12B.5

Progress in Joint OSSEs

A new nature run and international collaboration

Michiko Masutani[1#], Erik Andersson[4], Joe Terry[2+], Oreste Reale[2$],
Juan Carlos Jusem[2$], Lars Peter Riishojgaard[2$], Tom Schlatter[5], Ad Stoffelen[7],
Jack Woollen[1+], Steve Lord[1], Zoltan Toth[1], Yucheng Song[1*], Daryl Kleist[1+],
Yuanfu Xie[5], Nikki Prive[5], Emily Liu[2+], Haibing Sun[3%],
Dave Emmitt[6], Steve Greco[6], Sid A. Wood[6],
Gert-Jan Marseille[7], Ron Errico[2$], Runhua Yang[2&], Gail McConaughy[2],
Dezso Devenyi[5], Steve Weygandt[5], Adrian Tompkins[4], Thomas Jung[4],
Valentine Anantharaj[8], Chris Hill[8], Pat Fitzpatrick[8],
Fuzhong Weng[3], Tong Zhu[3@], Sid Boukabara[3]

[1]NOAA/NWS/NCEP/EMC, Camp Springs, MD
[2]NASA/GSFC, Greenbelt, MD
[3]NOAA/NESDIS/ORA, Camp Springs, MD
[4]**European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK**
[5]**NOAA/Earth System Research Laboratory, Boulder, CO**
[6]Simpson Weather Associates (SWA), Charlottesville, VA
[7]**Royal Dutch Meteorological Institute (**KNMI), DeBilt, Netherlands
[8]Mississippi State University/GRI, MS
[#]RS Information Systems (RSIS), VA
[+]Science Applications International Corporation (SAIC)
[$]Goddard Earth Science and Technology Center, University of Maryland, Baltimore, MD
[%]QSS Group, Inc.
*I. M. Systems Group, Inc. (IMSG)
[&]Science Systems and Applications Inc (SSAI). MD
[@]Cooperative Institute for Research in the Atmosphere (CIRA)/CSU

1. INTRODUCTION

Building and maintaining observing systems (OS) with new instruments is extremely costly, particularly when satellites are involved. Objective methods that can evaluate improvement in forecast skill due to the selection of instruments and configurations have long been sought. For future instruments, the forecast skill evaluation needs to be performed using simulation experiments, known as Observing System Simulation Experiments (OSSEs). The OSSE itself is a very expensive project; however, its cost is a small fraction of the total cost of actual OS.

Corresponding author address:. Michiko Masutani, NOAA/NWS/NCEP/EMC, 5200 Auth Road Rm 207, Camp Springs, MD 20746 Michiko.masutani@noaa.gov

By running OSSEs, current operational data assimilation systems (DAS) can be upgraded to handle new data types and volume, thus accelerating the use of future instruments and OS. Additionally, OSSEs can hasten database development, data processing (including formatting) and quality control software. Recent OSSEs show that some basic tuning strategies can be developed before the actual data become available. All of this will accelerate the operational use of new OS. Through the OSSEs future OS will be designed to be effectively used by DAS and forecast systems to improve weather forecasts, thus giving the maximum societal and economic impact (Arnold and Dey 1986, Lord et al 1997, Atlas 1997, Stoffelen 2006). NCEP conducted OSSEs and demonstrated that OSSEs are able to provide critical information for assessing observational data impacts (Masutani et al. 2006). The OSSE results have often been different from theoretical explanations or speculation.

Preparation of the Nature Run (NR) has been found to be a significant amount of effort, and it is important that many OSSEs with different DAS use a common NR. Based on the recommendation of the National Oceanic and Atmospheric Administration (NOAA) and National Aeronautics and Space Administration (NASA) scientists, the European Centre for Medium range Weather Forecasting (ECMWF) produced a 13-month-long nature run for an international joint OSSE effort. Operational centers are working on the evaluation of the new NR and simulation of observations. The research community for data assimilation and designing the future observing system are able to participate in international joint OSSEs using the same nature run. By using the same nature run, simulated observations can be shared and the results can be compared. Extended international collaboration within the meteorological community is essential for timely and reliable OSSEs.

