Co-Leads: Paula Davidson (NWS/OSTI) and Mike Ek (NCEP/EMC)

1.0 Testbeds

1.1 State of the Science

NWS provides detailed weather forecasts out to 7 days; however, there is intense demand and interest in extending daily weather forecasts to ten days. Implementing NGGPS will provide advances in forecast skill and reliability, including improving forecast predictions for Days 6-10; guidance, post-processing and forecaster service support for weeks 3-4 forecasts; and improving capabilities for storm-scale services with optimally coupled high-resolution nests, helping forecasters improve services for decision-makers needing more consistent, specific information to help prepare for high-impact weather. The NGGPS investment in coupling atmospheric and oceanic models, improving data assimilation, and increasing resolution will result in extending the range of skillful model forecasts. For NGGPS to become the very best modeling system and provide the best products to NWS users, there is need to evaluate the new NGGPS guidance from objective and subjective perspectives, during and after the NGGPS development. Further there is need to translate the raw model guidance into consistent, actionable products and services for decision makers.

Currently, NOAA/NWS does not issue explicit week 3 to 4 forecast products, though there is great demand from a wide-variety of stakeholders for such forecasts. Forecasts for this timescale lie outside the current understanding of the limit of predictability for initial value weather forecasts, as initially documented by Ed Lorenz and subsequently verified in numerous studies. Forecasts for this timescale also lie at the intersection between possible enhancement to forecast skill due to the representation of atmosphere-ocean coupling and possible degradation in forecast skill due to biases associated with coupled feedbacks between the atmospheric and oceanic models. Finally, forecasts at this time range are generally characterized as having a small signal and large noise. Therefore, there are a large number of science questions (gaps) that need to be answered in order to ultimately produce useful forecast products for NWS stakeholders. These include the model representations of key physical/dynamical processes, e.g. troposphere-stratosphere interaction, sea-ice, data assimilation strategy atmosphere-ocean coupling strategy, statistical correction/calibration methods, ensemble strategy and extraction of information from ensemble forecasts and associated reforecasts, and evaluation of forecast value in the context of specific decision making tools. Testing activities leveraging the capabilities of the testbeds can help close these gaps. For example leveraging the capability of the Climate Testbed by first evaluating the potential skill and proposed improvements for forecasts at weeks 3-4 will achieve reliably useful operational forecast products.

Operational storm-scale numerical weather prediction within the United States is in its formative stages. One critical need to support daily forecasts of high impact weather events, including tornado outbreaks, flash floods and major aviation disruptions due to thunderstorms, is for a rigorously formulated, operationally executed convection allowing ensemble. Investment in super computing resources will enable this capacity, but success will require systematic experimentation that leverages NOAA Testbeds, such as the Hazardous Weather Testbed, to enable collaboration between model developers, specialized forecast centers, and the research community. Design elements include the selection and refinement of dynamic cores, scale and phenomena appropriate physics, and ensemble perturbation strategy, with appropriate coupling to, and leveraging of, global scale predictive systems. A key first step will be to refine and test forecast evaluations (verification) systems for measuring success on convective scales, including establishing robust metrics to support effective modeling system improvement cycles and to measure overall project success. Establishing a storm-scale ensemble prediction system is essential to support United States goals for a Weather Ready Nation, and to enable next-generation severe storm warning capacities referred to as Warn-on-Forecast.

NOAA testbeds and proving grounds (NOAA TBPG), listed under www.testbeds.noaa.gov facilitate the orderly transition of research capabilities to operational implementation through development testing in testbeds, and pre-deployment testing and operational readiness/suitability evaluation in operational proving grounds. TBPG have consistent guidelines for governance, function and execution, and have traditionally conducted independent testing, with a recently increased focus on collaborative testing. Coordinated and collaborative testing activities leveraging the collective capabilities and capacity of NOAA TPBG are needed to support the effective, efficient transition of NGGPS into operations. Testing projects for components and coupled system prototypes of the NGGPS, and testing/evaluation of forecast service impacts will involve a diverse range of subject matter experts, forecasters and stakeholders.

1.2 Objectives

The overall objective to engage the TBPG is to tap into their capability and capacity to test and evaluate the developmental progress and the operational readiness of NGGPS. Many of the facilities have at least some base support and infrastructure (e.g. IT workstations) and some have programmatically sponsored testing and supercomputing resources (e.g. DTC). Leveraging coordinated and collaborative testing activities will be key to the success and smooth transition of the model and coupled components.

