower tropospheric winds which may indicate the potential for barotropic growth, while a third display no such structures. Storm-centered composites of sensitivity gradients and basic states for 'barotropic' and 'non-barotropic' cases are physically interpreted. 'Barotropic' cases tend to appear when the low-level jet is "thinnest" in the meridional direction and contains the most shear. Optimal perturbations to winds and temperatures for increasing TC intensity are also computed and

binned according to the possibility for barotropic growth. Finally, National Hurricane Center forecast error statistics are grouped into 'barotropic' and 'non-barotropic' bins to show that cases with the possibility for barotropic g rowth tend to be more accurately forecast at 48 hours, and a hypothesis is provided for the observed differences in predictability of 'barotropic' and 'non-barotropic' cases.

Title: An adjoint-based adaptive ensemble Kalman filter

Author(s): Hajoon Song (Scripps Institution of Oceanography, La Jolla, CA, USA), Ibrahim Hoteit (King Abdullah University of Sciences and Technology (KAUST), Thuwal, KSA), Bruce Cornuelle (Scripps Institution of Oceanography, La Jolla, CA, USA), Aneesh Subramanian (Scripps Institution of Oceanography, La Jolla, CA, USA)

This contribution presents a new hybrid EnKF (ensemble Kalman filter)/4D-VAR (four dimensional variational) approach to mitigate background covariance limitations in the EnKF. The work is based on the AEnKF (adaptive EnKF) method, which bears a strong resemblance to the hybrid EnKF/3D-VAR method. In the AEnKF, the representativeness of the EnKF ensemble is regularly enhanced with new members generated after back projection of the EnKF analysis residuals to state space using a 3D-VAR (or OI optimal interpolation) scheme with a preselected background covariance matrix. The idea here is to reformulate the transformation of the residuals as a 4D-VAR problem while constraining the new member with model dynamics and previous observations. This should provide more information for the estimation of the new member and reduce dependence of the AEnKF on the assumed stationary background covariance matrix. This is achieved by integrating the analysis residuals backward in time with the adjoint model.

Numerical experiments are performed with the Lorenz-96 model under different scenarios to test the new approach and to evaluate its performance with respect to the AEnKF and the hybrid EnKF/3D-VAR. The new method leads to the least root-squared-mean estimation errors as long as the linear assumption guaranteeing the stability of the adjoint model holds. It is also found to be less sensitive to choices of the assimilation system inputs and parameters than the other methods tested.

Title: Forecast error contribution of the global observing system using different energy norms and different representation of physical processes in the adjoint model

Authors: Marta Janiskova, Carla Cardinali (ECMWF, Shinfield Park, Reading, UK)

Over the years, a comprehensive set of the linearized physical parametrization schemes has been developed at ECMWF. These linearized schemes, operationally used in data assimilation, parametrize both the dry physical processes (vertical diffusion, gravity wave drag, shortwave and longwave radiation) and the moist processes (convection, large-scale condensation and clouds) consistently with the nonlinear model (though some simplifications are applied). In

particular, the moist processes representation in the adjoint model not only provide a better time evolution of the model state during the assimilation procedure compared to the adiabatic model or model with the dry processes only, but they also allow the assimilation of variables related to physical processes such as precipitations or clouds.

In this work, the representation of the moist physical processes in adjoint assimilation model is compared with the representation of humidity in the energy norm used to compute the forecast sensitivity to observations in the short-range forecasts. Observation forecast error contribution using the adjoint model with only dry processes (dry adjoint) but moist energy norm in the sensitivity gradient calculation is compared with error contribution obtained with moist processes (moist adjoint) and dry energy norm. Results will be presented and summarized.

Title: Comparing observation impact on low-level wind forecasts between an ensemble Kalman filter and a 3DVAR data assimilation scheme

Authors: Erin Kashawlic, Brian Ancell (Texas Tech University)

A variety of studies have been performed to determine the effectiveness of different data assimilation schemes within numerical weather prediction. For sequential schemes, previous research using mesoscale models at horizontal resolutions of tens of kilometers has shown that the ensemble Kalman filter (EnKF) has outperformed a 3DVAR system in producing both analyses and subsequent forecasts. However, the relative performance of these systems at very fine grid spacing is unclear. This study focuses on investigating the relative performance of a high-resolution EnKF and 3DVAR data assimilation scheme in producing low-level, 0-24hr wind forecasts over northwest Texas. This work employs a nested 12km/3km WRF-ARW modeling configuration and compares the Data Assimilation Research Testbed (DART) EnKF and the 3DVAR Gridpoint Statistical Interpolation (GSI) system over both domains.

Initial assimilation experiments using a variety of deployed sodar, radiosonde, and surface observations beyond the routine observational network are performed and results are presented here. The forecast quality of the EnKF and the GSI systems are compared, and the observational impacts within both assimilation systems are investigated to understand whether the most important observations vary among the two schemes. The ultimate goal of this study is to discover the best way to assimilate the most important observations in producing low-level wind and subsequent wind power forecasts. Future plans for additional assimilation experiments are discussed.

Title: Quantifying the sensitivity of nonlinear tides in the Philippine Sea

Authors: Colette Kerry, Brian Powell (Department of Oceanography, University of Hawaii at Manoa, Honolulu, Hawaii)

Adjoint sensitivity analysis is used to study the internal tides in the Philippine Sea. This work aims to investigate the predictability of the nonlinear tides and the interactions of the tidal energy and mesoscale oceanic conditions. We use an adjoint model to quantify the sensitivity of the baroclinic energy fluxes to model bathymetry, density and advective structure, as well as a selection of meaningful model parameters. This application demonstrates how the adjoint sensitivity method can simultaneously address the contribution of various background state and model variables to the baroclinic tides.

Complex mesoscale activity as well as significant internal tide energy makes the Philippine Sea a challenging oceanic region to predict. Numerical prediction of mesoscale ocean circulation typically focuses on the mesoscale and tidal forcing is omitted. However in regions with significant baroclinic tidal energy the tidal and mesoscale dynamics are often of comparable magnitude. The adjoint provides a quantitative evaluation of the relative roles of the dynamics and model configuration to the characterization of the baroclinic tides. This information helps us understand the effects of a varying mesoscale field on the internal tides and allows comparison of the importance of background oceanic state and uncertainties in bathymetric representation and model parameters. Our aim is to use this information to better predict the region.

Title: Evaluation of a hybrid ensemble-variational data assimilation scheme using an OSSE

Author(s): Daryl Kleist (NOAA/NCEP/EMC, Univ. of Maryland-College Park, IMSG), John Derber (NOAA/NCEP/EMC), David Parrish (NOAA/NCEP/EMC), Kayo Ide (Univ. of Maryland-College Park), Jeff Whitaker (NOAA/ESRL/PSD)

The capability to incorporate flow-dependent, ensemble-based representations of a background error covariance matrix into variational data assimilation has recently been developed for use in the GSI by utilizing the extended control variable method. Preliminary experimentation with minimal tuning using a coupled, dual-resolution 3DVAR-EnKF hybrid system with the NCEP Global Forecast System model have shown that utilizing the hybrid paradigm can yield substantial forecast error reduction relative to a 3DVAR-based control system. In the dual resolution experiments, the ensemble is run at a coarser resolution relative to the driving, high-resolution, deterministic (hybrid) analysis and forecast components.

To further understand the impact of hybrid data assimilation systems on the quality of analyses and subsequent forecasts, experiments will be carried out using an observing system simulation experiment (OSSE). By utilizing an OSSE and having access to a truth, it is possible to better understand how and why the supplementary ensemble information is reducing analysis and forecast error. Additionally, it is possible to investigate how well the ensemble is representing the actual analysis and forecast error covariances.

Experiments will be carried out using the NCEP GFS model and GSI-based dual-resolution hybrid by using the Joint OSSE nature run and simulated observations. In addition to hybrid 3DVAR-EnKF experiments (as have already been carried out for real observation impact evaluation), additional runs will be made to evaluate the importance of ensemble size, localization, and coupling of the EnKF to the hybrid analysis. The hybrid 3DVAR-EnKF results will be compared to results from using a GSI-based ensemble 4DVAR (without tangent linear and adjoint). The importance of time localization and supplementing the ensemble with a full-rank static estimate within the 4D paradigm will be explored.