2. NATURE RUN

The NR is a long, uninterrupted forecast by a model whose statistical behavior matches that of the real atmosphere. The ideal NR would be a coupled atmosphere-ocean-cryosphere model with a fully interactive lower boundary. Meteorological science is approaching this ideal but has not yet reached it. For example, it is still customary to supply the lower boundary conditions (SST and ice cover) appropriate for the span of time being simulated.

The advantage of a long, free-running forecast is that the simulated atmospheric system evolves continuously in a dynamically consistent way. One can extract atmospheric states at any time. Because the real atmosphere is a chaotic system governed mainly by conditions at its lower boundary, it does not matter that NR diverges from the real atmosphere a few weeks after the simulation begins *provided that* the climatological statistics of the simulation match those of the real atmosphere. A NR should be a separate universe, ultimately independent from but parallel to the real atmosphere.

One of the challenges of an OSSE is to demonstrate that the NR does have the same statistical behavior as the real atmosphere in every aspect relevant to the observing system under scrutiny. For example, an OSSE for a wind-finding lidar aboard a satellite requires a NR with a realistic cloud climatology because lidars operate at wavelengths for which thick clouds are opaque. The cloud distribution thus determines the location and number of observations.

The NR is central to an OSSE. It defines the *true* atmospheric state against which forecasts using simulated observations will be evaluated. This concept deserves more explanation. In Lorenc et al. (1986) suggested a definition of "truth": the projection of the true state of the atmosphere onto the model basis. As an example, if a spectral model produces a NR, the true atmospheric state might be represented by T511 spectral coefficients on 91 levels. Atmospheric features too small to be captured by the corresponding model resolution are not incorporated in this truth.

The NR is also the source of simulated observations. For each observing system, existing or future, a set of realistic observing times and locations is developed, along with a list of observed parameters. An interpolation algorithm looks at the accumulated output of the NR, goes to the proper time and location, and then extracts the value of the observed parameter. If the NR did not explicitly provide the observed parameter, the parameter is estimated from related variables that the model does provide. Because observations extracted from the NR are equal to the defined truth (they are 'perfect'), various sources of error must also be simulated and added to form observations with realistic accuracy with respect to the NR itself.

Some have suggested that a succession of atmospheric analyses could substitute for a NR. A succession of analyses is a collection of snapshots of the real atmosphere. Though (in the case of four-dimensional variational assimilation) the analyses may each be a realizable model state, they all lie on different model trajectories. Each analysis marks a discontinuity in model trajectory, determined by the information content extracted by a DAS from the existing Global Observing System (GOS). Residual systematic effects due to the spatially non-uniform and often biased observations, DAS or model state may either favourably or unfavourably affect the potential of the new OS to improve the forecasts. Considering a succession of analyses as truth seems to be a serious compromise in the attempt to conduct a "clean" experiment. For most applications, a NR is to be preferred to a succession of analyses.

What statistical measures show that the NR sufficiently replicates the true atmosphere? The nature of the OS to be tested in the OSSE partially dictates the answer. For example, an OSSE for a satellite-borne, wind-finding lidar requires an accurate cloud climatology in the NR.

2.1 The new nature run, at higher resolution

It has become evident that the relatively low resolution of the T213NR is a limiting factor. The impact of newly developed and future satellite observing systems needs to be tested at higher model resolution, but the T213NR produced by ECMWF is too coarse to be used in OSSEs with a higher resolution assimilation model. There is also a significant drift in the tropics during the first several weeks of the T213NR. Therefore, the one-month long NR cannot be used to evaluate data impact in the tropics. Since the data impact will depend on the season, it is important that future NRs cover a long enough period, preferably a whole year. We found that the preparation of a NR and simulation of data consumes significant resources. It is desirable to have one or two good NRs and to have one or two institutes simulating the data. If the NR is accessible, the simulation of observations may be distributed. Only for OSSE calibration will interaction be required.