Testbeds and proving grounds will focus on generalized phases of transition to operations as illustrated in Table X below. Developmental testing will be followed by experimental testing to demonstrate the project components' readiness to progress to the operations phase, and ultimately deployment into NWS operations.

Table X. Phased Implementation into NWS Operations

Phase	Key Q	Key Metric	Facility
Research & Development RL* 1-3	Does it work?	Peer-reviewed publication	Universities, government labs, private industry
Developmental Testing RL* 4-5	Does it work with operational systems?	Feasibility/ Engineering analysis successful	Testbed with operations-like environment
Experimental Testing RL* 6-8	Does it meet operational performance criteria?	Go/No Go based on: • Objective performance (e.g. accuracy) • Subjective feedback • Production readiness	Operational proving ground for clinical tests and full "dress rehearsal"
Operational RL* 9	Does it maintain required performance?	Objective criteria: • Accuracy • Reliability	Operations

*RL, Readiness levels for operations, are defined in the Appendix.

Working with the NGGPS team, test criteria will be established in three general categories, for both developmental and experimental testing:

1. Objective performance (e.g. specific NWP skill metrics), reliability, run-time, latency, etc.
2. Subjective performance (e.g. utility to forecasters, workforce/workflow impacts, user feedback (forecasters, external partners including decision-makers)
3. Production readiness (e.g. engineering criteria including H/W S/W protocols, back-up and data retention capabilities, ongoing near-real time verification)

1.3 Milestones and Outcomes for Near-Term Objectives

Short term objectives are to establish performance targets, refine and enhance testing infrastructure and testing approach. Initial generalized performance criteria and targets for service improvements have been incorporated in the testbeds priorities of the Round I NGGPS funding opportunity, for conducting developmental testing and evaluation, and experimental testing and evaluation at NOAA TBPG. Initial projects include additional specific test metrics in proposed test plans, which if found robust through the course of investigations, would be finalized in consultation with the appropriate TBPG and NGGPS managers.

Short-term objectives to be achieved during first two years (by end of FY16) support the overarching objective to advance forecaster tools and applications for NGGPS:

Establishing Verification Methods and Benchmark Skill:

- High-level testing criteria and performance targets for testing, identifying any critical gaps in metrics were developed and incorporated in the Round 1 FFO and initial implementation plan.
- Review and selection of proposed testbeds projects from Round I for support, along with enhanced support for testbed infrastructure/personnel for robust involvement of NOAA forecast expertise.
- Refinement and enhancement of the testing infrastructure and testing approach by adjusting the basic structure shown in Table XX. Further development of the evolved role of the DTC is discussed elsewhere in the plan.
- Completing (9) initial testing projects funded under Round I FFO associated with developing advanced forecaster tools and applications for NGGPS via establishing verification methods and benchmark skill assessments. See section 5.12.4 for associated activities, milestones and deliverables on these projects (completed by April, 2017).
- High-level testing criteria and performance targets for testing and assessing critical gaps in services developed and incorporated in the Round 2 FFO (Q1, FY 2016).
- Review and selection of proposed testbeds projects from Round II for support, along with enhanced support for testbed infrastructure/personnel for robust involvement of NOAA forecast expertise. Focus on filling gaps from Round I projects on identifying critical metrics and establishing service targets for forecasts of high-impact and severe/extreme weather. (Q2, FY2016)
- Sustain support for testing infrastructure in conjunction with funded NGGPS projects (Q2 2016).
- Launch second cycle of test projects from Round II FFO (Q4 2016).
- Preliminary recommendations for forecast elements (e.g. impact-based thresholds) and associated skill targets for days 8-10 (e.g. winter weather and quantitative precipitation); and for weeks 3-4 (e.g. drought, heat).
- Incorporating results from Round 1 projects, provide preliminary recommendations to improve consistency of forecasts for 0-6 hours and days 1-3 (severe convection, aviation weather, fire weather) from analyses of baseline skill using 13km resolution GFS

Table 1. TBPG Testing Infrastructure and Coordination for NGGPS Testing

Potential contributions of NOAA TBPG in R2O coupled model initiative:

1. CODE testing:

Global components — DTC (1, new role) — potentially TBA(2):
COMT,JCSDA,SWPT,CTB,GRPG
collaborative NWP testing approach(3)

2. Regional & Storm-scale components:

HWT ⟷ AWT (?on storm-scale convection?)

