SUMMARY OF TELECONFERENCE DISCUSSION
July 29 2005, 11 am – 1 pm EST

Present: J. Whitaker (CDC), J. Anderson (NCAR), C. Bishop (NRL), M. Zupanski (CIRA/CSU), I. Szunyogh, E. Kostelich (UM), Y. Song, and Z. Toth (NCEP/EMC) (T. Hamill missed the call, M. Wei on travel)

Agenda:

1) Brief description of activities in first year - Each group described their work and main results

Jeff Whitaker
Successfully compared their ensemble data assimilation results with T62 NCEP SSI benchmark, using same or less amount of data (in fact, only 75-150k pieces of data were used, compared with ~300k in SSI – only data from +/-1 hr window used, and even that was thinned). No radiance, radar, or scatterometer data used. Results are very encouraging, 5-10% rms error reduction compared with SSI results (see his slides). Processing of remotely sensed data with same sequential algorithm is not practical, looking for alternative solutions (ETKF?) Tested 3 types of variance inflation methods, difference between successive archived analysis fields may work best, simple inflation by a coefficient almost as good

Istvan Szunyogh
Started work with simulated obs, worked well. Adapted ETKF formulation, still applied locally (region by region), very efficient algorithm. When switched to assimilation of real observations some bugs got into code, working on clearing up software. Expects some results by end of summer 2005. Looking into use of additive inflation procedures. Analyzing effect of imperfect model on DA results, using bias estimation ideas. Discussed a slide indicating that a relatively small ensemble may be able to well describe low dimensional dynamics for global circulation (PECA-type analysis).

Craig Bishop
Was unable to hire post-doc, working with Master level student, had to adjust research plans somewhat. Worked on producing large ensembles with ETKF. In parallel, work on generating ensemble perturbations to be centered around NAVDAS variational analysis, similar to M. Wei's research at NCEP: use estimate of analysis error variance derived from NAVDAS to constrain initial ensemble variance using ET algorithm. Reports that successfully used ET technique to inflate covariance: uses ET to transform old archived ensemble data tfor inflating variance in tropics of current ensemble Toth points out link between this work and that of D. Hou at EMC who plans to use similar technique to introduce stochastic perturbations. Plans to experiment with combining ensembles from different sources.

Milija Zupanski
Hired postdoc, secured accounts on NCEP IBM computers. Plans to test his method, similar to ETKF, except solves for mode (instead of mean) of distribution. Currently setting up software on NCEP machines, will start testing with simulated obs soon, in couple of mos will start using real obs. Plans using bos operators and other applicable software from NCEP SSI code. This will enable quick technology transfer to NCEP operations if research is successful.

Jeff Anderson (Unfunded collaborator)
Made an attempt to port GFS system to NCAR. Work is not complete, no funding from NOAA. Worked on generic filters, looked at sampling error in ensemble filters. Found a solution where no inflation is needed in perfect model setup. Ran some experiments, without much tuning, with NCAR T85 CAM climate model, real observations, January 2003 cases, using radiosonde and other traditional data, but no radiances. Compared results with GFS T254 operational system (p. 16 of his slides). Very encouraging results, 5-10+% rms error reduction for temperature, even larger reduction for low level wind errors. Problem with winds higher up traced to use of inaccurate obs error variances with ACAR data.

Yucheng Song
Briefly described NCEP T62 SSI benchmark analysis/forecast data set that he prepared for use by other groups (see below)

Zoltan Toth
Pointed out few links between external research and NCEP development activities: Connection between model error studies of B. Hunt (UM) and M. Pena (EMC); ET initialization by C. Bishop (NRL) and M. Wei (EMC); Inflation with ET method by C. Bishop (NRL) and D. Hou (EMC).