The NRs and simulated data ought to be shared between many institutes carrying out the actual OSSEs. OSSEs with different NRs are difficult to compare but OSSEs using different DA systems and the same NR can provide a valuable crosscheck of data impact results.

The primary specifications of the new NR are thus that it:

a. Covers a long enough period to span all seasons and to allow selection of interesting sub-periods for closer study;
b. Provides data at a temporal resolution higher than the OSSE analysis cycle;
c. Simulates the atmosphere at scales compatible with the main OS;
d. Uses daily SSTs;
e. Has user-friendly archiving.

Based on the recommendations from NOAA and NASA, ECMWF produced a new NR in July 2006 at T511 (40 km) spectral truncation with 91 vertical levels with the output saved every 3 hours. The version of the model used was the same as the interim reanalysis at ECMWF (cy30r1). The initial condition is the operational analysis on 12Z May 1st, 2005 and the NR ends at 00Z June 1st, 2006. The model was forced by daily SST and ICE provided by NCEP (used also in the operational forecasts) which is used throughout the experiments. It is planned that for one or two sub-periods (to be selected), a NR with T799 (25 Km) and 91 vertical levels will be produced and archived hourly.

The set of archived variables is enhanced to accommodate the need for OSSEs. Geopotential height at model levels has been computed and archived. This is to help in the simulation of observations based on height coordinates, such as DWL and profilers. In order to help in the evaluation of NR and OSSEs, pressure level data for 31 levels are also generated on a 1 by 1 degree regular grid. A few selected variables are also saved at the isentropic levels 315K, 330K, 350K, 370K, and 530K. In addition, for selected variables, time series of fields were generated and used to evaluate the T511NR.

At ECMWF, the T511NR is available from their MARS archive. Copies of the complete T511NR at model, including surface, pressure and isentropic levels are being saved at NCEP, NASA/GSFC and NOAA/ESRL. In the US, model level data and surface data are saved in a reduced Gaussian grid instead of spectral components. Pressure level data and potential temperature level data are saved at a 1 x 1 degree regular grid.

2.2 Diagnostics of the new NR

a) Overview

Tropical rainfall showed some dumping during the first few weeks of the T511NR, and it takes about 10 days for the convective rainfall to settle down and 20 days for the large scale rainfall. The amount of area average rainfall does not show any apparent drift in the midlatitudes.