Title: On the influence sampling of atmospheric microstates

Authors: Paul Krause (U of Santa Catarina, Mathematics, Brazil), Pedro L.S. Dias (U of Sao Paulo, Atmospheric Sciences, Brazil)

A novel path sampler for large dimensional nonlinear dynamical systems is presented. It is inspired by Chorin's "Optimal prediction with memory" on model reduction per the Mori-Zwanzig formalism of Statistical Mechanics and Kalnay's "Ensemble forecasting at NMC: the generation of perturbations" on the sampling of growing modes per state perturbations. In the Influence Sampling scheme, Monte Carlo sampling is used to draw the path statistics of a set of variables and an influence formula applied on the top of it to draw the path statistics of the remaining variables. Communication disregarded, the proposed scheme has a numerical cost comparable to Monte Carlo sampling and is parallelizable. With tests conducted on a system of ordinary differential equations obtained by non-autonomous perturbations of the Lorenz '63 equations, conclusions are drawn about the modeling and predictability issues of complex dynamical systems.

Title: Which matters more in Hybrid Ensemble 4D-VAR, variances or correlations?

Authors: David D. Kuhl (Naval Research Laboratory Washington, D.C.), Tom Rosmond (Science Application International Corp., Monterey, CA.), Craig H. Bishop (Naval Research Laboratory, Monterey, CA), Elizabeth Satterfield (Naval Research Laboratory Monterey, CA)

Recently there has been growing interest in the use of covariance models that linearly combine static and ensemble flow dependent covariances in 4D-VAR. The performance differences between the Hybrid and standard 4D-VAR can be partially attributed to a change in the forecast error variance field and partially attributed to a change in forecast error correlations. In addition, the sensitivity of Hybrid Ensemble 4D-VAR to ensemble size and/or the quality of the static covariance used in the Hybrid has not been thoroughly tested. Indeed, the theoretical results of Bishop et al’s presentation suggest that a superior static covariance model for use in a Hybrid would be one based on a very large historical collection of ensemble perturbations. Here, using NRL's recently developed observation space Hybrid Ensemble 4D-

VAR, a series of careful experiments are performed to determine the extent to which differences in performance are associated with either (a) changes in the forecast error variances or (b) changes in the forecast error correlations. Results from a low-resolution version of the system show a high degree of sensitivity to the specification of variances. By the time of the conference, we hope to be able to compare these low resolution results with higher resolution counterparts and to have also explored sensitivities to changes in ensemble size, and changing to a static covariance model based on historical ensemble perturbations.

Title: On ensemble forecasts, singular vectors and reliability

Author: Martin Leutbecher ECMWF

Perturbations of the initial conditions based on the leading singular vectors of the forecast model propagator are used to represent initial uncertainties in some ensemble prediction systems. The presentation will reflect on the rationale for this approach and the implications it has for the reliability of the ensemble forecasts, i.e. their statistical consistency. A comparison of ensemble variance with the variance of the ensemble mean error in the subspaces defined by the leading evolved singular vectors will play a key role for evaluating and understanding the representation of initial uncertainties.

Title: Variational assimilation of MODIS aersosol optical depth over east Asia region

Authors: Zhiquan Liu, Quanhua Liu (JCSDA), Hui-Chuan Lin (NCAR), and Craig Schwartz (NCAR)

Capability of assimilating NASA MODIS Aerosol Optical Depth (AOD) product is developed in NCEP GSI 3DVAR data assimilation system. Community Radiative Transfer Model (CRTM) is extended to include aerosol's absorption and scattering effects for visible and near infrared bands. This allows CRTM served as the AOD observation operator and used in data assimilation system. The corresponding Jacobian model of AOD forward operator is also developed to facilitate the gradient's calculation of the cost function as required by variational iteration procedure. In addition, analysis variables in the GSI system are expanded to include 14 aerosol species in order to initializing GOCART aerosol module, which is built in the WRF-Chem model. The newly developed AOD assimilation system is applied to dust storm case occurred in March 2010 over East Asia region. This dust storm is under-predicted by the WRF-Chem model initialized from NCEP global meteorological field. The statistics of the amplitude and length-scale of the aerosol background error is obtained. This is important for determining the relative weighting of the AOD observation and the background from a short-term forecast as well as the propagation of observation information from observed areas to unobserved areas. Assimilating MODIS AOD observation improves clearly the aerosol analysis and forecast when comparing independent AERONET AOD observations.

Title: The Met Office's hybrid ensemble-4D-Var scheme

Author(s): Adam Clayton, Dale Barker, Neill Bowler, Peter Jermey, Andrew Lorenc, Rick Rawlins, Mike Thurlow (Met Office, Exeter, UK)

The Met Office has developed a hybrid DA scheme, where perturbations from the MOGREPS ensemble are used to supplement the background error covariance matrix used in 4D-Var, and the 4D-Var analysis is used to recentre the MOGREPS ensemble. The ensemble perturbations are localised in VAR using the "alpha control variable method" applied to transformed variables so as to minimise the introduction of imbalance.

Early trials demonstrated a clear benefit of hybrid-3D-Var over 3D-Var, as found at other centres. Demonstrating a similar benefit from hybrid-4D-Var has been harder; after careful tuning of 3-dimensional localisation, and weighting of ensemble and climatological covariances, full resolution trials now demonstrate that the hybrid gives significant improvement in a range of scores. Final pre-operational trials are underway with the expectation of implemention in our operational global NWP system before the Workshop.

Title: Forecast-error-sensitivity to observations in the UM

Author(s): Richard Marriott, Andrew Lorenc (UK Met Office)

4D-Var in the UM utilises a regularised linear forecast model (and adjoint) known as the Perturbation Forecast (PF) model. Its regularised design means that forecast-error gradient calculations using the adjoint PF model are relatively insensitive to changes in the nonlinear forecast trajectory allowing efficient calculation of analysis-sensitivity vectors from single adjoint-model integrations. We show, as noted by Gelaro et al. (2007) and Tremolet (2007), that the dominant nonlinearity of this problem is the quadratic measure of forecast error and that this can be taken account of by averaging forecast-sensitivity vectors for the background and updated forecasts prior to adjoint-model integration.

Met Office 4D-Var is nonlinear and we minimise a non-quadratic cost-function. We show that the effect of these nonlinearities is not large enough such that the adjoint of Var does not represent the forward process well. We also show that, as a result of our variational quality control, the accuracy of observation-sensitivities from the UM can be improved by linearising observation operators in adjoint 4D-Var about analyses rather than about background states.

The common result that only ~51% of observations are found to be beneficial has also been observed in our system. We explain this as a result of inevitable errors in observations and in verifying analyses but also as a consequence of climatological background error covariances causing partitioning of analysis increments between error modes to be incorrect in any particular instance. This is demonstrated with examples from a toy model. I will also show a selection of real impact statistics from the Met Office's global 4D-Var and hybrid ensemble-4D-Var schemes.

Title: Variational ensemble-based forecast error variance maps filtering, a toy-models approach.

Author(s): Benjamin Menetrier, Thibaut Montmerle, Loik Berre, Yann Michel (CNRM-GAME (Meteo-France/CNRS))

A variational ensemble, based on an explicit perturbation of assimilated observations and on an implicit perturbation of the background through the cycling, allows background error covariances to be estimated (e.g. Belo Pereira and Berre, 2006). However, the high computational cost of such ensembles in operational applications restricts the ensemble size, leading to a significant sampling noise. The work of Raynaud et al. (2009) has shown that an objective spectral filtering of the variance maps can reduce the noise while keeping rich and robust information about the forecast error flow-dependency. This technique is now used operationally at Meteo-France to provide variances of "errors of the day" to the global NWP model ARPEGE. Our plan is now to adapt this approach at the convective scale for the operational cloud resolving model AROME. However, the heavy handling of such an operational code have led us to design a variational data assimilation testbed running cheap limited-area toy-models. It is specially intended to work on forecast error covariances modelling algorithms using variational ensembles. It will be shown that spectral filtering still is efficient to reduce the sampling noise of error variances, but that a wavelet filtering is better adapted to detect and preserve significant heterogeneous structures. An inflation of forecast perturbations, tuned using a posteriori diagnostics in observation space, is also investigated in order to represent model error.

