Lead/Author: Dr. John Derber (NWS/EMC)

5.x Atmospheric Data Assimilation

5.x.1 State of the Science

The primary goal of data assimilation is to produce the best possible initial conditions for the forecast model. Here the emphasis will be on the data assimilation in the atmospheric component of the global system. The data assimilation system will be the Hybrid system that has been under development for several years at EMC and ESRL, where the background error is defined from an ensemble system. These situation dependent background errors from the ensembles were introduced into the Grid-point Statistical Interpolation (GSI) data assimilation system used at EMC. This type of a hybrid system has been shown to be superior to either a 3-D variational system or an Ensemble Kalman Filter system used individually. The EMC operational 3-D hybrid system is expected to be upgraded to a 4-D hybrid system at the beginning of FY16. The primary foci of the atmospheric data assimilation development are shown in Table 1. Data assimilation specific to ocean/wave/ice models and land models is described in the ocean and land team plans of the overall NGGPS Implementation Plan.

Some of the data assimilation projects in Table 1 were supported by Sandy Supplement funding. They are long-term efforts that are the highest priority (only Sandy Supplemental items marked – other high priority activities may exist from other funding sources) for continuing into the future and are indicated by an asterisk (*) in Table 1.

Table 1. Data Assimilation Development Plan

Data Assimilation Development Plan
1. * The enhancement of the ensemble forecasts. This work will be done in coordination with the ensemble group, other divisions in NCEP and ESRL/PSD. Improving the ensemble forecasts will allow an improved analysis through the hybrid system. The primary work in improving the ensembles for data assimilation will be directed towards estimation of the forecast error. Much of this work will be directed towards developing appropriate representations of model uncertainty (i.e. stochastic physics parameterizations). 2. The data assimilation will continue to improve the use of current observations a. Improved specification of observation and representativeness error b. Improved quality control (including the use of ensemble information) c. Improved forward models d. * Inclusion of additional channels (especially water vapor) from currently used instruments, especially hyperspectral IR instruments e. Use of aerosol and trace gas information in radiative transfer when available

    f. Bias correction of conventional observations
    g. Correlated observation errors
    h. Intelligent use/thinning of conventional and satellite observations in time.
3. Incorporation of additional data types in the global assimilation system
    a. Inclusion of surface observations over land (can be done more appropriately with hybrid system than with earlier systems). This may include coupled assimilation of the land surface and atmospheric states.
    b. * Inclusion of new satellite instruments such as: Himawari-8, GOES-R, METOP-C, Joint Polar Satellite System (JPSS) instruments, Imager observations (including Atmospheric Motion Vectors, AMVs), COSMIC-2, Chinese and Russian satellites, GPM satellites, and others
    c. Ground, satellite and aircraft-based Radar data
    d. Ground and satellite-based Lightning data
    e. Ground based cloud observations
    f. * Aircraft humidity observations
    g. Aura MLS real time (and other) ozone profile information (delayed retrievals of q and T may also be useful for validation)
    h. Wind observations from turbine hubs from wind energy sector
    i. Others
4. * Inclusion of cloudy radiances and cloud observations such as microwave observations, IR observations, and visible observations
5. * Improved assimilation techniques
    a. 4-D ensemble-var hybrid and EVIL (variational ensemble update)
    b. Improved static background-error term
    c. Inclusion of improved variational quality control
    d. Techniques to account for displacement errors in the background
    e. Improved choices of analysis variable (especially moisture variables)
6. Improved data and assimilation monitoring - flexible diagnostic file structure
7. Bias correction of background field
8. Increased resolution consistent with forecast model
9. Reduction in model spin-up from analysis (improved balance, in both the model dynamics and the hydrologic cycle)
10. Improved minimization techniques (faster convergence)
11. Refactoring of the assimilation code to improve modularity/extensibility as well as scalability, efficiency and portability.
12. Extension of techniques developed for the global atmosphere to other scales (e.g., regional models) and variables (ocean, land, ice, etc.).