Forecaster and service impacts (including Post-processing and forecaster apps):

Service centers — JHT HMT/WPC HWT AWT

↕ ↕ ↕ ↕

Local offices — OPG (for WFOs, RFCs)

Notes:

1. DTC in partnership with OAR & NCEP serves as/ becomes testing environment for NGGPS- generating tests output for model components, and (eventually) integrated system. Interface for external R&D partners to test selected potential component upgrades. NCEP is partner or lead for additional O2R functions, e.g. maintaining code repository for current operational codes, supporting user requests for downloads, helpdesk etc.
2. COMT, JCSDA, SWPT, CTB, GRPG potential roles would grow in Yrs2-5, with additional resources for testing special model components: coastal/ocean components (COMT); week 2-4 testing in conjunction with IS/IA timeframes (CTB); and high-altitude extension (SWPT). Additional work on testing expanded satellite data assimilation during years 2-5 would be expected, depending on whether analyses (of early developmental testing of the NGGPS) show critical observational gaps.
3. Collaborative NWP testing could involve sharing (real-time) predictions generated in DTC for global coupled to additional regional or storm-scale components

1.3a Milestones: Programmatic

- High-level testing criteria and performance targets for testing, identifying any critical gaps in metrics were developed and submitted to the R2O Initiative lead. Q4 FY14. Completed
    - Outcome: Used in Round 1 FFO and project selection.
- Review and selection of proposed testbeds projects from Round I for support, along with enhanced support for testbed infrastructure/personnel for robust involvement of NOAA forecast expertise. Q2 FY15. Completed
    - Outcome: Staffing enhancements ongoing.
    - Requests for enhanced outyear support launched.
- Refinement and enhancement of the testing infrastructure and testing approach by adjusting the basic structure shown in Table 2. Further development of the evolved role of the DTC is discussed elsewhere in the plan. Q4 FY15. In progress
    - Outcome: Charter for evolved DTC to support NGGPS (in work).

- Launching initial testing projects (9) associated with service improvements evaluations at NOAA testbeds and proving grounds. Q2 FY15. Completed launch: May 1, 2015.
    - Outcome: Leveraged expertise of wider community in partnership with NWS science services teams in developing advanced forecaster tools and applications for NGGPS.
- Completing annual announcement of opportunity for 2nd cycle of test projects, launch 2nd-cycle test projects. Q2 FY16 (Est)
    - Outcomes: Broader engagement of NOAA scientists and their partners in NGGPS (anticipated). Robust testing and service targets for NGGPS for high-resolution, high-impact forecasts (days 0-3), and extended high-impact forecast information for days 6-10 and weeks 3-4.

1.3b Milestones and Outcomes: Funded projects

Milestones and outcomes are presented for each of the nine projects supported in Round I. Projects are grouped below under three generalized priority areas. Each project will test specific aspects of service improvements at identified NWS operational forecast units, and in addition each will contribute to achieving outcomes for multiple additional priorities.

Advanced Forecaster Tools and Applications for NGGPS: Establishing Verification Methods and Benchmark Skill

i. Global-scale and extended-range weather applications

a. Exploitation of Ensemble Prediction System Information in support of Atlantic Tropical Cyclogenesis Prediction
b. Application of a Hybrid Dynamical-Statistical Model for Week 3-4 Forecast of Atlantic/Pacific Tropical Storm and Hurricane Activities
c. An Investigation of the Skill of Week Two Extreme Temperature and Precipitation Forecasts at NCEP-WPC
d. Validation of Significant Weather Features and Processes in Operational Models Using a Cyclone Relative Approach

ii. Storm-Scale and High-Resolution Applications

a. Test and Evaluation of Rapid Post-Processing and Information Extraction from Large Convection 3hr Tornado Outlooks
b. Data Mining of High-resolution Storm-scale Data Sets
c. Information Extraction and Verification of Numerical Weather Prediction for Severe Weather Forecasting
d. Improvement of Convective/Severe Weather Prediction through an Integrative Analysis of WRF Simulations and NEXRAD/GOES Observations over the CONUS

iii. Cross-cutting

a. Incorporation of near real-time Suomi NPP Green Vegetation Fraction and Land Surface Temp data into the NCEP Land modeling suite