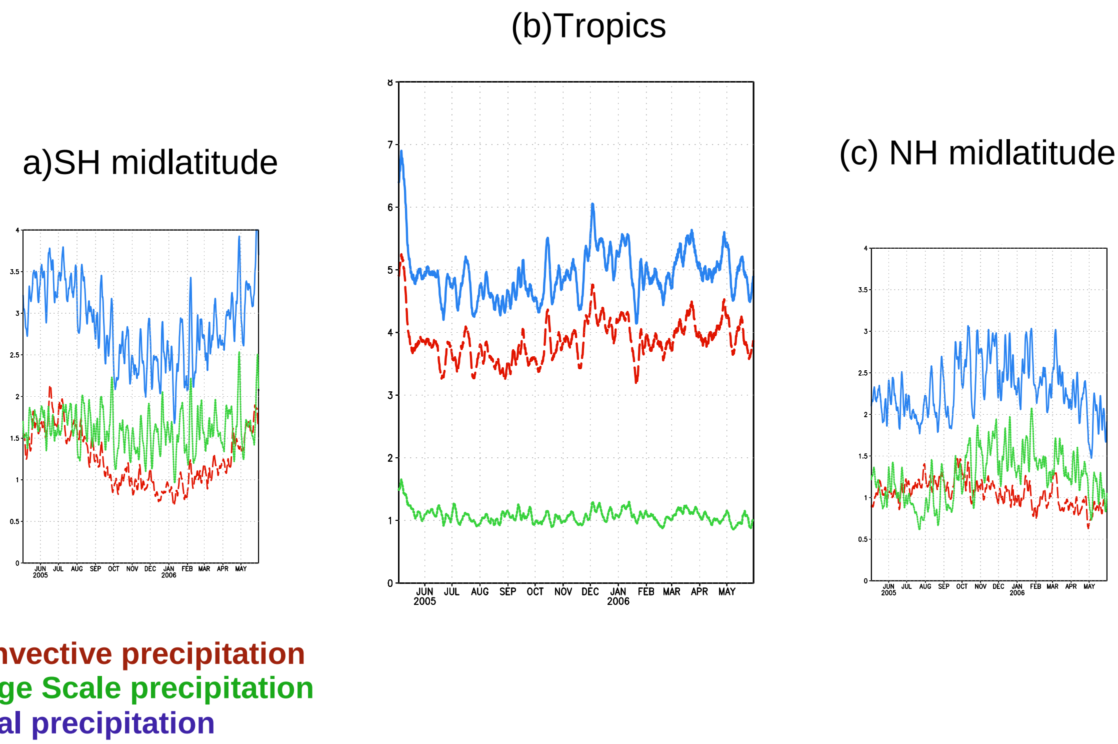

Fig.1: Time series, Zonally averaged 3 days running mean precipitation from May 01,2005 to May 31, 2006. (a) Averaged over SH midlatitude (60S-30S); (b) Averaged over tropics (10S-10N); (c) Averaged over NH midlatitude (30N-60N).

The large scale structure of the T511 NR is very realistic (Fig.2). At some times, smaller scale structures in the NR are more realistic than in the reanalysis, which is processed by a much lower resolution model.

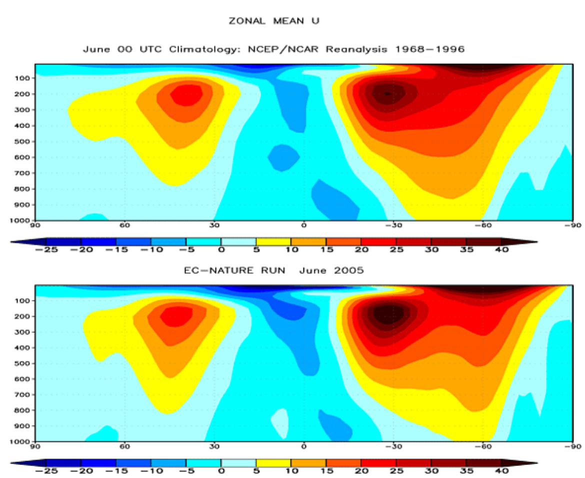

Fig.2 Zonal mean wind for June 2005. Top panel is from NCEP reanalysis. Bottom panel is computed fromT511 NR.

b) Midlatitude cyclone statistics

Midlatitude cyclone statistics were produced using Goddard's objective cyclone tracker. The cyclone tracker produces:

- Distribution of cyclone strength across the pressure spectrum;
- Cyclone lifespan (Fig. 3);
- Cyclone deepening (Fig. 4);
- Regions of cyclogenesis and cyclolysis;
- Distribution of cyclone speed and direction.

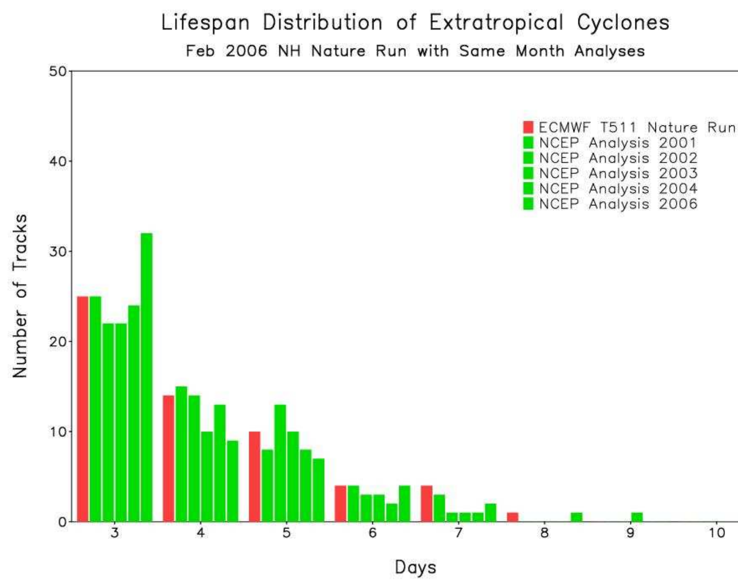

Fig. 3 . Lifespan distribution of extratropical cyclone during February 2006 in the Northern Hemisphere. Red bars are for NR. Green bars are for the NCEP analysis.