* Continuation of project initially supported by Sandy Supplement

5.x.2 Milestones, Resource Requirements, and Outcomes

The goal of data assimilation is to provide the initial state for the numerical forecast models which will produce the best possible forecast. To do this, it is necessary to improve all aspects of the current assimilation systems since any individual weakness will result in a degraded analysis. The R2O effort has funded a few external proposals to cover a portion of the necessary work. The priorities from the DA team are:

1. Upgrade and maintain use of current observations
   - **Lead Organization:** EMC
   - **Activities:** Maintenance and upgrade of current observation system (aircraft, AMV, GPS-RO) in GSI analysis.
   - **Milestones and deliverables:** + 1 year past project start: perform work to improve usage of aircraft AMV, GPS-RO and other observations. +2 years past start: deliver code for pre-operational testing and continue to improve use of aircraft, AMV, and GPS-RO observations. + 3+ years past start: continue to improve use of aircraft, AMV, GPS-RO systems. Note that higher resolution observations in both space and time pose some difficult assimilation issues.
   - **Anticipated collaborating organizations:** EMC, JCSDA, ESRL for GPS-RO
   - **Priority:** High. The bias correction, observation error, quality control and forward models for these data are inadequate. Because of the Sandy Supplemental project many additional aircraft observations and AMVs are becoming available. It is necessary to treat all of these observations carefully and completely to extract the maximum amount of information from these data and to maximize the impact from the investment in these resources.
   - **Duration:** 5 years.
   - **Points of contact:** John Derber

2. Continue to develop GSI Hybrid
   - **Lead Organization:** EMC
   - **Activities:** Continuation of hybrid system development.
   - **Milestones and deliverables:** + 1 year past project start: perform work to enhance the hybrid system. +2 years past start: deliver code for pre-operational testing and continue to improve hybrid system + 3+ years past start: continue to improve hybrid system.
   - **Anticipated collaborating organizations:** ESRL/PSD, EMC
   - **Priority:** High. The hybrid data assimilation system has resulted in significant improvements in the quality of the analysis fields and resulting forecasts. Many possible improvements are still possible on top of the initial implementation including the improved vertical localization of satellite radiance, variable weighting of the static and ensemble terms, variable localization length scales, etc. This project is to examine the potential benefits and implement those which are promising.
   - **Duration:** 5 years.
   - **Points of contact:** John Derber

3. Enhance use of cloud-impacted ("all sky") radiances
   - **Lead Organization:** EMC
   - **Activities:** Use of all-sky radiances
   - **Milestones and deliverables:** + 1 year past project start: perform work to improve usage of all sky radiances. +2 years past start: deliver code for pre-operational testing and continue

development of all-sky radiance assimilation. + 3+ years past start: continue development of all-sky radiance assimilation and operational upgrades
- **Anticipated collaborating organizations:** EMC, JCSDA, NASA GMAO
- **Priority:** High. The use of all-sky satellite radiances is still in it's infancy. Clouds and precipitation are extremely important weather variables to analyze and predict. The most complete observation of the clouds and to some extent the precipitation fields is from satellite radiances. The challenges in the use of infrared and microwave radiances will be quite different. Sufficient resources must be expended towards both microwave and infrared assimilation development.
- **Duration:** 5 years.
- **Points of contact:** John Derber

4. Improve specification of observation errors
    - **Lead Organization:** EMC
    - **Activities:** Improved specification of observation errors including correlated errors.
    - **Milestones and deliverables:** + 1 year past project start: perform work to improve specification of observation errors. +2 years past start: deliver system for pre-operational testing. + 3+ years past start: continue to refine definition of correlated errors for all observation types (including new instruments).
    - **Anticipated collaborating organizations:** EMC
    - **Priority:** High. The specification of the observational error variance determines the relative weighting of individual observations. The off-diagonal components of the error have been shown to have a significant impact on the quality of analyses and resulting forecasts. Included within this project is also the definition of the representativeness error. This project would have one person working on conventional observations, and the other working on satellite radiances. For both components, the specification of the off-diagonal components should be important, but not straightforward.
    - **Duration:** 5 years.
    - **Points of contact:** John Derber