2) Preliminary discussion on plans for 2nd year

ZT commented that the results by JW-TH & FA are very encouraging, and warrant continuation of ensemble-based DA research work. There was general agreement on this. JW and CB discussed potential for using ensemble covariance information for improving variational schemes. They pointed out the demonstrated ability of variational schemes to process large amounts of data. JA made the point that ensemble-based DA is a new field and there is no evidence that these methods could not be modified to cope with heavy data volume, all agreed on this. MZ mentioned that after working on 4DVAR for 10 yrs, he switched to ens-DA methods because he believes they offer a theoretically more appealing approach. IS & ZT pointed out that CPU limitations on current operational machines should not constrain research aimed at 3-5 years into the future. Focus should be on understanding whether and how much improvements can be gained by using ens-DA methods compared to variational methods. Algorithms should be built with resource limitations in mind, but that should not be the primary consideration at this stage. Optimization of procedures can be considered and will become more important as the research evolves. ZT

suggested each group to continue their work under their proposal, and the project to keep focusing on ensemble-based data assimilation methods. Work on hybrid methods (where information from a set of ensemble members are used in variational DA) is encouraged but the THORPEX ens-DA funds should support the development and testing of ensemble-based DA schemes. This research has a horizon of 3-5 yrs, as compared to hybrid applications that if funded through other mechanisms may bring some benefits on shorter time scale.

3) Collaborative work within the project

ZT discussed the possibility of JW-TH, beyond their own research, playing a central role in trying to build prototype ens-DA system as time goes that would include useful and new results from any of the participating groups. This will be further discussed at next meeting.

Proposed dates/time for next meeting:
Sept 7, 1-3 pm eastern time
Sept 9, 11 am – 1 pm eastern time
Proposed agenda:

1) Review detailed plans for yr2 (each group present their plans)
2) What should be our stated goal for yr2 as a group? Like for first year, we wanted to generate a benchmark, have initial comparison; what should we aim to accomplish by end of yr2?
3) How to enhance collaboration?

The NCEP T62 Benchmark run

Yucheng Song

This document summarizes the benchmark experiment done at NCEP in preparation for the inter-comparison of different ensemble-based data assimilation schemes.

MODEL VERSION

To be comparable to the four independent groups that work on ensemble-based data assimilation (EBDA), we used the executable of global forecast model (named global_fcst6228, which is Triangular truncation T62 with 28 levels) archived at NCEP high performance storage system (HPSS). For the assimilation, we used the executable compiled on March 2 (which is named as global_ssi). For interested users who have accounts at NCEP, the source is

/hpssprod/runhistory/rh2004/200405/20040530/nwprod.tar

TEST PERIOD

From January 1 to the February 29 of 2004, the first ~15 days are also archived, thought they might be excluded from the evaluation.

EXPERIMENT PLATFORM

The benchmark experiment is finished on NCEP IBM BLUE machines.

OBSERVATIONS TO BE USED

The post quality control (post-qc) files are used for the experiment. A note here is that the data has been processed by comparing with the high-resolution gdas guess files. The input data files used are archived onto HPSS as well which is named /hpssuser/g01/wx20ys/Benchmark/dump.tar. The file also contains SST, ICE and SNOW data files used for the experiment.

ARCHIVED FILES

Every 6 hours, for the 00Z and 12Z cycle, pgb (pressure level grib) files are archived on the NCEP HPSS system. Bias correction, satellite angle, surface analysis as well as sigma analysis files are also archived. There are 31 levels in the pgb file, they are:
1000 975 950 925 900 850 800 750 700 650 600 550 500 450 400 350 300 250 200 150 100 70 50 30 20 10 7 5 3 2 1mb

HOW TO GET THE FILES

Data are archived on HPSS by day, for example, to get the data file for Feb 24, in your desired directory, you can issue command like –

hpsstar get /hpssuser/g01/wx20ys/Benchmark/20040224.tar

The command hpsstar is Mark Iredell's version of tar which is convenient to use. Users can also use htar to get the files.

QUESTIONS or COMMENTS

If after going through this entire document, you still have questions, please let me know. I can be reached at yucheng.song@noaa.gov.