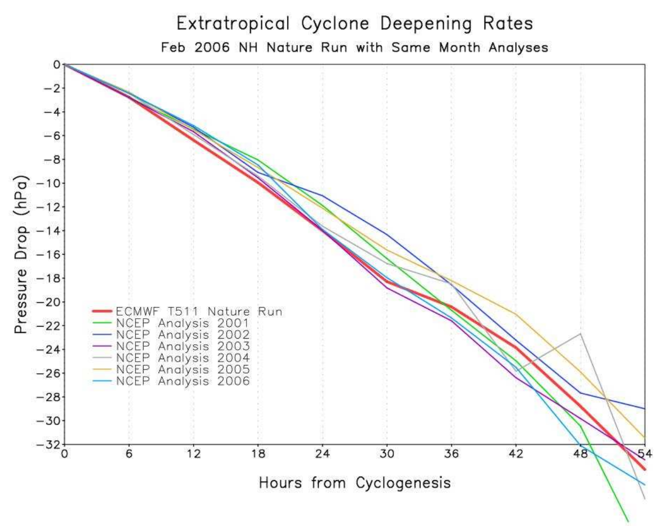

Fig. 4. Extratropical cyclone deepening rate during February 2006 in the Northern Hemisphere.

c) Tropics

Some preliminary analyses performed over the first four months of the ECMWF NR for the African Monsoon and tropical Atlantic regions are presented. The analyzed data are the 1x1 degree resolution pressure-level fields.

The overall representation of the African Easterly Jet (AEJ) is realistic, and a number of important well-known observational features are observed, such as the axis of the AEJ core slightly tilted northward and westward, a clear separation from the low-level Harmatthan flow, and a stronger low-level monsoonal flow on the western side.

Disturbances resembling African Easterly Waves are being produced in NR. The propagation speed (of about 5-9 degrees/day) and the amplitude appear realistic, as the evident modifications occurring at about transition (approximately 15W) when some waves intensify and most accelerate.

Once over the Atlantic Ocean, signs of the development and organization of some waves into smaller-scale circulations are observed. In particular, the ECMWF NR seems to also show the capability of spontaneously (without any form of vortex bogus using, relocation or ad-hoc data assimilation) producing realistic Atlantic hurricanes. One example is depicted in Fig 5, where zonal and meridional vertical cross-sections are shown relative to a vortex positioned at 29N 83W at 21z 26 Aug 2005. Despite of the interpolation at 1x1 degrees, the vortex looks very realistic, with a prominent warm core and vertically aligned isotachs indicating an eye-like feature. Low-level winds exceed 45 m/s, which is a very respectable speed for a 1x1 degree resolution data set.