Title : Estimating deformations of random processes for correlation modelling in data assimilation

Author: Dr. Yann MICHEL Meteo-France, CNRM-GAME/GMAP/RECYF, Toulouse, France

Data assimilation makes an extensive use of covariance models in order to describe the statistical structure of errors that are present in the observations, in short term forecasts, and in the numerical model. Building on work in the computer vision community, we introduce the shape from texture approach for the modelling of covariances and correlations in data assimilation with large dimensions. In this framework, the covariance model is obtained as the deformation (coordinate transform) of a stationary covariance model. Contrary to some coordinate changes already proposed in the literature, here the deformation is objectively estimated from the data. The energy of the deformed process is measured at different scales and orientations by a continuous wavelet analysis. The scalogram is shown to obey a Texture Gradient Equation that relates its derivatives to local metric changes. The deformation gradient can be estimated from a single realization of the deformed process, and integrated to recover the deformation. Estimating the inverse of the deformation allows to spatially stationarize the process. These steps define linear interpolating operators for the direct, adjoint and approximate inverse deformation, which can be used in the control variable transform in variational data assimilation schemes. We also highlight that the representation of spatially deformed background error covariances is in agreement with the propagation equation of the Kalman filter

when advection is a leading phenomenon in the model, and when the analysis error covariance matrix is more homogeneous and isotropic than the background error covariance matrix, which is the case when the observing network is dense. We show that this modeling of the correlations has interesting properties such as it does not require accurate re-normalization of the variances, it allows geographical variability of the correlations and it is computationally feasible. The algorithm is applied to the modelling of background error covariances in a convective scale model and is compared to other approaches such as the recursive filters and the diagonal assumption in a wavelet basis.

Title: Information-based localization for ensemble data assimilation

Author: Stefano Migliorini, Department of Meteorology, University of Reading

A well known limitation of ensemble filtering is the inability of providing a sufficiently accurate representation of the posterior density function in the presence of a large number of sources of observational information. This problem is usually dealt with through procedures -- e.g., Schur-product localization and local ensemble-transform Kalman filtering (LETKF) -- that limit the number of measurements that are allowed to constrain the analysis at a given location. One of the drawbacks of the Schur-product approach is that the radius of localization should be large enough not to disturb the balances that act at given spatial scales and that are well represented by the ensemble error covariance. Also, in the LETKF case, the radius of localization should be large enough to include enough observations to provide a meaningful analysis. A radius of localization that is too large may limit the ability of current localization procedures to reduce the amount of observational information, particularly over data-dense areas. Another limitation of localization procedures is in dealing with non-local observations (e.g., satellite radiances), particularly in the vertical.

In this talk we will first show results illustrating the effects of not using localization at all for a case study where the forecast ensemble is determined with an ETKF at convective scale, with and without radar data. Also, we will discuss the theoretical basis of a new localization procedure based on the information content of the measurements and show results of a number of ensemble data assimilation experiments with a two-dimensional (horizontal and vertical) linear advection model using both local and non-local observations.

Title: Information-based localization for ensemble data assimilation

Author: Stefano Migliorini, Department of Meteorology, University of Reading

A well known limitation of ensemble filtering is the inability of providing a sufficiently accurate representation of the posterior density function in the presence of a large number of sources of observational information. This problem is usually dealt with through procedures -- e.g., Schur-product localization and local ensemble-transform Kalman filtering (LETKF) -- that limit the number of measurements that are allowed to constrain the analysis at a given location. One of the drawbacks of the Schur-product approach is that the radius of localization should be large enough not to disturb the balances that act at given spatial scales and that are well represented by the ensemble error covariance. Also, in the LETKF case, the radius of localization should be large enough to include enough observations to provide a meaningful analysis. A radius of localization that is too large may limit the ability of current localization procedures to reduce the amount of observational information, particularly over data-dense areas. Another limitation of localization procedures is in dealing with non-local observations (e.g., satellite radiances), particularly in the vertical.

In this talk we will first show results illustrating the effects of not using localization at all for a case study where the forecast ensemble is determined with an ETKF at convective scale, with and without radar data. Also, we will discuss the theoretical basis of a new localization procedure based on the information content of the measurements and show results of a number of ensemble data assimilation experiments with a two-dimensional (horizontal and vertical) linear advection model using both local and non-local observations.

Title: Use of heterogeneous background error covariances accounting for precipitations at convective scale

Author(s): Thibaut Montmerle, Loik Berre (CNRM-GAME (Meteo-France/CNRS))

This study focuses on the simultaneous use of different background error covariances matrices that are representative of different meteorological phenomena in the incremental 3DVar of the French operational AROME model at convective scale.

At first, background error covariances have been computed for precipitating cases using forecasts provided by an AROME ensemble 3D-Var assimilation, coupled with the operational French global ensemble 4D-Var assimilation AEARP. Results obtained from 3h forecast differences performed separately for precipitating and non-precipitating columns will be discussed. Explicit convection and comprehensive microphysical parameterization explain in particular the large discrepancies in correlation lengths, error variances and in the coupling between humidity, temperature and divergence errors between the different areas.

These covariances are then used in different geographical locations, following a formulation that allows heterogeneity. This consists in expressing the analysis increment as the sum of two terms, one for precipitating and one for non-precipitating areas, making use of a mask deduced

from observations. The interest of considering background error covariances and balance relationships that are representative of rainy areas for the assimilation of precipitation-related observations, e.g reflectivities and radial velocities from radars, will be discussed. This method allows more localized anisotropic increments to be retrieved in a way coherent with convection processes in precipitating areas, and spin-up is mitigated. Impacts on analyses and on forecasts will be shown in a cycled assimilation framework and compared to the operational AROME suite.

Title: Estimates of analysis error, forecast error, and predictability derived from the adjoint of 4D-Var

Authors: Andy Moore (Dept of Ocean Sciences, University of California at Santa Cruz, USA), Hernan Arango (Institute of Marine and Coastal Sciences, Rutgers University, USA), Gregoire Broquet (Laboratoire des Science du Climat et de l'Environnement, France)

The adjoint of the Regional Ocean Modeling System (ROMS) 4D-Var data assimilation system has been used to compute the expected analysis errors and forecast errors in linear functions of the California Current (CC) circulation. ROMS 4D-Var has been run sequentially for the CC for a period spanning several years, and error estimates will be presented for various linear functions that characterize different aspects of the CC coastal upwelling circulation, namely upper ocean transport and heat content. Our methodology is based on the recently proposed approach of ensemble 4D-Var for estimating the analysis error covariance. However, when using the adjoint of the entire 4D-Var system, the explicit generation of an ensemble of 4D-Var analyses is circumvented. In addition, the adjoint of 4D-Var can be used to quantify the impact of each observation on the predictability of the circulation resulting from data assimilation. Given that the analysis and forecast error covariance estimates are intimately related to the prior error covariances, a series of consistency checks will also be presented which allows the efficacy of the prior covariances to be assessed.

Title: An adjoint description of geostrophic adjustment

Author: Michael C. Morgan (University of Wisconsin – Madison)

The notion of geostrophic adjustment is an important concept in geophysical fluid dynamics as it offers insight into the relationship between the slowly evolving dynamics of rapidly revolving fluid flows (i.e., low Rossby number flows) and the potential vorticity distribution (the conserved dynamical variable for rotating fluids) of those flows. Further it provides a means of addressing the scale-dependence of the evolution of mass and wind perturbations to rotating fluids, and the emergence of a nearly balanced wind and mass field. Many extant adjoint sensitivity studies have revealed that over sufficiently long integrations of adjoint models, a relationship, resembling geostrophic balance, emerges between sensitivities to mass and sensitivities to

wind. In this presentation, an analytical and numerical study of the linearized shallow water system is discussed and the conserved dynamical variables of the adjoint to the shallow water system are revealed. The scale-dependence of this "adjoint adjustment" is also explored. The implications of these results to sensitivity studies and data assimilation are discussed.