Establishing Skill Benchmarks and Verification Protocols for High Impact Weather Applications

i.a. Exploitation of Ensemble Prediction System Information in support of Atlantic Tropical Cyclogenesis Prediction

- **Lead testbed:** Joint Hurricane Testbed. PI: Christopher Thorncroft, SUNY Albany, CO-I: Jason Dunion, University of Miami/RSMAS/CIMAS
- **Activities:**
  - Developing dynamical ensemble prediction tools, based on GFS and ECMWF ensemble prediction systems for tropical cyclogenesis associated with African Easterly Waves (AEWs)
  - Extending the JHT-funded objective Tropical Cyclone Genesis Index (TCGI) to include probabilistic information provided by the GFS and ECMWF ensemble prediction systems.
- **Milestones and deliverables:**
  - Year 1
    - (05/01/2015−11/1/2015) Provide quantitative assessment of model skill in prediction of tropical cyclogenesis associated with AEWs, as a function of region and lead time and for both the GFS and ECMWF ensemble prediction systems.
    - (08/01/2015−4/30/2016) Provide quantitative assessment of the key factors influencing predicted tropical cyclogenesis associated with AEWs, as a function of region and lead time and in both the GFS and ECMWF ensemble prediction systems.
    - (2/1/2016-4/20/2016) Finish initial assessment of the physical reasons for good and bad forecasts of tropical cyclogenesis associated with AEWs, as a function of region and lead time and in both the GFS ad ECMWF prediction systems.
  - Year 2
    - (05/01/2016-10/31/2016) Establishment of objective tool for predicting probable forecast skill of tropical cyclogenesis associated with AEWs, as a function of region and lead time and in both the GFS and ECMWF ensemble prediction systems.
    - (05/01/2016-01/31/2017) Establishment of forecast tools and related graphics to provide probabilistic guidance of tropical cyclogenesis probability to NHC forecasters.
    - (11/01/2016-04/30/2017) Establishment of real-time website to communicate forecast tools and related graphics created in preceding milestone.
    - (08/01/2016-01/31/2017) Establish a methodology for incorporating ensemble information into the TCGI product.
    - (11/01/2016-04/30/2017) Establish a beta ensemble based TCGI product to be run in real-time both on pre-invests and invests.

- **Anticipated collaborating organizations:** NCEP/NHC, GFDL
- **Priority (Low, medium, high):** High
- **Duration:** 05/01/2015-04/30/2017
- **Points of contact:** Christopher Thorncroft, SUNY Albany , CO-I: Jason Dunion, University of Miami/RSMAS/CIMAS

i.b Application of a Hybrid Dynamical-Statistical Model for Week 3 to 4 Forecast of Atlantic/Pacific Tropical Storm and Hurricane Activities

- **Lead organization:** Climate Testbed. PI: Jae-Kyung E. Schemm, NOAA/CPC, CO-I: Hui Wang, Innovim
- **Activities:** Develop a dynamical–statistical prediction system for week 3 and week 4 tropical storm and hurricane activities in the tropical North Atlantic and eastern and western tropical North Pacific and implement the model for useful operational forecasts over the two ocean basins.
  - Explore and better understand the impact of the MJO cycle on the sub-monthly variability of tropical storms and hurricanes in the tropical North Atlantic, eastern and western tropical North Pacific regions, respectively, and to assess implications for week 3 to 4 predictions.
  - Develop a hybrid dynamical–statistical model for week 3 to 4 tropical storm and hurricane forecasts with the multiple linear regression method and cross-validate the model over the 1999- 2015 period.
  - To test the model for real-time week 3 to 4 forecasts for the 2016 hurricane season and implement the model into operations at NCEP/CPC starting from 2017 hurricane season.
- **Milestones and deliverables:**
  - Year 1
    - (05/01/2015−10/31/2015) Statistical analysis of the NOAA Hurricane Best Track Dataset and the daily CFSv2 Reanalysis data over the 1999-2014 period, and the 45-day CFSv2 reforecasts (1999−2010) and the 45-day CFSv2 real-time forecasts (2011−2014).
    - (11/01/2015−04/30/2016) Development of the dynamical−statistical forecast model for the week 3 to 4 forecasts of tropical storms and hurricanes including the cross-validation for the 1999−2015 period.
  - Year 2
    - (05/01/2015-10/31/2016) The real-time forecast will be tested for the 2016 hurricane season
    - (11/01/2016-04/30/2017) The model will be implemented for operational forecasts at NCEP Climate Prediction Center. The procedures will include transferring finalized model codes to the CPC's computer farm, as well as the scripts for extracting the 45- day CFSv2 real-time forecast data, running the forecast model, and post-processing. Prepare the documentation for the model, write up a manuscript summarizing the project, including the model development and validation, and submit it to a journal for publication.
- **Anticipated collaborating organizations:** CPC
- **Priority (Low, medium, high):** High