SUMMARY OF TELECONFERENCE DISCUSSION
April 7, 2004, 2-3 pm EST

Present: J. Whitaker, T. Hamill (CDC), J. Anderson, J. Tribbia (NCAR), C. Bishop (NRL, joined later), M. Zupanski (CIRA/CSU), I. Szunyogh, E. Kostelich (UM), M. Wei, R. Wobus, Y. Zhu, and Z. Toth (NCEP/EMC)

General issues:

4) Funding

ZT mentioned that two participating groups will likely receive more funding than initially thought, restoring funding level to that originally requested by one group.

2) Use of real observations or simulated observations in a perfect model framework?

ZT emphasized the main focus should be on inter-comparison of different methods using real data. Several participants pointed out science advantages of carrying out perfect model data assimilation experiments as well. It was agreed that if resources permit, the project would include perfect model DA experiments as well. Rest of the discussion focused on real obs experiments, since this is the primary interest from NCEP's point of view, and this is the setup that should drive basic experimental design etc. It was noted that the addition of perfect model DA experiments would not double resources needed for the project. JA noted that NCAR could generate data based on a model integration. ZT pointed to the NCEP OSSE software that should preferably be exercised for generating simulated data (with realistic observational error). Participants (including EMC) are asked to assess whether the addition of perfect model experiments is within their reach. In case it is, details of the perfect model setup will be discussed after plans are fleshed out, and work begins with real observations.

3) Software infrastructure to be used.

ZT recalled that whenever possible, software available from NCEP (NWP model, observation operators, file format, verification routines, etc) should be used. In case new software needs to be developed in the inter-comparison project, it should be compatible with existing NCEP software and practices. These practices will insure that as the project progresses, participants can easily exchange parts of their software, can start working jointly on a prototype software, that can later be tested in an operational environment.

Experimental design:

1) Test period

After short discussion, participants agreed to use Jan-Feb 2004 as a test period. The first ~15 days will be excluded from the evaluation.

2) Observations to be used

Data types. After some discussion, participants agreed that in the main experiment, the following data types will be used:

Surface observations, radiosondes, ACAR winds, cloud drift winds. Participants can ignore some of the observation types as they wish.

Data files. The NCEP prebufr files from the final gdas analysis cycle will be used. EMC is going to make the gdas1 prepbufr files (including restricted access data, please confirm you have privileges to use that, and whether you can all use blocked data format) available on the IBM machine in a few days. There was some discussion about using CDAS data files. These files use a much longer data cut-off time that does not allow for their use in real time weather forecasting. Also, they may not contain some new data types that in further analysis we may want to consider.
In addition to the basic experiment, participating groups can also run a second experiment where they include additional data types.

Observational period. Following the usual practice (also reflected in +/-3 hrs time window for data included in prepbufr files) analyses performed at the nominal 0000 UTC time, for example, will use data up to 3 hrs after the nominal analysis time (0300 UTC). It was noted that unlike 3DVAR where through time interpolation, "future" data are used (currently up to 3 hrs into the future), ensemble-based schemes are filters that may use data only up to the time of the analysis. Therefore, if participants desire, they can choose to perform an analysis step at 0300 UTC (using data up to that valid time), for a comparison with SSI forecasts initialized at 0000 UTC (that also use data up to 0300 UTC).

Issues not discussed/settled yet:

Observational error statistics. This has not been discussed. *Suggestion* – use observational error statistics as used in operational 3DVAR, given in prepbufr files.

Quality control. This issue has not been discussed yet. *Suggestion* – use operational QC marks as given in prepbufr files. Agree about cut-off value regarding QC mark below/above which data will/will not be used.

3) Data assimilation

Cycling frequency: Each group decide on their own. *Suggestion* - required minimum analysis frequency every 6 hrs (available at 00, 06, 12, and 1800 UTC).

Next meeting, tentative: Monday, 11 April, 10 am Pacific, 11 am Mountain, 1 pm Eastern time
Continue with discussion of remaining experimental design issues in strawperson plan.