Title: Assimilation of Earth Rotation Parameters into an Atmosphere Model

Authors: Lisa Neef (GFZ German Research Centre for Geosciences), Katja Matthes (GFZ German Research Centre for Geosciences, Free University Berlin)

Observations of Earth rotation parameters (i.e. anomalies in the length-of-day and the orientation of the rotational pole) have a achieved a high degree of precision in recent years due to the space-geodetic observation techniques. By exchange of angular momentum with the solid earth, the atmosphere is a primary source of excitation of these Earth Rotation Parameters (or ERPs). Therefore, ERPs represent an integral measure of atmospheric dynamics, and a source of information that is independent of standard meteorological observations. The question is whether and how we can use this integral quantity to inform atmosphere/climate models. Because the observations are integrals of the atmospheric state, and also connected to ocean dynamics and the continental hydrosphere, the data assimilation problem presents unusual challenges. Moreover, the relative constraint offered by these observations, and the added value relative to meteorological observations, has been found to depend strongly on the timescale considered. In this talk we will discuss theses and other lessons learned in implementing the assimilation of ERPs into the atmospheric component of the Community Earth System Model (CESM) using an Ensemble Kalman Filter.

Title: Resolution of sharp fronts in the presence of model error using L1-regularized variational assimilation

Author(s): M.A. Freitag (Department of Mathematical Sciences, University of Bath), N.K. Nichols (Department of Mathematics and Statistics, University of Reading), C.J. Budd, (Department of Mathematical Sciences, University of Bath)

Abstract: Strong-constraint 4D-variational (4DVar) data assimilation methods do not perform well where there are sharp gradients, such as fronts or shocks, in the dynamics - particularly in cases where the background state contains a displacement error in the position of the front. Where model errors are present, the strong-constraint variational method also produces inaccurate analyses. In both cases the analysis is chosen to compensate on average for the background and model errors over the time window. The analysis therefore may smear the shock front and may include over/undershoots, leading to oscillations and phase errors in the forecast.

Here we introduce an alternative form of the variational problem. We show first that the 4DVar scheme can be interpreted as a form of Tikhonov, or L2 - regularization, commonly used to treat ill-posed inverse problems. The regularization term constrains the analysis to remain close to the background in the L2-norm {least-squares sense). An alternative process that has proved valuable in image processing for recovering sharp edges is L1-norm regularization. We apply this approach to the variational assimilation problem in cases where shocks are present and give examples where the L1-norm technique performs more effectively in the presence of model error than the standard L2-norm regularization.

Title: Sensitivity of the KFS to the trajectory of reference

Authors: Oger N., Pannekoucke, O., Doerenbecher, A., and Arbogast, P.

Advanced observation targeting techniques attempt to account for both model state uncertainties and properties of the data assimilation schemes.To achieve this, technical implementation exist either with ensemble-based or variational methods. Kalman Filter Sensitivity (KFS) is a data targeting technique mainly based on the adjoint technology applied to a variational data assimilation scheme. In practice, the (linear) targeting products are computed with the help of a non-linear trajectory. Through the linearisation process the targeting products depends on the non-linear trajectory. This poster provides results on the sensitivity of the KFS to the non-linear trajectory and to approximations within the data assimilation scheme.

Title: Accounting for linearisation error in the Extended Kalman Filter and 4D-Var

Author: Tim Payne, Met Office

Both the Extended Kalman Filter and incremental 4D-Var make use of a linear model, to propagate covariances explicitly in the first case and implicitly in the second. Experience with incremental 4D-Var at major operational centres suggests that the relative error in the linear model can approach 100% after 12 hours in several fields and model levels; this may be contrasted with the full model which shows significant skill even at 5 days. It would appear that linear model error can be far larger than the error in the full model. However, while much work has been done on the inclusion of full model error in data assimilation, no one has yet attempted to account for error in the linear model. This is ironic, as while full model error is largely unknown (even unknowable), we have in principle unlimited knowledge of every aspect of the errors in the linear model.

The sources of error in the linear model, on top of the error in truncating the Taylor expansion of the nonlinear model, are several. In operational DA probably the largest source of error comes from physical processes which are either not included at all, or only very approximately. This may be due to the large number of thresholds in their formulation, the complexity of the code, or their perceived cost. Another source of error is the lower resolution of the linear compared with full model, typically by a factor of between two and four.

We present an as far as possible rigorous analysis of linear model error, and how it may be accounted for in the Extended Kalman Filter and incremental 4D-Var.

Title: Model-reduced 4D-Var data assimilation in application to 1D ecosystem model

Authors: Joanna S. Pelc (Delft University of Technology, Delft, The Netherlands; Deltares, Delft, The Netherlands), Ehouarn Simon (Nansen Environmental and Remote Sensing Center, Bergen, Norway), Laurent Bertino (Nansen Environmental and Remote Sensing Center, Bergen, Norway), Ghada El Serafy (Deltares, Delft, The Netherlands), Arnold W. Heemink (Delft University of Technology, Delft, The Netherlands)

Nowadays, ecosystem models become more and more sophisticated. The number of their biological components, as well as their parameters is increasing. Even simple ecosystem models have strong nonlinear behavior. For such challenging environment data assimilation can be an extremely difficult task. Especially when one considers the adjoint-based techniques, obtaining the adjoints of such models is nontrivial. Also the model resolutions are much finer than in the past, which results in large sizes of the model states. This introduces a limitation for using the finite difference gradient approximations, since these ones are not suitable for large problems. The model-reduced 4D-Var (Vermeulen and Heemink, MWR, 2006) is a method proposed to deal with this problem. Based on a number of simulations of the original model, proper orthogonal decomposition is used to obtain a reduced model. The model-reduced 4D-Var is performed in the reduced space, therefore, the implementation of the adjoint of the tangent linear approximation of the original model is not required. Instead, it is approximated by

the adjoint of the tangent linear approximation of the reduced model. The method is easily extended to the initial condition estimation, hence the parameter calibration is coupled together with the initial condition estimation. This method was effectively used in several applications. An advantage was shown especially for models characterized by periodical behavior. Since for these type of models, the number of required model simulations is relatively small. The ecosystem models also have periodic characteristics, therefore the method is a potential tool in ecological applications. Twin experiments have been conducted in a 1D ecological model. The study highlights the ability of the method to tackle the problem of combined initial condition and parameter estimation in such nonlinear framework.

Title: High resolution dynamic data assimilation

Author(s): Ruth Petrie, Ross Bannister (University of Reading)

An important ingredient in operational data assimilation problems is the dynamic forecast error covariance matrix (P^f). This matrix describes statistically the degree of confidence in the prior state, and describes univariate and multivariate couplings. The dimensions of the P^f-matrix prohibits explicit calculation, storage or propagation with current resources. In conventional variational data assimilation this matrix is approximated by the background error covariance matrix (B). The B-matrix is typically modelled using a Control Variable Transform (CVT) and it is a static approximation to the dynamic P^f.

Leading edge forecast models for operational Numerical Weather Prediction (NWP) can operate at resolutions of 0(1 km) and at these resolutions it becomes possible to resolve small scale weather features such as thunderstorms. The transient nature and nonlinear error growth of such systems implies that a static description of their forecast error characteristics is sub-optimal.