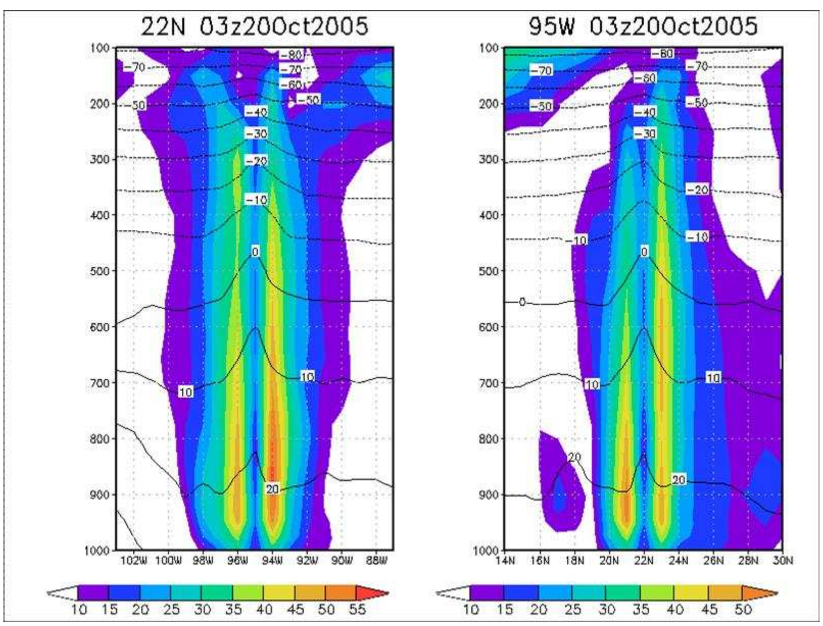

Fig 5. Zonal and meridional vertical cross-sections across the center of a Hurricane-like vortex detected at 22N 95W on 03z 20 October 2005. Plot shows wind speed (shaded, m/s, shading interval 5 m/s) and temperature (solid lines, C, contour interval 10C). Vertical structure of a HL vortex shows, even at the degraded resolution of 1 deg, a distinct eye-like feature and a very prominent warm core. Low-level wind speed exceeds 55 m/s.

These findings, albeit preliminary, are suggestive that the ECMWF NR simulates a realistic meteorology over tropical Africa and the nearby Atlantic and may prove itself beneficial to OSSE research focused over the AMMA or the Atlantic Hurricane regions.

3. Roles of OSSEs in NWP

It is a challenging task to evaluate the realism of impacts from OSSEs. Due to uncertainties in OSSE, the differences between the NR and real atmosphere, the process of simulating data, and the estimation of observational errors all affect the results. Evaluation metrics also affect the conclusion. Consistency in results is important. Some results may be optimistic and some are pessimistic. However, it is important to be able to evaluate the source of errors and uncertainties. As more information is gathered we can perform more credible OSSEs. If the results are inconsistent, the cause of the inconsistency needs to be investigated carefully. If the inconsistencies are not explained, interpreting the results becomes difficult.

NCEP's OSSEs have demonstrated that carefully conducted OSSEs are able to provide useful recommendations which influence the design of future OSs. OSSEs for DWL and for TOVS radiance showed that OSSEs can provide an evaluation of:

- whether scanning significantly improves the data, particularly in the upper atmosphere;
- the relative importance of upper or lower atmospheric data;
- the evolution of data impact with forecasts;
- the balance between model improvement and data improvement;
- the combined impacts of radiance data and wind data;
- the development of bias correction strategies.

Much research has shown that wind information has a much stronger impact on weather forecasts as compared to temperature (Arnold and Dey, 1986; Halem and Dlouhy, 1984). The results from NCEP OSSEs support these results in many ways. If DWL provides three dimensional wind data, it would cause a fundamental advance in the prediction of weather (Baker et al. 1995). Another advantage of DWL is its ability to take direct measurements of the wind, unlike extracting temperature information from radiance data, which involves radiative transfer models and many other complicated processes. Since a space based DWL is a costly instrument, a careful evaluation through OSSEs is extremely important before the investment of a large amount of resources.

As models improve, there is less improvement in the forecast due to the observations. Sometimes the improvement in forecasts due to model improvements is much more than the improvement due to observations. However, even in the NH, forecasts for sub-synoptic scales require much better observations. In the tropics, models need to be improved to retain the analysis improvement for more than a few days of forecasts.

4. Collaboration and coordination in OSSEs

Since OSSEs are a very labour-intensive project, efficient international collaborations are essential to produce timely and reliable results. Using the same NR will help in comparing the results from various DA systems and will enhance the credibility of the results. Simulation of observations requires access to the complete model level data and a large amount of resources. Each type of observed data needs to be simulated by a single institute, with the data from many institutes shared with all the OSSEs. In this way, OSSEs will be able to produce results which can be compared and this will also enhance the credibility of the results.