5. Improve hydrological and dynamical balance in analysis increments
    - **Lead Organization:** EMC
    - **Activities:** Improved hydrological and dynamical balance in analysis.
    - **Milestones and deliverables:** + 1 year past project start: begin work to improve hydrological and dynamical balance in the analysis. +2 years past start: continue development. + 3+ years past start: deliver codes for first pre-operational testing and continue development.
    - **Anticipated collaborating organizations:** EMC
    - **Priority:** High. Without adequate balance between the model variables, the forecast model will lose the information within the analyses very quickly. For example, if the cloud field is not in proper balance with the wind, temperature and moisture fields, the changes to the cloud field done in the analysis will be lost quickly within the forecast process. The goal of this project is to improve this balance so that all information in the fields is retained and forecasts are improved.

- **Duration:** 5 years.
- **Points of contact:** John Derber
-

6. Improve stochastic physics in ensemble members
    - **Lead Organization:** ESRL/PSD
    - **Activities:** Improved stochastic physics in ensembles.
    - **Milestones and deliverables:** + 1 year past project start: perform work to improve specification of stochastic physics in ensembles. +2 years past start: deliver code for pre-operational testing and continue development. + 3+ years past start: continue to develop improved specification of stochastic physics in ensembles.
    - **Anticipated collaborating organizations:** ESRL/PSD, EMC
    - **Priority:** High. The ensembles within the hybrid system are extremely important in determining how the information from the observations is projected into the analysis variables. At this time, the uncertainties in the physics and dynamics are parameterized in a simple way. If these uncertainties can be better modeled, the projection of the information would improve and the resulting analyses and forecast skills enhanced. Note, that this project also would have important impacts on the ensemble prediction problem as well.
    - **Duration:** 5 years.
    - **Points of contact:** Jeff Whitaker

7. Add new satellite instruments, particularly hyperspectral IR instruments
    - **Lead Organization:** EMC
    - **Activities:** Additional instruments and channels from IR and microwave satellites.
    - **Milestones and deliverables:** + 1 year past project start: deliver code for pre-implementation testing and continue to work to incorporate additional satellite data. +2 years past start: deliver code for pre-implementation testing and continue to work to incorporate additional satellite data. + 3+ years past start: continue to deliver code for pre-implementation testing and continue to work to incorporate additional satellite data.
    - **Anticipated collaborating organizations:** JCSDA, NASA GMAO
    - **Priority:** High. The United States Government has invested and continues to invest a huge amount of resources into new satellite observations. Unfortunately, the use of these observations has not been adequately funded. Currently many channels from the hyperspectral infrared observations are not being used. With some effort, we should be able to extract some additional information from these observations. In addition, new satellites (such and Himawari, GOES-R, etc.) will continue to be launched. To make best use of the data from these new satellites careful examination of the data must be performed and improve techniques to use the information must be developed.
    - **Duration:** 5 years.
    - **Points of contact:** John Derber

8. Improve scalability and efficiency of GSI code
    - **Lead Organization:** EMC
    - **Activities:** Refactorization and optimization of GSI.

- **Milestones and deliverables:** Ongoing project which should continually improve structure of the GSI code and improve its efficiency.
- **Anticipated collaborating organizations:** JCSDA, NASA GMAO
- **Priority:** High. Changes in High Performance Computing (HPC) over the next 5-10 years necessitate improvements to GSI scalability and efficiency. At the same time, it is necessary to enhance the readability and understandability of the code for new and current developers. Code restructuring upgrades to meet all of these goals (refactorization) are essential to ensure the continued viability of the GSI analysis code. Enhancements in the forecast model and assimilation system result in demands on the computational resources that outpace the enhancement of available resources. Thus, the only way to improve take advantage of the enhancements is to optimize the efficiency of the system. In addition, the code must continually be refactored to make it (and keep it) sufficiently logical and simple for external collaborators and new developers to use. This is an ongoing project requiring continual effort as new capabilities are added and old capabilities are no longer used. **Duration:** 5 years.
- **Points of contact:** John Derber