In this work we present the Ensemble Reduced Rank Kalman Filter (EnRRKF) which is a timely extension to the RRKF of Fisher (1998). In the EnRRKF an ensemble of states, forecast by the dynamic model, are used to define a subspace. The ensemble is used to explicitly sample the forecast error statistics of that space. The EnRRKF seeks to explicitly represent the part of the state that is most "dynamically active" and merge these dynamic statistics in a formal way with the static statistics, which apply in the remainder of the space. The description of this subset of error statistics applies at the forecast time even to structures which have grown in a non-linear fashion as the ensemble trajectories are found with the non-linear forecast model. The hybrid statistics are designed to be used in a practical variational data assimilation setting. It is hoped that by incorporating a subset of flow dependent background error statistics, which are appropriate for small scale features, the analysis (and hence forecast) scores will be improved.

Reference:Fisher, M.: Development of a simplified Kalman Filter, ECMWF Technical Note 1998, 260

Title: Domain localization in ensemble based Kalman filter algorithms

Authors: T.Janjic[1], L.Nerger[1], A. Albertella[2], J.Schroeter[1], S. Skachko[3]

[1]Alfred Wegener Institute, Bussestrasse 24, D-27570 Bremerhaven, Germany

[2] Institute for Astronomical und Physical Geodesy, Muenich, Germany

[3] University of Quebec in Montreal, Montreal, Canada

Ensemble Kalman filter methods are typically used in combination with one of two localization techniques. One technique is covariance localization, or direct forecast error localization, in which the ensemble derived forecast error covariance matrix is Schur multiplied with a chosen correlation matrix. The second way of localization is by domain decomposition. Here, the assimilation is split into local domains in which the assimilation update is performed independently. Domain localization is frequently used in combination with filter algorithms that use the analysis error covariance matrix for the calculation of the gain like the ETKF and the SEIK filter. Further, domain localization methods are used with method of weighting of the observations, or localization of the observation error covariance matrix.

In this work we focus on explaining the effects of domain localization in ensemble based Kalman filter algorithms and in particular effects of weighting of observations. We introduce a new method for the localization and compare it first to the already existing methods on Lorenz 40 system. On this simple example, the method of weighting of observations is less accurate than the new method, particularly if the observation errors are small.

We apply our finding to assimilation of geodetic dynamical ocean topography (DOT) data into the global finite element ocean model (FEOM) using the local SEIK filter. The geodetic DOT was obtained by means of the geodetic approach from carefully cross-calibrated multi-mission-altimeter data and GRACE gravity fields. We show that, depending on the correlation function used for weighting, the spectral properties of the solution can be improved.

Title: Adaptive mesh method in the Met Office variational data assimilation system

Authors: Chiara Piccolo and Mike Cullen (UK Met Office)

A frequent problem in forecasting fog or icy roads in a numerical weather prediction system is attributed to the misinterpretation of the boundary layer structure in the assimilation procedure. Case studies showed that much of the misinterpretation of temperature inversions and stratocumulus layers in the assimilation is due to inappropriate background error covariances. This paper looks at the application of adaptive mesh methods in the Met Office variational assimilation system to modify the background error correlations in the boundary layer when temperature inversions or stratocumulus layers are present in the background state. Results show improvements in the analysis root mean square errors with respect to radiosonde

observations and surface observations and improvements in forecast errors in two metre temperature in the presence of low clouds. This enhancement in two metre temperature forecast is attributed to reduced background vertical correlations and increased temperature background error variances in the assimilation due to the movement of the grid near the surface. Impacts above surface will be also discussed.

Title: Errors of representivity

Author(s): J. A. Pocock, A. S. Lawless, S. L. Dance, N. K. Nichols (University of Reading)

The observation error covariance matrix is an important component of any data assimilation scheme. This error matrix consists of two parts; one contains the information on the instrument error, the other contains information on the errors of representativity. Errors of representativity are errors that arise when observations can resolve spatial scales that the model cannot or when the observation operator does not correctly model the observations. Currently these errors of representativity are not modelled correctly in data assimilation schemes. In this study we seek to understand the general structure of errors of representativity. We generate solutions of the Kuramoto Sivashinsky equation at high resolution and use this as our truth. We then create pseudo-observations of this state using different integral weighting functions and calculate the errors of representativity for observations calculated at a lower spectral resolution. We also consider observation operators that both correctly and incorrectly represent these observations. As expected we find that the errors of representativity are reduced when the observation operator correctly models the observations. We also see that the structure of the error of representativity changes when the observation weighting is altered. Finally we see that the calculated errors of representativity are sensitive to the assumed covariance of the true state.

Title: Dealing with Nonlinearities in Data-Space Assimilation of Oceanic Flows

Authors: Brian Powell (University of Hawaii), Bruce Cornuelle (Scripps Institute of Oceanography)

In state-estimation of geophysical flows, we aim to maximize the time window to improve our forecasting ability; however, as the window increases, nonlinearities within the system begin to violate the linear assumptions. We investigate data-space methods that allow us to properly account for these nonlinearities. Outer-loops implement a Picard (or fixed-point) iteration scheme, which allow a linear formulation to converge to a weakly nonlinear problem. We present a method for updating the increments in data-space to maintain proper constraint to the initial error. The Observational Error of Representativeness, R, has three sources: measurement error, nonlinear model error due to missing or incorrect physics, and linear model error due to an incorrect reference for linearization. Oftentimes, this error is prescribed without consideration to the flow characteristics; however, the third component is determined by the error in the initial estimate. If the initial error is small, the growth of nonlinear errors will be small, and the linearization error contribution to R will be smaller. Conversely, if the initial error is large,

the growth of error over the window will be much larger, which must be accounted for in R. We examine the role of nonlinearities and methods for accounting for them with both the Lorenz (1963) problem as well as to an ocean state-estimation problem in the Philippine Sea---an important component of the North Pacific Sub-Tropical gyre where Kuroshio western boundary current forms creating a highly dynamic environment with strongly nonlinear flows.

Title: ON THE POSSIBILITIES AND LIMITS OF DIRECT PHYSICAL INTERPRETATION AND SYNOPTICAL USE OF MATHEMATICAL OBJECTS RELATED TO THE ADJOINT HYDRO-THERMODYNAMICAL EQUATIONS

Author: Tamas Prager, Éva König, and Fanni Kelemen (Dept. of Meteorology, Eötvös Loránd University of Budapest)

Adjoint applications have approximately 25 years history. However the solution of the tangent linear equations, the adjoint equations and, also the solution of the equations with composite operators L*L and LL*, i.e. the right and left singular vectors appear in most applications only as pure mathematical tools, or intermediate steps of a complex chain of computations. This is the case with EPS forecasts, variational data assimilation, or forecasts using targeted observations, etc.

The main question which the presentation intended to be focused on, is, whether there is a possibility of direct physical or in other words synoptico-dynamical interpretation of these objects, which is simple enough for synopticians to use in their everyday work, or so to say in their conceptual models. In other words, is there a possibility to gain directly some additional information from these „intermediate products", which can be straightly used in everyday forecasting. The philosophy behind all this is similar to that of Hoskins, McIntytre and others in the 1980s, when they initiated the direct use of potential vorticity, Q-vector and other mathematical objects, more or less new for the synopticians at that time, in everyday synoptic work. They also recommended a methodology for use. The main difference from our trial is that their objects had absolutely clear physical meaning, while the physical meaning is not so clear in our case, even using the energy norm and scalar product.

We investigated the above in a number of case studies of extreme and near-to average weather situations and received partly positive and partly negative results. We try to give a short picture of our methodology and then present some results. In summary we try to assess the power and limits of these direct applications.

Title: Observing System Simulation Experiments (OSSEs) as tools for the investigation of data assimilation systems

Author(s): Nikki Privé, Ronald Errico, (GESTAR/GSFC)

Observing System Simulation Experiments (OSSEs) offer a unique setup for evaluation of the behavior of data assimilation systems. As the "true" state of the atmosphere is completely known in the OSSE via the Nature Run, some of the difficulties encountered when investigating the impact of observations in the real world may be circumvented. By verifying against the Nature Run, the influence of observations on the analysis can be quantified as useful or detrimental impacts. The Nature Run may be used as the initial condition for the forecast model to determine the role of model error; likewise perfect observations can be ingested into the data assimilation system to explore observation error.