- **Duration:** 05/01/2015-04/30/2017
- **Points of contact:** Dr. Jae-Kyung Schemm, Jae.Schemm@noaa.gov

i.c. An Investigation of the Skill of Week Two Extreme Temperature and Precipitation Forecasts at the NCEPWPC

- **Lead testbed:** Hydromet Testbed. PI: Lance Bosart, SUNY Albany , CO-I: Daniel Keyser, SUNY Albany
- **Activities:**
  - Identify cases of extreme weather events (EWEs) for the CONUS on the basis of extreme precipitation and/or temperature
  - Analyze identified EWEs for associated flow regimes and regime transitions that can lead to formation of high-amplitude Rossby waves on timescales of 5-10 days
  - Develop ensemble-based probabilistic operational forecast procedures for predicting likelihood of EWEs over the CONUS for the 8-10 day time period
- **Milestones and deliverables:**
  - Year 1 (05/01/2015−4/30/2016)
    - Begin U Albany research effort on EWEs over the CONUS.
    - Spin up collaborative R2O effort with NCEP-WPC.
    - Present preliminary project results at conferences and workshops.
  - Year 2 (05/01/2016−4/30/2017)
    - Continue UAlbany research effort on EWEs over the CONUS.
    - Conclude components of UAlbany research effort ready for transfer to operations.
    - Continue collaborative R2O effort with NCEP-WPC.
    - Transition testing at HMT completed of candidate forecast guidance for 8-10 day period
- **Anticipated collaborating organizations:** NWS/WPC, Hydromet Testbed
- **Priority (Low, medium, high):** High
- **Duration:** 05/01/2015-04/30-2017
- **Points of contact:** Lance Bosart, lbosart@albany.edu

i.d. Validation of Significant Weather Features and Processes in Operational Models Using a Cyclone Relative Approach

- **Lead testbed:** Hydromet Testbed; also DTC PI: Brian Colle, SUNY Stony Brook, CO-I: Edmund Kar-Man Chang, SUNY Stony Brook
- **Activities:**
  - Analyze the cyclone features in the operational GFS and Global Ensemble Forecast (GEFS) using one or more tracking algorithms.

    - Complete cyclone relative verification for different cyclone stages (from genesis to decay) and intensity of the temperature, moisture, precipitation and winds around the cyclone. Include mean stats; use thresholds to separate the stronger cyclone events. Separate bias results by flow regimes using compositing approaches.
    - Map the cyclone relative data back to the Earth-relative grid to determine geographical variations in the errors, which will help link results to various terrain/water features.
    - Develop cyclone-relative verification for key physical features: stability, surface fluxes, cloud cover, moisture flux, etc.
    - Develop a software interface for operational experiments (e.g., Winter Weather Experiment) and for EMC for ongoing evaluation efforts after the end of this project.
    - Develop a representative suite of cases and diagnostic tools that allow model developers and forecasters to better understand the origin of model errors of additional and future cases.
- **Milestones and deliverables:**
    - Year 1 (05/01/2015−4/30/2016)
        - Dataset collection (GFS and GEFS forecasts to day 10). SBU
        - Tracking of cyclones and matching of obs/model (SBU and EMC)
        - Basic verification metrics for cyclones (intensity, speed, track) using analyses from CFS Reanalysis and from RUC (after 2006). (SBU and EMC).
        - Collection of other observational datasets and interpolation to grid (multi---sensor precipitation, cloud products, etc.). (SBU and DTC)
        - Track other important cyclone features instead of central pressure (e.g. moisture plumes, low-level jets). (SBU and DTC)
        - Complete cyclone relative verification for different cyclone stages (from genesis to decay) and storm intensity of the temperature, moisture, precipitation, and winds around the cyclone. Include mean stats, also various threshold-based probabilistic metrics. (SBU – with visits to DTC, and DTC)
        - Separate the verification results by regimes using large-scale and regional flowcomposites. (SBU)
        - Develop MET module (METViewer) to composite statistics around cyclone (plotspatially and calc metrics), and ability to map results back to geographic grid. Potential approaches include
            - extending tropical cyclone QPF compositing capability currently under development within the DTC
            - development of percentile thresholding capability to be able to identify key synoptic features for both categorical statistics calculations and the MODE tool;
            - extending current series analysis capability to plot time-series analysis geographically;