INTERCOMPARISON OF DIFFERENT ENSEMBLE-BASED DATA ASSIMILATION SCHEMES (Jan. 21, 2004)

GENERAL DESCRIPTION

BACKGROUND. Currently there are four groups that work independently with different ensemble-based data assimilation (EBDA), or ensemble generation schemes using the NCEP global assimilation/forecasting software:

Jeff Anderson and David Parrish – developed EBDA software (Ensemble Adjustment Filter, EAF) and are in the process of testing it with T62 GFS system

Mozheng Wei et al. – developed/adapted ensemble perturbation software based on EBDA concept using Ensemble Transform Kalman Filter (ETKF) algorithm, in the process of testing initial perturbation scheme. This scheme will first be tested only as an initial perturbation generation tool. Later it can be adapted and tested for data assimilation applications, too.

Istvan Szunyogh et al. – developed EBDA software (regional variant of ETKF, RETKF) using GFS model, tested in perfect model environment, in the process of setting up software to assimilate actual observations

Whitaker and Hamill – developed EBDA software (Ensemble Square Root Filter, ESRF, related to EAF), and tested it using limited surface observations

OBJECTIVE. Compare the performance of the different schemes in terms of the quality of their data assimilation and ensemble initial perturbations results. Identify the strengths and weaknesses of each scheme.

WORK SCHEDULE. The inter-comparison part of the project will be completed in 12 months, depending on available resources. Suggested starting date is Febr. 1, 2004.

Febr. 04	Work plan finalized
March-May	Prepare software
June-Oct.	Run (and if necessary, rerun) experiments
Nov.-Dec.	Verification/evaluation
Jan. 05	Short summary report, with recommendations for future work

LONG TERM BENEFITS. The most promising elements identified in the different EBDA schemes will be used to construct an EBDA scheme to be further developed and later tested in a quasi-operational environment. It is anticipated

that most or preferably all groups will contribute to the design of a scheme for future development.

PROJECT CREDO. The four groups will seek agreement on the specifics of an experimental design that each group can accept and adopt, allowing for a fair comparison of the different techniques. The four groups will agree to use a common set of values/procedures for several important aspects of the experiments. The use of a *common experimental design* will allow the identification and attribution of differences in performance that arise due to the differences between the various *EBDA schemes themselves*.

WORK PLAN. Each group will make the necessary modifications to their software and will perform the experiments on their own computers. Verification software and benchmark experiments will be provided by NCEP.

EXPERIMENTAL DESIGN

BACKGROUND. The specifics below are based on initial discussions with some project participants and are presented here only as a starting point for discussions. The four groups are expected to develop and agree on the details of the inter-comparison study.

EXPERIMENTS	*Test period:*	2-month period from most recent winter (Jan-Febr 2004?)
OBSERVATIONS	*Data set:*	NCEP operational prepbufr and satellite data files, containing data in +/-3hr chunks
	Data selection:	There are two options here. Each group uses (1) all data they can; (2) the largest set of observations that every group can assimilate (ie, model variables with height assignment)?
	Error statistics	As given in operational data files
DATA ASSIMILATION	*Output frequency:*	6 hourly?
	Model error:	Common approach (eg, multiplicative noise), or allow different treatment

by each group?

FORECASTS *Model version*: NCEP GFS, operational at start of project

Resolution: T62L28

Digital filter: Turned off?

ENSEMBLE *Size*: 50-100 members (40?)

Surface fields: Same for all members?

Frequency/length: Once per day (00Z), output every 12 hrs out to 7.5 (or 16?) days

Output format: Enspost files

Output variables: 1000 & 500 hPa geop. height, 850 hPa temp, u&v winds at 850 & 250 hPa, and precip

VERIFICATION *Data assimilation* Fit of mean of first guess forecasts (in sigma or standard pressure GRIB file) to observations, based on standard statistics used at NCEP

Initial perturbations Reposition initial ensemble members to be centered around operational SSI analysis. Run set of ensemble forecasts and evaluate them using standard NCEP probabilistic verification measures

BENCHMARK *Data assimilation*: Operational SSI analysis, run at equivalent resolution

Initial perturbations: Operational perturbations bred (at equivalent resolution