Title: Accounting for model error in global and regional ensemble data assimilation systems

Author(s): Laure Raynaudm, Loik Berre Gerald Desroziers (Meteo-France),

Ensemble data assimilation (EDA) is now a widespread technique to sample analysis and background uncertainties. A variational EDA system has been running operationally at Meteo-France since july 2008 for the global Arpege model. It is used to calculate flow-dependent background-error statistics, and it also provides perturbed initial states for the global ensemble prediction system. Another important area of ongoing research at Meteo-France is the experimentation of ensemble variational assimilation with the high resolution regional model Arome (with a 2.5 km resolution).

An important aspect in both systems is the representation of model error uncertainty. An adaptive multiplicative inflation has first been investigated in the Arpege EDA. An observation-based estimate of background-error variances is extracted from diagnostics relative to the minimum of the variational cost function. This can be compared to ensemble-based variances in order to estimate the contribution of model errors to background-error variances. This model-error information is then used to implement a multiplicative inflation of background perturbations after each 6h forecast step. This approach leads to a more realistic ensemble spread and the new ensemble-based background-error statistics have a positive impact on the forecast skill. It is thus considered to implement this approach operationally in the near future. Similar techniques are being tested in the regional EDA. First results will be presented.

Title: Examining tropical cyclone predictability using a mesoscale-model adjoint

Authors: Carolyn A. Reynolds, James D. Doyle, Clark Amerault (Naval Research Laboratory, Monterey, CA)

Adjoint sensitivity calculations illustrate complex influences on tropical cyclone (TC) evolution that occur over a wide range of scales from local storm-scale phenomena to remote synoptic-scale weather systems. In this study, the adjoint and tangent linear models for the atmospheric portion of the nonhydrostatic Coupled Atmosphere/Ocean Mesoscale Prediction Sys-tem (COAMPS) are used to explore the sensitivity and predictability characteristics associated with TCs during different phases of their life cycle. COAMPS simulations are performed using a range of resolutions (15, 45, and 135 km) for both moist and dry adjoint calculations. Optimization time-scales are limited such that finite perturbation evolution remains quasi-linear, leading to shorter optimization times (12-h) for fine-scale moist calculations and longer optimization times (48-h) for larger-scale dry calculations.

We focus on Northwestern Pacific Super-typhoon Lupit, which dramatically veered to the northeast, sparing the Philippines from the landfall that was predicted by several forecast models. The characteristics of the sensitivity of the Lupit forecasts vary significantly over the lifecycle of the storm. During the intensification time period, fine-scale adjoint calculations indicate that substantial strengthening of Lupit may be achieved through warming and moistening in the lower-mid troposphere in regions that would promote a closing of the circulation, and cooling and drying outside the circulation. Small initial perturbations (less than 1 g/kg, 1 m/s and 1 K) can lead to an increase in maximum wind speed of over 50% in 12 hours. Large-scale simula-tions indicate strong sensitivity to synoptic-scale phenomena to the northwest during rapid changes in the direction of motion. During extratropical transition, the strongest sensitivity is as-sociated with a ridge to the northeast of the storm. These simulations illustrate how changes in the strength of the storm are efficiently modulated through changes to the track of the storm and the details of the extratropical transition process.

Title: Balanced Ensemble Localization with Normal Mode Initialization

Authors: Tom Rosmond (Science Application International Corp, Forks, WA), Craig Bishop (Naval Research Lab, Monterey, CA), Dave Kuhl (Naval Research Lab Washington, DC), Liz Satterfield (Naval Research Lab, Monterey, CA)

Ensemble based covariances provide flow dependence that makes them are an attractive feature for 4-D VAR data assimilation. In most applications, however, computational costs limits the size of ensembles, and the resulting covariances have non-local features that cause problems for data assimilation. A pragmatic solution to this problem is applying localization functions to the covariances that filter the non-local structures, yielding covariances that are a reasonable approximation to those possible with very large ensembles.

A troubling by-product of localization is that the wind-mass balance present in the original covariance structures is essentially lost, and this imbalance propagates to the increment fields

produced in the data assimilation system. This imbalance can compromise the effectiveness of that system.

At NRL-Monterey this problem has been solved by adding the TLM and adjoint of the NOGAPS non-linear normal mode initialization to the accelerated representer cost function upon which our ensemble DA system is based. Each conjugate gradient iteration the interim solution is 'balanced', and at convergence the increment fields are as well. There are several variations possible with the procedure, including special treatment in the stratosphere to prevent spurious inflation of increments. A mathematical discussion of the method will be presented, along with preliminary results of system performance for some of the variations.

Title: Properties of discrete adjoints for adaptive models

Authors: Adrian Sandu Virginia Tech

In this talk we discuss the properties of discrete adjoints for models that use time step and grid adaptation to control forward numerical errors. We show how the differentiation of the time step control mechanism can destroy the consistency of the discrete adjoint, and propose ways to overcome this. For adaptive grids, we show that discontinuous Galerkin provides a flexible framework for both h- and p-refinement, while preserving dual consistency. Apriori and aposteriori error estimates for the solution of inverse problems treated in a fully discrete setting are discussed.

Title: Deriving optimal weights for combining static and flow-dependent covariance models

Authors: Elizabeth Satterfield (Naval Research Laboratory, Monterey, CA), Craig H. Bishop (Naval Research Laboratory, Monterey, CA), David D. Kuhl (Naval Research Laboratory, Washington DC), Tom Rosmond (Science Application International Corp., Forks, WA)

Hybrid 4D-VAR schemes employ forecast error covariance models that are linear combinations of static covariance models and flow-dependent ensemble based covariance models. Currently, to determine the optimal weights for the flow-dependent and static parts one must perform computationally expensive trials to compare each set of plausible weights. Since it is likely that the accuracy of ensemble based error covariances varies with vertical level and latitude, it also seems reasonable to assume that the optimal weights for a Hybrid covariance model would be a slowly varying function of latitude and height. However, the amount of tuning required to optimize such spatially varying weights would be prohibitively expensive. Here, we introduce and demonstrate a new method for deriving spatially varying weights for Hybrid error covariance models. This method is based on a univariate model of the stochastic relationship between error variance predictions and true error variance (discussed in detail in Bishop et al's presentation). If the assumptions of the univariate model were true, the approach would recover optimal weights for linear combinations of ensemble variances and climatological variances from archives of innovation and ensemble variance pairs. Our first experiment evaluates the

assumptions of the univariate model applied to the Lorenz-40-variable model. The second set of experiments creates spatially varying weights for our newly developed NAVDAS-AR Hybrid by dividing the atmosphere into six regions: low-level N.Hemisphere, Tropics and S. Hemisphere and upper-level N. Hemisphere, Tropics and S. Hemisphere. The performance of the NAVDAS-AR Hybrid using these spatially varying weights is then compared against its performance using non-spatially varying weights derived by brute force tuning. At low resolution, we found that the spatially varying weights derived using the univariate theory match the performance the invariant weights obtained from tuning. If available, higher resolution results will also be presented.

Title: Application of weak constraint dual formulation 4D-Var to the California Current System

Authors: Polly Smith, Andrew Moore (Ocean Sciences Department, 1156 High Street, University of California, Santa Cruz)

Model errors are a significant source of uncertainty in analysis and forecast systems. The character, structure, and properties of these errors are generally poorly known; it is common practice in both oceanographic and meteorological data assimilation applications to neglect these errors and assume that the numerical model is perfect.

Here we describe work to attempt to quantify the magnitude and structure of model errors in the Regional Ocean Modeling System (ROMS) configured for the California Current System (CCS). It is known that ROMS estimates of sea surface temperature along the California coast are in error, particularly during the peak upwelling season. Independent analyses of the surface forcing data used to drive ROMS suggest that this tendency is related to deficiencies in the model rather than errors in the data. Previous work with the ROMS strong constraint 4D-Var system found that to compensate for this model error and produce a better fit to the observations the data assimilation acts to reduce the rate of upwelling by adjusting the alongshore surface wind stress. Clearly, this effect is undesirable; an appropriate solution is to account for the model error during the assimilation process by employing a weak constraint approach.