9. Improve quality control and bias correction of conventional data

- **Lead Organization:** EMC
- **Activities:** Improved quality control and bias correction of conventional data.
- **Milestones and deliverables:** + 1 year past project start: develop automatic update of station blacklist based on station history.. +2 years past start: deliver code for automatic update of station blacklist for pre-operational testing. Begin development bias correction techniques + 3+ years past start: continue to improve quality control and bias correction of conventional data and transition to operations.
- **Anticipated collaborating organizations:** EMC
- **Priority:** High. The bias correction and quality control of observations is extremely important for the proper use of conventional data. Currently stations with a constant measurement or incorrect location may occasionally be close enough to the background to make it through the quality control. It has been seen that by looking at the history of the differences between the observations and background fields, many of the bad observations can be detected. These observations can then be removed by use of a blacklist and then reinstated when they are fixed. Also, consistent biases between the background and observations at certain locations have been noted. With bias correction, these observations can be rehabilitated. Many additional enhancements to the system can be made to ensure that the analysis system is receiving the best set of observations.
- **Duration:** 5 years.
- **Points of contact:** John Derber

10. Enhanced variational quality control for all observations

- **Lead Organization:** EMC
- **Activities:** Enhanced variational quality control in analysis.

- **Milestones and deliverables:** + 1 year past project start: complete initial work to apply enhanced variational quality control to conventional data and deliver code for pre-operational testing. +2 years past start: extend use of enhanced variational quality control to satellite data and radar data. + 3+ years past start: deliver code for pre-operational testing and continue development.
- **Anticipated collaborating organizations:** EMC
- **Priority:** High. The current variational quality control being applied in the operational analysis system must be turned on slowly within the minimization and can cause strong nonlinearities in the minimization. An enhanced variational quality control scheme has been developed by R.J. Purser at EMC. Preliminary tests of this system, shows significantly greater flexibility and minimization characteristics while applying the same quality control decision throughout the minimization. This project would be to complete this work and transition it to operations for conventional data. Then attempt to apply this technique to satellite data which often exhibit much more strongly correlated observation errors . **Duration:** 3 years.
- **Points of contact:** John Derber

11. Bias correction of background field
    - **Lead Organization:** EMC
    - **Activities:** Bias correction of background field.
    - **Milestones and deliverables:** + 1 year past project start: perform work to enhance and test background bias correction capability in GSI. +2 years past start: deliver code for pre-operational testing. + 3+ years past start: continue to improve background bias correction techniques and ensure no negative interaction with observation bias correction.
    - **Anticipated collaborating organizations:** EMC, NASA GMAO
    - **Priority:** High. Biases in the background field are not properly accounted for in the data assimilation theory normally applied in operational or research systems. To theoretically account for this bias resulting from errors in the forecast model properly, one should correct for the bias before using the background. A technique for estimating and correcting for the bias within the GSI exists and is applied for the RTMA. There is a difficulty in separating the effect of observational bias from background bias. This project will attempt to enhance the system to reduce the bias in the background field to improve the use of the information in the observations.
    - **Duration:** 5 years.
    - **Points of contact:** John Derber

12. Assimilation of aerosol and trace gas information through CRTM
    - **Lead Organization:** EMC
    - **Activities:** Inclusion of low resolution trace gas and aerosol data into high resolution assimilation system through radiative transfer model

- **Milestones and deliverables:** + 1 year past project start: perform work to use low resolution aerosol and trace gas information in high resolution analysis. +2 years past start: deliver code for pre-operational testing. + 3 years past start: operational implementation
- **Anticipated collaborating organizations:** EMC
- **Priority:** Medium. The simulation of satellite radiances are impacted by the presence of trace gases and aerosol fields. Ideally, aerosols and trace gases would be analyzed and predicted quantities in the forecast model. However, computational cost prohibit including these fields at high resolution. Operational systems are currently running at lower resolution to predict these quantities. By using the lower resolution fields in the simulation of the radiances, the larger spatial scales of the aerosol and trace gas signals can be incorporated in the simulation of the observations and therefore more information can be extracted from the radiance observations.
- **Duration:** 5 years.
- **Points of contact:** John Derber