- use MODE-Time Domain algorithms on cyclones to diagnose additional track attributes. (DTC)

- Year 2 (05/01/2016−4/30/2017)
  - Apply MODE to validate important features around the cyclone (e.g. jet streaks, low-level jets, heavy precipitation, strong surface winds, etc.) (DTC).
  - Map the cyclone relative verification results back to the Earth-relative grid, so the results can be related to various terrain, coast and sea-surface temperature features. (SBU and DTC).
  - Compute cyclone relative verification for relevant physical processes: stability, surface fluxes, temperature gradients, cloud cover, and a moisture budget around the storm (flux into/out of box, surface moisture flux, and precipitation fallout). (SBU).
  - Provide MET software tools for various operational centers to continue the validation efforts after this project for next generation of models. (DTC and EMC).
  - Use software and scripts for operational ensembles and in WPC Winter Weather Experiment. (SBU and HMT/ WPC).
  - Add several additional cases to the Mesoscale Model Evaluation Testbed (MMET) based on results of the analyses, to make them available to others in the community. (DTC).

- **Anticipated collaborating organizations:** DTC, HMT, NWS/NCEP/EMC, NWS/WPC
- **Priority (Low, medium, high):** High
- **Duration:** 05/01/2015-04/30-2017
- **Points of contact:** Brian Colle: brian.colle@stonybrook.edu

ii.a. Test and Evaluation of Rapid Post-Processing and Information Extraction from Large Convection Allowing Ensembles Applied to 0-3hr Tornado Outlooks

- **Lead Testbed:** Hazardous Weather Tested. PI: James Correia, University of Oklahoma Cooperative Institute for Mesoscale Meteorological Studies (CIMMS), and NOAA/SPC , CO-I: Daphne LaDue, University of Oklahoma; Christopher Karstens, Dustan Wheatley, and Kent Knopfmeier, CIMMS
- **Activities:**
  - Relate Probabilistic Hazard Information (PHI) variables to specific severe weather phenomena, (e.g., relate low-level vorticity to mesocyclones or tornadoes).
  - Provide context of severe weather proxy variables thereby relating model ensemble variables to observed phenomenon
  - Quantify capabilities/skill of guidance (prototype PHI tools) built on proxy variables allowing the forecaster to maximize use of the guidance, and to identify weaknesses so the forecaster can add value.

- Analyze impacts and effectiveness of use of prototype PHI tools to forecasters in identifying problem(s) of the day and in communicating forecast probabilities for tornadoes.

- **Milestones and deliverables:**
  Year 1 (May 1, 2015-April 30, 2016)
  - Complete development of WoF 3km full grid processing for 2d and 3d objects out to 2–3 hours
  - Test data flow in the HWT for full post-processing timing, data shipping probability generation (on the fly server client versus automatic generation)
  - Complete simple "cognitive interview" experiment designed to test protocols and survey instruments to ensure that we are getting information that we need to conduct a full test of prototype PHI tools (new post-processing and display systems) in 2016.
  - Assess forecaster work flow in an experiment to measure workload in closed vs open probability generation
  - Design, prepare, and conduct survey on task load analysis within the HWT Spring Experiment
  - Test current visualization system and adapt to meet needs of the project, iterate during the HWT Spring Experiment

  Year 2 (May 1, 2016-April 30, 2017)
  - Code refinement (speed up, more diagnostics) in object based post processing, data flow timing adjustments
  - Modify task load analysis survey protocols to help design next experiment
  - Derive contextual information regarding storm character including mesocyclone characteristics related to tornadic and non-tornadic supercells.
  - Analyze and adapt cognitive task analysis for final HWT Spring Experiment
  - Complete full experimental testing of prototype PHI tools in 2016 and 2017 HWT Spring Experiments
- **Anticipated collaborating organizations:** OAR/NSSL, NWS/SPC, Hazardous Weather Testbed
- **Priority (Low, medium, high):** High
- **Duration:** 05/01/2015-04/30-2017
- **Points of contact:** james.correia@noaa.gov