Practical implementation of weak constraint 4D-Var requires prescription of the prior model error covariance and offers a considerable challenge. Due to the large dimension of the problem, weak constraint 4D-Var is only possible in ROMS when working dual form. Here we will discuss issues related to application of the ROMS weak constraint dual form 4D-Var system to the CCS. We will present results from a series of experiments that aim to identify the nature of the CCS model errors and thus inform the design of suitable prior model error covariance structures.

Title: File-based model connections for data-assimilation with OpenDa

Author(s): Julius Sumihar (Deltares, Delft, The Netherlands), Martin Verlaan (Deltares & TU Delft), Stef Hummel (Deltares), Nils van Velzen (Vortech & TU Delft)

When starting a new data-assimilation application, the starting point is typically an existing model, some observations and a limited amount of time and funding to make a first assessment of the potential benefits of data-assimilation. In some applications the source code is available, but for many commercial models only the executable is available. In these cases, it is often a good solution to connect the data-assimilation methods to the model through the input files and output files of the model.

In the opensource data-assimilation toolbox OpenDA (www.openda.org), a generic approach, called black-box wrapper was developed for this purpose. The aim was to reduce the implementation time for new applications as much as possible. However, some parts of the code remain model dependent, such as the routines to read and write the input files, output files and restart files of the model. For some models these files are similar to earlier applications, so the adapter source code can be reused at least partially. The developments focussed on ensemble based techniques and model calibration, since the availability of an adjoint model is less likely in this scenario.

The use of the file-based black-box wrapper avoids several potential problems that may delay the early stages of development of a data-assimilation application. An important advantage is that the files that link the data-assimilation method to the model can be inspected to detect errors. It is also easier to use existing parallel computing capabilities of the model.

Ongoing developments try to extend the number of options for the description of uncertainty for model and observations. Although the number of options is limited at the moment, the noise properties can be specified through the OpenDA configuration files without writing additional code or compilation.

Title: 4-Dvar spectral covariance with horizontal anisotropic transformation

Author(s): Xudong Sun and Peter Steinle (Centre for Australian Weather and Climate Research)

The use of limited anisotropic horizontal covariances in spectral form in variational data assimilation will be presented. The importance of background error covariances to data assimilation are well established. Based on the previous studies of anisotropic covariances in physical space and its statistical evidence calculated in Australia region, the anisotropic horizontal covariance parameters are defined and the subsequent transform is performed from physical space to spectral space mathematically by using two-dimensional Fourier and Hankel transformations. We use SOAR correlation relationship as an example of these processes. This provides a way for correlations in 4Dvar to be calculated that not only it relies on distance but also on its horizontal directions. It is found that: with its Fourier transformation, the initial anisotropic angle can communicate in both the physical space and spectral space. Furthermore, such scheme is successfully applied in ACCESS regional 4D variational data assimilation and the model forecast verifications has shown some notable improvements in NWP model forecast accuracy.

Title: On the use of Lagrangian Coherent Structures in direct assimilation of ocean tracer images.

Author(s): Olivier Titaud (CERFACS, Toulouse, France) , Jean-Michel Brankart (LEGI, CNRS & Universit\'e de Grenoble, France), Jacques Verron (LEGI, CNRS & Universit\'e de Grenoble, France)

Satellite ocean tracer images, of Sea Surface Temperature (SST) and Ocean Colour images, for example, show patterns like fronts and filaments that characterise the flow dynamics. These patterns can be described using Lagrangian tools such as Finite-Time Lyapunov Exponents (FTLE) or Finite-Time Lyapunov Vectors (FTLV). In recent years, several studies have investigated the possibility of directly assimilating structured data from satellite images into numerical models. We exploit specific properties of FTLE and FTLV to define observation operators that can be used in a direct ocean tracer image assimilation scheme. In an idealised context, we show that high-resolution SST and Ocean Colour images can be exploited to correct velocity fields using FTLE or FTLV.

Title: Towards a longer 4D-Var assimilation window

Author: Yannick Tremolet, ECMWF

For a linear system, weak constraints 4D-Var with a long enough assimilation window is equivalent to a full rank kalman smoother. Good results have been obtained with simple systems that confirm that the analysis can be improved with a long assimilation window. However, there are still difficulties in using long assimilation widows for operational systems.

In this talk, we will present results obtained with the full ECMWF system with 24-hour and 48-hour assimilation windows. The formulation of the weak constraint term, the model error covariance matrix and the cycling with overlapping windows will be discussed.

Title: Modelling of flow-dependent ensemble-based background error correlations using a wavelet formulation.

Author(s): Hubert Varella, Loik Berre, Gerald Desroziers (Meteo-France)

A wavelet formulation on the sphere is considered for modelling heterogeneous background error correlations for the Meteo-France global Numerical Weather Prediction model. This approach is compared to the operational spectral formulation, which is horizontally homogeneous to a large extent. Moreover, while this wavelet formulation is used operationally at ECMWF to specify heterogeneous but static correlations, a flow-dependent wavelet representation of correlations is investigated in this study.

Firstly, diagnostic studies have been conducted to examine geographical variations of 3D correlations over the whole globe. The results show the ability of wavelets to represent horizontal and vertical heterogeneities of correlations. For instance, the contrast between relatively broad horizontal correlations in the tropics and sharp ones in the mid-latitudes is well represented by the wavelet formulation.

The impact of the flow-dependent formulation on the forecast quality has also been investigated during two different periods, with different levels of cyclonic activity in the North Atlantic. While the impact of the flow-dependent formulation is globally positive during the two periods, the impact over Europe tends to be more spectacular during the cyclonic period. This seems to be related to spatio-temporal variations in correlation length-scales over the North Atlantic, which are well captured by the wavelet formulation.

Title: Generic parallelization strategies for data assimilation applications.

Author(s): Nils van Velzen (Vortech & TU Delft, The Netherlands), Martin Verlaan (Deltares & TU Delft)

Parallel computing is used in many complex simulation models in order to reduce the computational time to acceptable size. The computation time is often increased by several orders when a simulation model is extended with data assimilation techniques. This increase of computational time can be reduced to an acceptable level by using parallel computing for the data assimilation algorithm.

Parallel computing can be used within the model, but also in the data assimilation algorithm. An efficient parallel computing strategy depends on various aspects such as the kind of model and the parallelization that it uses, the data assimilation algorithm and the amount of observations that are assimilated.

The open-source data assimilation toolbox OpenDA (www.openda.org) offers support for various kinds of parallel computing. It can accommodate models with various forms of parallel computing. OpenDA aims at providing a generic parallel performance that is comparable with dedicated implementations of parallel data assimilation systems for a single model taking all complexities and challenges of parallel data assimilation methods into account.

Here, the efficiency of parallel computing is illustrated with data-assimilation using Kalman type algorithms and calibration algorithms for the SWAN wave model, the WAQUA shallow water model and the LOTOS-EUROS air quality model. For most applications some degree of parallel computing can be achieved with the generic tools independent of the specific model or application.

For the attainable speedup, the type of data assimilation algorithm is very important. In general, algorithms that are very computationally efficient when run sequentially are not always the best choice for parallel computing. They may contain dependencies that can not be parallelized efficiently.

Title: Ensemble based observations sensitivity applied to storm surge forecasting

Author(s): Martin Verlaan (Deltares & TU Delft, The Netherlands), Julius Sumihar (Deltares)

Adjoint based observation sensitivity is rapidly becoming a standard tool for monitoring data-assimilation in an operational system. For ensemble based data-assimilation the tools are still in an early stage of development. In this work we apply ensemble based observation sensitivity to storm surge forecasting in the North Sea. The aim is to optimize the monitoring network and to monitor the data-assimilation process of the operational system.