13. Improve static background error in GSI hybrid
    - **Lead Organization:** EMC
    - **Activities:** Improved static background error in hybrid system
    - **Milestones and deliverables:** + 1 year past project start: perform work to improve specification of static background error in hybrid system. +2 years past start: deliver first set of upgrades for pre-operational testing and begin development of improved system. + 3+ years past start: continue to improve specification of static background error.
    - **Anticipated collaborating organizations:** EMC, ESRL/PSD
    - **Priority:** Medium. While much of the background error comes from the ensemble in the hybrid system, the static term still plays an important role specifying the parts of the background error that are not resolved by the ensemble. This project will attempt to improve the specification of the static term of the background error to improve the operational analysis.
    - **Duration:** 5 years.
    - **Points of contact:** John Derber

14. Increase resolution for data assimilation ensemble members
    - **Lead Organization:** EMC
    - **Activities:** Higher resolution ensembles for operational use
    - **Milestones and deliverables:** + 1 year past project start: perform work to examine the impact of higher resolution ensembles on model skill. +2 years past start: estimate trade-offs of increasing ensemble resolution in data assimilation cycle vs other enhancements (e.g., high resolution deterministic run). + 3+ years past start: repeat as needed as resolution of models increase and compute resources change.
    - **Anticipated collaborating organizations:** ESRL /PSD
    - **Priority:** Medium. Because of computational limitations, the ensembles used in the analysis are at a lower resolution than the deterministic forecast. This results in aliasing of the analysis increments into the lower resolution scales. As computer resources and the resolution and

skill of the deterministic model increase and as data assimilation techniques evolve, it will be necessary to evaluate the trade-offs between the potential ensemble resolutions and other changes in the assimilation system and deterministic forecast model. Note this project would require very significant compute resources.

- **Duration:** 5 years.
- **Points of contact:** John Derber

15. Modification of GSI for new dynamic core - Phase 3

- **Lead Organization:** EMC
- **Activities:** Work to modify GSI to work with the selected, follow-on NGGPS dynamic core.
- **Milestones and deliverables:** + 1 year from start (for low level capability: complete work.
- **Anticipated collaborating organizations:** ESRL/PSD
- **Priority:** High.
- **Duration:** 1 year. Possibly longer
- **Points of contact:** John Derber

Externally funded projects

The NGGPS Program has funded projects led by external (non-EMC) investigators that will contribute in a complementary manner to those listed above.

A. Improved tropical cyclone initialization for NCEP operations through direct assimilation of storm information

- **Lead Organization:** PI: Daryl Kleist, University of Maryland, Co-Is: Kayo Ide, Da-Lin Zhang
- **Activities:** To improve the suite of operational NWS NWP models through a new algorithm to initialize tropical cyclones by designing a consolidated tropical cyclone initialization scheme that is suitable across a wide variety of numerical guidance at NCEP, including the GFS and HWRF.

Year 1 Activities include:

- Develop observation operator for position of advisory TC position in GSI
- Validate observation operator using single observation tests
- Finalize position assimilation in EnKF (Ota work extension)
- Select period(s) for data assimilation cycling tests in coordinate with colleagues at NHC and EMC, collect relevant data, and set up experiments for NCEP GDAS/GFS
- Conduct baseline GFS tests using pseudo-operational GFS configuration including dual-resolution hybrid 3D EnVar assimilation and vortex relocation pre-processing
- Carry out experiments with GFS using the position assimilation procedure as replacement for vortex relocation using observation errors that have been established in the literature
- As results justify, deliver code and documentation
- Present results at relevant meetings and workshops
- Prepare results for publication