ii.b. Data Mining of High-Resolution Storm-Scale Datasets

- **Lead Testbed**: Hazardous Weather Tested. PI: Travis Smith, University of Oklahoma, CIMMS; Co-PIs: Dr. James Correia, Univ. Okla. CIMMS, CO-I: Dr. Christopher Karstens University of Oklahoma, Kiel Ortega, University of Oklahoma
- **Activities:** Perform pattern recognition and data mining on a subset of the MYRORSS dataset
  - Establish a methodology of objectively identifying storms of varying intensity in different environmental regimes.
  - Establish a methodology of objectively classifying storms.
  - Demonstrate the use of MYRORSS to validate storm mode and severity in a convection-allowing model.

        - Conduct real-time testing and evaluation with forecasters.
    - **Milestones and deliverables:**
        - February 2016- Initial storm type classification algorithm complete
        - April 2016 - Initial storm type classification algorithm operational on NSSL MRMS development system; visualization software prepared for HWT experiments
        - July 2016 - Testing of storm type classification algorithm by forecasters in HWT complete
        - February 2017 - Initial comparative study of the differences between model-based and MYRORSS-based storm type concluded
        - April 2017 - Visualization software updated to include model-based storm classification and methods to compare model-based and MYRORSS-based storm type distributions
        - July 2017 - Testing of storm type classification model/MYRORSS comparison by forecasters in HWT complete
    - **Anticipated collaborating organizations:** OAR/NSSL, NWS/SPC, NWS/NWSFOs, Hazardous Weather Testbed
    - **Priority (Low, medium, high):** High
    - **Duration**: 05/01/2015-04/30-2017
    - **Points of contact:** Travis Smith, tms@ou.edu

ii.c Information Extraction and Verification of Numerical Weather Prediction for Severe Weather Forecasting

- **Lead Testbed:** Hazardous Weather Testbed. PI: Israel Jirak, Storm Prediction Center (SPC), Co-PIs: Christopher Melick, SPC, Harold Brooks, NSSL, Matthew Pyle, EMC
- **Activities:** There are two primary components to this proposal for severe weather forecasting: verification and information extraction.
    - Verification
        - Identify the current accuracy and relative skill of experimental and operational CAMs for predicting organized severe-wind-producing mesoscale convective systems (MCSs), including derechos, using existing metrics.
        - Evaluate the membership of the SPC SSEO to determine the relative value of time-lagged members and usage of multiple model cores.
        - Explore and develop new methods, techniques, and datasets for verifying CAMs and SPC forecasts for organized severe-wind-producing MCSs.
    - Information Extraction
        - Develop and test model diagnostic variables for identifying organized MCSs with the potential for producing damaging surface winds in CAMs.
        - Incorporate diagnostic variables into operational CAMs for identifying organized severe-wind-producing MCSs.
- **Milestones and deliverables**:

Year 1 Verification

- Identify Current Accuracy and Skill of CAMs
    - Develop climatology of severe-wind-producing MCSs
    - Identify MCS cases for further inspection and re-running

        - Evaluate SSEO Membership – Assess skill of full SSEO in predicting sever-wind-producing MCS's
        - Explore New Metrics and Methods – Compare standard metrics to SEDI for CAM and SPC forecasts

Year 1 Information Extraction

- Develop and Test model diagnostics
    - Run cases using NSSL-WRF and examine 3-D structure
    - Formulate diagnostic variables for identifying sever wind-producing systems in CAMs

Year 2 Verification

- Identify Current Accuracy and Skill of CAMs – Assess current skill of CAMS in predicting severe-wind-producing MCSs
- Evaluate SSEO Membership – Assess skill of subset SSEO in predicting severe-wind-producing MCSs

- Explore New metrics and Methods – test and evaluate methods Year 2 – Information Extraction
    - Incorporate Diagnostic Variable into Operations – transfer code for diagnostic variables to operational models
- **Anticipated collaborating organizations:** EMC, Hazardous Weather Testbed
- **Priority (Low, medium, high):** High
- **Duration:** 05/01/2015-04/30/2017
- **Points of contact:** Israel Jirak, Israel.jirak@noaa.gov

ii. d Improvement of Convective/Severe Weather Prediction through an Integrative Analysis of WRF Simulations and NEXRAD/GOES Observations over the CONUS