Ideally, all new instruments should be tested by OSSEs before they are selected to be built. OSSEs will also be important in influencing the design of the instrument and the configuration of the OS. While the instruments are being built, OSSEs will help to prepare the DA systems for the new instruments. We have to realize that developing a DA system to assimilate the simulated data is also a significant task. However, this effort has traditionally been done after the data become available. The OSSE effort demands that this same work be completed earlier and that will speed up the actual use of the new data.

OSSEs performed with various models will enhance confidence in the results. The NCEP model is known to be very different from the ECMWF model. An OSSE done using a model more similar to the ECMWF model used to generate the NR will produce different results. Although the so called fraternal twin problem

(Arnold and Dey, 1986) must be considered when interpreting the results, it is worthwhile to try to confirm the positive impact of the data. When other models are used for an OSSE, the perceived analysis quality will be reduced and be more realistic for a given set of observations. This needs to be checked by an OSSE calibration experiment.

OSSEs will be conducted by various scientists with different interests. Space-borne OSs are expensive to design, develop, deploy and operate, and typically a considerable amount of effort goes into promoting and assessing the merits of a proposed new system. Therefore, the main interest is from agencies responsible for space-based systems for civilian earth remote sensing, such as NASA and ESA.

Organizations such as NOAA/ESRL are interested in performing an OSSE whereby Unmanned Aircraft Systems (UAS) would sound the atmosphere at a regular set of points and at regular intervals of time, to augment existing climate and weather OSs with high-resolution (in the vertical) in situ observations over the oceans and polar regions. We would simulate the UAS observations necessary to perform this OSSE.

The assimilation of remotely sensed satellite data into regional modeling frameworks has been demonstrated to have helped improve atmospheric forecasts and enhance the understanding of physical processes in hurricane forecasting (Chou et al. 2007 and Zhang et al. 2007). Developing methodologies for regional OSSEs will be valuable for domain specific meteorological applications as well as in evaluating the added value of sensors onboard geostationary satellites, such as GOES-R, that are designed to provide data at high spatial and temporal resolutions.

Operational centers such as NOAA/NCEP will perform the role of finding the balance among conflicting interests to seek an actual improvement in weather predictions. Another important role of an operational center is to maintain parallel structures between OSSEs and the operational analysis.

The role of the Joint Center for Satellite and Data Assimilation (JCSDA) will be to coordinate resources among all participating institutes. Joint methodologies and software for observation simulation, including error modeling and error calibration, include:

- A common NR which is jointly maintained, evaluated and calibrated;
- A common understanding, guidelines and methodologies for designing, executing and assessing experiments;
- An integrated planning approach that ensures robustness checks of experimental results while avoiding unnecessary duplication of effort;
- A science steering group that also includes representatives from the satellite meteorology and NWP community outside of the key participants.

Actual data assimilation and evaluation of OSSE results require only simulated observed data and pressure level data. Therefore, universities with fewer computational resources will be able to participate in OSSE activities. The evaluation of OSSE results produced by other institutes is also a valuable contribution. This will help to train future scientists who are ready to work with operations.

Finally, international collaboration will become important part of OSSEs. Since real data are shared internationally, simulated data need to be shared internationally as well. If the same data are repeatedly simulated by many institutes, a major part of OSSE resources could be spent just on simulating the data.

5. Concluding remarks

OSSEs are a challenge to the scientists in the field of meteorology. Operational centers are busy in getting the best possible value out of existing instruments. Carefully designed OSSEs will enable scientists to make strong and important contributions to the decision making process for future OSs. Time will be saved in utilizing the new data compared to the work required for OSs that were built without any back up from OSSEs. However, there is a serious dilemma in spending resource on OSSEs. If you spend your resources on getting the greatest benefit out of existing data sources, you miss the opportunity to critique future OSs and then you have to live with what you get ten years later. If you spend all your resources on OSSEs, you will be criticized for letting today's valuable data fall on the floor.

OSSEs will be a great challenge to the leadership and organizations. OSSEs require strong leadership with a clear vision, because many of the efforts have to be performed against short-term benefit. Although operational systems are supposed to benefit from a development OSSE system which will reduce the cost of implementation, there are immediate costs to OSSEs.