Year 2 activities include:

  - Perform similar sets of HWRF tests with position assimilation operator
  - Test use of higher temporal frequency advisory information in hybrid 4D EnVar for GFS
  - Compare results from high frequency observation experiment with experiment that uses more frequent updating (3 hourly instead of 6 hourly)
  - Perform estimation and calibration of observation errors
  - Explore TLNMC improvements for applications of TC initialization in the GFS
  - Explore balance related issues, especially in context of GFS which utilized TLNMC
  - Finish consolidation of TC initialization across NCEP models
  - Present results at relevant meetings and workshop
  - Prepare results for publication

- **Milestones and deliverables:** The following deliverables are anticipated:
  - A new observation operator in GSI to handle the direct assimilation of advisory TC position estimates. The software will be delivered adhering to all coding standards and testing using the EMC subversion server.
  - Results from impact tests of the new algorithm within cycled data assimilation experiments using the NCEP GFS for at least one active Atlantic hurricane season.
  - Results from impact tests of the new position assimilation algorithm for the NCEP HWRF system for at least one active Atlantic hurricane season to coincide with the same period used for the NCEP GFS.
  - Recommendations and assistance in consolidating TC initialization across all of the operational NCEP models.
  - Calibrated observation error estimates for the advisory-based observations.
  - Scripting and software related to testing and development of position assimilation in EnKF (extension of work by Ota).
  - Software modifications in GSI to make the TLNMC more applicable in the tropics.
  - Documentation and publications.
- **Anticipated collaborating organizations:** Co-Investigators: Michael Brennan (NHC), EMC - Emily Liu, Qingfu Liu, and Mingjing Tong
- **Priority:** High
- **Duration:** 05/01/2015 – 04/30/2017
- **Points of contact:** Daryl Kleist, dkleist@atmos.umd.edu

B. Testing and implementation of a cycling ensemble data assimilation system for operational hurricane prediction

- **Lead Organization**: PI: Jeffrey. S. Whitaker, ESRL/PSD, Co:I: Dr. Vijay Tallapragada, NCEP/EMC

- **Activities**: To implement a fully cycled, high resolution HWRF-based ensemble within the data assimilation system to improve the analysis at the vortex scale. The HWRF-based GSI/EnKF, currently being integrated into the operational python-based workflow, will be tested during 2015 and 2016, with the goal of operational implementation in 2017. Various configurations of the cycled EnKF configuration will be assessed with regard to forecast.
- **Milestones and deliverables**:
  - Dec 2014 - Feb 2015: Testing of the initial 6-h cycling HWRF EnKF-based assimilation system performed for the 2014 hurricane season. Performance will be evaluated relative to the 2014 operational HWRF system by comparing track and intensity forecasts initialized from the operational HWRF analyses with those initialized from analyses from the cycled HWRF EnKF-based system. Tests will be performed with and without a vortex adjustment step
  - March - June 2015: Based on the results obtained in the first series of tests, changes will be made to the system, and components will be upgraded to match a baseline HWRF configuration that will be used in multi-season pre-implementation tests for the 2015 HWRF upgrade. The 6-h cycled HWRF EnKF-based assimilation system will be run through the same series of tests as the candidate 2015 HWRF system. Currently, the HWRF 2015 system is slated to be run at 18/6/2km resolution, and will include a 40 member 6-km ‘warm start’ HWRF ensemble (initialized from GFS EnKF analyses at the previous analysis time, but not cycled within an HWRF-based EnKF system).
  - July - November 2015: Depending on the results of the multi-season tests, the cycled HWRF EnKF system will be run either in stream 1.5 or (more likely) stream 2 mode in the HFIP realtime demo.
  - November 2015 - March 2016: Tests of hourly cycling will be run, utilizing high-frequency observations. Stochastic physics parameterizations to represent model uncertainty developed at ESRL and implemented in the GFS will be ported to HWRF. An incremental analysis update (IAU) procedure to gradually introduce increments into the forecast model will be implemented in HWRF to reduce spin-down/spin-up effects.
  - March - May 2016: Based on the previous set of tests, a configuration will be chosen for stream 1.5 HFIP multi-season retrospective testing. The baseline for test will be the HWRF 2016 system.
  - July - November 2016: Stream 1.5 HFIP realtime demo.
  - November - December 2016: Configuration of an enhanced HWRF-EnKF based system for HWRF 2017 implementation. The enhanced system may include enhanced resolution, hourly cycling, stochastic physics and/or a 4D-IAU update.
- **Anticipated collaborating organizations**: CIRES, University of Colorado
- **Priority**: High
- **Duration:** 12/01/2014 – 12/01/2016