In order to apply the approach of Liu&Kalnay, several modifications were needed. First, tide gage observations are frequent in time, which makes the sensitivity to the background dominant if only one observation time is considered. Asynchronous filtering (Sakov2010) of multiple times in one batch was used to avoid this problem. In addition, the forecast accuracy was measured as the misfit between observations and the matching forecasts instead of the more common approach that uses the verifying analysis as a proxy to the truth. The background for this is that in our application the analysis increments may not be beneficial in data sparse areas far from the area of interest. It also avoids the introduction of a norm for the misfit between forecast and analysis, which is not an easy task for some applications. This is also practical since the routines for computing a costfunction based on the observation to model misfits are usually available from the data-assimilation.

The application shows that observations at locations in or near the area of interest provide large improvements for small lead times. Locations further from the area of interest have a smaller impact, but provide longer lead time. There is also a dominant direction to the flow of information caused by the counter clockwise propagation of Kelvin waves for the Northern Hemisphere. Finally, observations from some locations do not improve the forecast. This may be caused by local lack of resolution near the observation location, e.g. when it is situated in an estuary.

Title: On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation

Authors: Anthony Weaver, Isabelle Mirouze (CERFACS, Toulouse)

Differential operators derived from the explicit or implicit solution of a diffusion equation are widely used for modelling background-error correlations in variational assimilation. In the isotropic case, the correlation functions implied by explicit diffusion are approximately Gaussian, whereas those implied by implicit diffusion belong to the larger class of Whittle-Matern functions which contains the Gaussian function as a limiting case. Anisotropic Gaussian and Whittle-Matern correlation functions can be constructed from their isotropic counterparts by replacing the standard normalized Euclidean distance measure with a Mahalanobis distance measure involving the inverse of an aspect tensor that accounts for directionality. Likewise, a `Daley tensor' can be defined that generalizes the standard definition of the Daley length-scale to the anisotropic case. Relationships between the aspect tensor and Daley tensor of the Gaussian and Whittle-Matern functions are established. These tensors are in turn related to the parameters of anisotropic formulations of the explicit and implicit diffusion operators. Methods to estimate the elements of the Daley tensor from an ensemble are presented and compared in idealized and real-data experiments. Since the number of independent parameters needed to specify the local Daley tensor is of the order of the total number of grid points N, sampling errors are inherently much smaller than those involved in the order N2 estimation problem of the full correlation matrix. The approach can be viewed as a hybrid technique for combining variational and ensemble methods.

Title: Evaluation of WRF 3/4D-Var and En3/4D-Var data assimilation methods for the antarctic applications

Authors: Qingnong Xiao, Chengsi Liu, Kekuan Chu (College of Marine Science, University of South Florida, St. Petersburg, FL), Bill Kuo (National Center for Atmospheric Research, Boulder, CO) , David Bromwich (Byrd Polar Research Center and Atmospheric Sciencess Program, Dept. of Geography, The Ohio State University, Columbus, OH)

Abstract: Data assimilation over Antarctica and the Southern Ocean faces exceptional challenges, among which the scarcity of traditional meteorological observations is the major one. However, the demanding for improved analysis and forecasting in the region increases after the International Polar Year (IPY, 2007-2008). Data assimilation technique becomes more important in order to extract the limited information in Antarctica for the atmospheric analysis. In this research, we examine the performance of WRF 3/4 dimensional variational (3/4D-Var) and the ensemble-based 3/4 dimensional variational (En3/4D-Var) data assimilation techniques in Antarctic applications. The technical details for the WRF En3/4D-Var are formulated and incorporated into WRF variational data assimilation system. We conducted two cases studies and evaluated the performances of these date assimilation techniques over a full month of October 2007 (in the IPY period). The observations collected for the study are conventional observations, Automatic Weather Station (AWS) data, GPS (Global Positioning System) refractivity and some satellite-retrieved products. WRF 3D-Var and En3D-Var assimilate one-

time observational data, and WRF 4D-Var and En4D-Var assimilate observations in a 6-hour window. The domain configuration is exactly the same as the Antarctica Mesoscale Prediction System (AMPS), which has a mother domain with a resolution of 45 km and a 15-km nested domain covering all of Antarctica and the parts of the Southern Ocean. Numerical experiments indicate that each data assimilationmethod has its advantage and disadvantage. Based on their performance, some conclusions and insights will be summarized in our presentation.

Title: Optimal Sensor Placement for Data Assimilations

Authors: LIANG XU (Marine Meteorology Division, Naval Research Laboratory, Monterey, CA), Wei Kang (Naval Postgraduate School, Monterey, CA, USA)

We explore the theoretical framework as well as the associated algorithms for the problem of optimally placing mobile observation platforms to maximize the improvement of forecasts. The current adjoint based observation monitor system can provide the observation impact only after the fact. A comprehensive solution to the planning of optimal sensor placement is still an open problem. The approach in this paper is based on the concept of observability, which is a quantitative measure of the information provided by sensor data and user-knowledge. The observability can be numerically computed based on the systems dynamic model. It provides the cost function for the optimization of sensor locations. The Burgers equation is used to verify this approach. In the computation, the systems observability is numerically approximated using empirical covariance matrix method. The observabiltiy is maximized using gradient projection method to find the optimal sensor location. To prove the optimality of the method, Monte Carlo simulations are carried out using standard 4DVAR algorithms based on two sets of data, one from equally spaced sensors and the other from the optimal sensor location. The results show that, relative to equally spaced sensors, the 4DVAR data assimilation achieves significantly improved estimation accuracy if the sensors are placed at the optimal location. Robustness study is also carried out in which the error covariance matrix is varied by 50% and the senor noise covariance matrix is varied by 100%. In both cases, the optimal sensor location results in improved estimation accuracy.

To conclude, the optimal sensor placement is able to significantly improve the atmospheric analysis. The concept of observability is a fundamental property of the system that does not rely on the choice of the data assimilation system to be used. For future research, optimal sensor path planning will be developed using the same concept of observability.

Title: Comparison of balance and flow-dependency of large-scale background-error variances in two ensembles

Authors: Nedjeljka Zagar (University of Ljubljana, Ljubljana, Slovenia)

Time-averaged and time-dependent structure of short-term forecast-error variances is investigated in two global ensemble data assimilation systems, the ECMWF 4D-Var ensemble and the ensemble adjustment Kalman filter DART/CAM. The applied methodology of the normal-mode function expansion provides an attractive way to measure the balance by splitting forecast-error variances into parts projecting on the balanced and inertio-gravity (IG) circulations, the approach particularly suitable for the tropics where the IG circulation dominates on all scales. The flow dependency is quantified by the flow-dependency coefficient which measures correlation between the forecast-error variances and the mean energy in two-dimensional modal subspaces. Similarities and major differences between the tyo systems over the same time period are presented and discussed. In particular, it is described how the applied inflation field in the ensemble Kalman filer assimilation has a major impact on the structure of the background variance field and its reduction by the assimilation step. A perfect-model assimilation experiment supports the findings from the real-observation experiment.

Title: Control of lateral boundary conditions in WRF 4D-Var

Author(s): Xin Zhang and Xiang-Yu Huang, (NESL, National Center for Atmospheric Research, Boulder, CO), Nils Gustafsson (Swedish Meteorological and Hydrological Institute

The regional WRF forecasting problem is a lateral boundary condition problem in addition to the initial condition problem. For regional WRF application, the lateral boundary conditions are generally provided by a global model. One may argue that the available boundary conditions from the host model have to be accepted (without change) for usage in the WRF. However, the observations close to the lateral boundaries might be assimilated in analysis stage and they might influence the initial conditions. In WRF model, the initial conditions are used as the lateral boundary conditions at the initial time, the subsequent forecast will also be influenced by these observations. With application of 4D-Var for the WRF forecasting, the lateral boundary conditions can be controlled in minimization during the period of the data assimilation window. Both ideal and realistic experiments with WRF 4D-Var show that this may be particularly important for assimilation of phenomena that are observed well inside the domain during the later part of the data assimilation window, while being propagated through the lateral boundaries during the early part of the data assimilation window. If we do not control the lateral boundary conditions, observed information related to these phenomena may be lost.