OSSEs will require the development of good communication. It is great deal of work to implement a new element to operational suites. It is important to maintain the parallel structure between the DA system for OSSEs and the operational DA system as much as possible and keep updating and testing operational facilities in OSSEs. Many experimental facilities could be developed through OSSEs. However, these facilities cannot be used if they are not developed according to the operational coding standards. Very close coordination between the OSSE effort and operational work is very important to having a valuable outcome from OSSEs.

Finally, as the experience in OSSEs at NCEP has demonstrated, OSSEs often produce unexpected results. Theoretical prediction of the data impact and theoretical back up to the OSSE results are very important. On the other hand, unpredicted OSSE results stimulate further theoretical investigation. When all efforts come together, OSSE will help with timely reliable recommendations for future OSs. At the same time, OSSEs will prepare the operational DA system to promote the prompt and effective use of the new data.

Acknowledgments

Throughout this project many staffs at NOAA/NWS/NCEP, NASA/GSFC, NOAA/NESDIS, NOAA/ESRL and ECMWF contributed. We appreciate extensive support from the data support section and interim reanalysis project at ECMWF. Finally we appreciate the efficient technical support of M. Hart in the preparation of the manuscript.

REFERENCES

Arnold, C. P., Jr. and C. H. Dey, 1986: Observing-systems simulation experiments: Past, present, and future. *Bull. Amer., Meteor. Soc.,* **67**, 687-695.

Atlas, R. 1997:Atmospheric observation and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan,* **75**,111-130.

Baker, W.E., G.D. Emmitt, F. Robertson, R.M. Atlas, J.E. Molinari, D.A. Bowdle, J. Paegle, R.M. Hardesty, R.T. Menzies, T.N. Krishnamurti, R.A. Brown, M.J. Post, J.R. Anderson, A.C. Lorenc and J. McElroy, 1995: Lidar-measured winds from space: An essential component for weather and climate prediction. Bull. Amer. Meteor. Soc., 76, 869-888.

Chou, S-H., B. Zavodsky, G. J. Jedlovec, W. Lapenta, 2007: The Impact of Atmospheric InfraRed Sounder (AIRS) Profiles on Short-term Weather Forecasts 11th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Ocean, and Land Surfaces (IOAS-AOLS), CDROM, AMS

Halem, M. and R. Dlouhy, 1984: Observing system simulation experiments related to space-borne lidar wind profiling. Part 1 Forecast impact of highly idealized observing systems. *AMS preprint volume for the conference on Satellite Meteorology/Remote Sensing and Applications. June 25-29, 1984, Clearwater, Florida, 272-279.*

Lord, S. J., E. Kalnay, R. Daley, G. D. Emmitt, and R. Atlas 1997: Using OSSEs i.n the design of the future generation of integrated observing systems. AMS Preprint volume, 1st Symposium on Integrated Observation Systems, Long Beach, CA, 2-7 February 1997.

Lorenc, A. 1986:"Analysis Methods for Numerical Weather Prediction," *Quarterly Journal Royal Meteorological Society*, , **112**, 1177-1194)

Masutani, M., J. S. Woollen, S. J. Lord, T. J. Kleespies, G. D. Emmitt, H. Sun, S. A. Wood, S. Greco, J. Terry, K Campana, 2006: Observing System Simulation Experiments at NCEP, . *NCEP Office Note* No.451.

Stoffelen, A., G. J. Marseille, F. Bouttier, D. Vasiljevic, S. De Haan And C. Cardinali 2006: ADM-Aeolus Doppler wind lidar Observing System Simulation Experiment, *Quar.J.Roy. Metorol. Soc*. , 619, 1927-1948

Zhang, X., Q. Xiao, and T. J. Fitzpatrick, 2007: The Impact of Multisatellite Data on the Initialization and Simulation of Hurricane Lilis (2002) Rapid Weakening Phase. Mon. Wea. Rev., v135, pp. 526-548.