- **Points of contact**: Jeffrey. S. Whitaker, jeffrey.s.whitaker@noaa.gov

C. Development of Advanced Data Assimilation Techniques for Improved use of Satellite-Derived Atmospheric Motion Vectors

- **Lead Organization**: James Jung, CIMSS/SSEC, University of Wisconsin – Madison, Co:I: Steven A. Ackerman, CIMSS
- **Activities**: To successfully use the new GOES-R data stream to develop new data assimilation techniques to such as a nested tracking Atmospheric Motion Vector (AMV) algorithm to provide the opportunity to generate AMVs at high spatial and temporal resolution potentially increasing the positive impact of this data type on the skill of the higher resolution assimilation systems such as the HWRF.
- **Milestones and deliverables**:
    - Task 1: Clear Air Water Vapor AMV Assimilation in the Tropics - develop quality control procedures and assimilation techniques for these clear air water vapor AMVs starting with the current 3D+ENKF-VAR system and switch to the developmental 4D+ENKF-VAR system when it becomes available. The first year will focus on the development of quality control procedures for the clear air water vapor winds. The second year will focus on assimilating these AMVs in the tropics from the GOES (east and west) satellites over two seasons and assessing the impact on forecast skill. Year 1 deliverables include:
        - Completion of the quality control procedures for the clear air water vapor AMVs. May 2015 - September 2015
        - Completion of tests involving the spatial and temporal thinning and review of assimilation weights September 2015 - May 2016
        - Quarterly progress reports May 2015 - May 2016
        - Presentation to NCEP/EMC at NCWCP - May 2016
        - Presentation to International Winds Working Group - June 2016

        Year 2 deliverables include:
        - Completion of the two months during two seasons experiments May 2016 - January 2017
        - Completion of GSI software review and regression tests February 2017 - April 2017
        - Quarterly progress reports May 2016 - May 2017
        - Presentation to NCEP/EMC at NCWCP - May 2017
    - Task 2: Transition HWRF to hourly AMV assimilation and GOES-R data format – develop quality control procedures and assimilation techniques that will enable effective use of the new GOES-R like AMVs in the operational NCEP HWRF system. The first year will focus on transitioning the HWRF from using the current heritage GOES AMV

product to the new hourly GOES-R like AMVs whose spatial resolution will vary from 19km (Visible AMVs) to 76km (infrared AMVs) and whose temporal resolution is hourly. The second year will focus on transitioning to higher spatial and temporal GOES-R like AMVs. Year 1 deliverables include:

- Complete the generation of GOES-R like AMVs for the case storms May 2015 June 2015
- Complete the development of HWRF specific quality control procedures. Compute forecast impact statistics for case storms and portions of the 2014 hurricane season May 2015 - January 2016
- Complete the generation of GOES-R like AMVs during the previous three seasons for NCEP/EMC/HWRF implementation tests. July 2015 - November 2015
- Quarterly progress reports May 2015 - May 2016
- Presentation to NCEP/EMC at NCWCP - May 2016
- Presentation to International Winds Working Group - June 2016

Year 2 deliverables include:

- Generate higher spatial and temporal GOES-R like AMVs for the case storms. May 2016 - June 2016
- Complete the development of HWRF specific quality control procedures. Compute forecast impact statistics for case storms and portions of the 2015 hurricane season May 2016 - January 2017
- Generate higher spatial and temporal GOES-R like AMVs for the NCEP/EMC/HWRF implementation tests. July 2016 - Sept 2016
- Quarterly progress reports May 2016 - May 2017
- Presentation to NCEP/EMC at NCWCP - May 2017

- **Anticipated collaborating organizations**: John Derber, NWS/NCEP/EMC, Andrew Collard, NWS/NCEP/IMSG, Vijay Tallapragada, NWS/NCEP/EMC, John Le Marshall, Australian Bureau of Meteorology
- **Priority**: High
- **Duration**: 05/01/2015 - 04/30/2017
- **Points of contact**: James Jung, Jim.Jung@noaa.gov

D. Improving Global and Hurricane Prediction by Using Minimum-Cost Large Ensemble in GFS 4DEnVar Hybrid Data Assimilation System

- **Lead Organization**: Xuguang Wang, University of Oklahoma
- **Activities**: A time-lagged sampling approach is proposed to:
    - Extend the infrastructure of the GSI-based 4DEnVar system for GFS ingesting lagged forecasts;

- Investigate the optimal way to use the lagged reduced resolution GEFS ensemble in 4DEnVar;
- Investigate the optimal way to use lagged full resolution GFS forecasts in 4DEnVar and the optimal way to combine lagged forecasts at different resolutions including lagged full resolution GFS forecasts and lagged reduced resolution GEFS forecasts in 4DEnVar;
- Systematically explore the impact of using lagged forecasts to increase ensemble size in 4DEnVar on the analyses and subsequent global and hurricane predictions;
- Working with EMC to transition the development and research on lagged forecasts 4DEnVar to NCEP operational global 4DEnVar hybrid DA system if results warrant.

o **Milestones and deliverables**:
Year 1 activities:

- Select a winter period and collect operational observations.
- Setup and conduct baseline dual resolution 4DEnVar experiment with regular 80-member EnKF.
- Extend the 4DEnVar infrastructure with lagged GEFS ensemble using the equal weighting method.
- Conduct experiments using lagged ensemble 4DEnVar developed in the previous bullet and compare with the baseline 4DEnVar using the general evaluation metrics.
- Further extend 4DEnVar ingesting lagged GEFS ensemble with skill weighting method and conduct experiments to compare with the equal weighting method using general evaluation metrics.
- Compare performance of the optimized lagged ensemble 4DEnVar ingesting GEFS ensemble with 4DEnVar where the EnKF ensemble size is directly increased.
- Repeat (2) - (6) for a summer period corresponding to an active hurricane season and conduct evaluation using both general evaluation metrics and hurricane forecast evaluation metrics.
- Prepare manuscripts documenting the experiments and scientific findings

Year 2 activities:

- Using capabilities developed in year 1, further develop the multi-resolution lagged ensemble 4DEnVar where both lagged GEFS ensemble and lagged full resolution GFS forecasts will be ingested.
- Implement scale dependent weighting method.
- Conduct experiments using 4DEnVar in previous bullets for the same winter period selected in year 1.
- Conduct experiments using 4DEnVar in previous bullets for the same summer period and hurricane cases selected in year 1.

    - Explore and determine the optimal scale dependent weighting for multiresolution lagged ensemble 4DEnVar using comprehensive evaluation metrics including general evaluation metrics, scale dependent evaluation metrics and hurricane forecast evaluation metrics.
    - Deliver the lagged ensemble 4DEnVar system to EMC and assist EMC to conduct further testing for potential operational implementation.
    - Prepare manuscripts documenting the experiments and scientific findings
- **Anticipated collaborating organizations**: Jeffrey Whitaker (NOAA/ESRL/PSD), Daryl Kleist (UMD), Rahul Mahajan (NOAA/NCEP/EMC) and Yuejian Zhu (NOAA/NCEP/EMC)
- **Priority**: High
- **Duration**: 05/01/2015 - 04/30/2017
- **Points of contact**: Prof. Xuguang Wang xuguang.wang@ou.edu