- **Lead Testbed**: HWT. PI: Dr. Xiquan Dong, University of North Dakota; CoPIs: Dr. Aaron Kennedy, University of North Dakota, Dr. Matthew S. Gilmore, University of North Dakota
- **Activities:**
    - Evaluation of WRF simulated convective systems and precipitation: Existing HWT simulations from NSSL and NCEP will be evaluated via objective verification software utilizing NEXRAD reflectivity and Q2/Q3 precipitation products provided by the NSSL.
    - Develop and determine best practices for a microphysics based WRF ensemble: The development of an ensemble forecasting system for WRF where the current WRF default (WSM6) and numerous other microphysical schemes will be evaluated.
- **Milestones and deliverables:**

Year 1

- Provide NSSL-WRF ensemble and deterministic run output and details needed to reproduce these runs
- Provide necessary code for reproducing existing NSSL-WRF output products
- Assessment of HWT runs (tracking/classification/verification
- Development and assessment using SOMs
- Evaluation of WRF simulated Precipitation using Q2 product
- Hire web developer to assist with scripted model-to-webpage data transfers

- UND ensemble output webpage
- Microphysics ensemble development and running
- Satellite verification development
- Present the results to scientific meetings, such as AMS, AGU, EGU, and NOAA R2O meetings, and submit three papers to refereed journals

Year 2

- Assessment of microphysics ensemble (first half of year)
- Satellite evaluation
- SOM/HWT-style assessment
- Operational modeled ensemble runs
- Final assessment using previously developed techniques/
- Present research results to scientific meetings, and submit four papers to refereed journals
- Incorporate important findings and products into the future NSSL-WRF ensemble

- **Anticipated collaborating organizations:** NSSL/CIMMS, NASA LaRC
- **Priority (Low, medium, high):** High
- **Duration:** 05/01/2015-04/30/2017
- **Points of contact:** Dr. Xiquan Dong, dong@aero.und.edu

iii.a. Incorporation of near real-time Suomi NPP Green Vegetation Fraction and Land Surface Temp data into the NCEP Land modeling suite

- **Lead Testbed**: JCSDA. PI: Ivan Csiszar, NESDIS STAR. Co-PIs: Marco Vargas and Yunyue Yu (NESDIS/STAR), Michael Ek (NWS/NCEP/EMC), Zhuo Wang (ESSIC/U MD), Yihua Wu and Weizhong Zheng (IMSG)
- **Activities:**
  - Improve satellite data assimilation into land surface models (LSM) with two data products from Suomi NPP observations: the Green Vegetation Fraction (GVF) and the land surface temperature environmental data record (LST - EDR).
  - Calibrate and ingest of the GVF to incorporate daily updates to a weekly grid of greening indices, at 1km resolution over North American and at 4km resolution globally.
  - Integrate Suomi NPP EDR observations into global grids and evaluate against existing, coarser LST products from GOES observations to determine suitability for further use in land surface models and evaluating new operational models.
  - Quantify impacts of incorporating the new GVF and LST products with evaluations of experimental forecasts of severe weather that incorporate the new products in land
- **Milestones and deliverables:**
  - Year 1

GVF –Task 1

- May 2015 − April, 2016: acquire VIIRS GVF data product from CLASS or NDE PE-1 distribution zone
- May 2015 − April, 2016: acquire AVHRR derived GVF data product from CLASS
- August 2015 − September, 2015: develop software application to ingest GVF datasets (AVHRR and VIIRS), perform statistical analyses and implement data display tools.
- October 2015 − March, 2016: perform statistical analysis to establish the relationship between GVF datasets (AVHRR and VIIRS) and the 5 year AVHRR derived monthly climatology currently used in NCEP models.
- March 2016 − April, 2016: summarize results.

LST-Task 1

- May − November, 2015: develop the methodology and software package to retrieve and match up LST from all four datasets, and then composite all VIIRS granules on each day. Generate 24 hourly global gridded VIIRS LST files on each day in grib2 format at resolution of 0.036°.
- December, 2015 − May, 2016: on weekly basis, download all VIIRS granule data including all VIIRS granule-level temperature and geo-location data, and run the software package to produce the continuous global gridded VIIRS LST data at 0.036° and hourly resolution.
- By May 31, 2016, generate the hourly global gridded VIIRS LST product at resolution of 0.036°.

Year 2

GVF- Task2

- May 2016 − June, 2016: select extreme weather events from years 2014- 2015 to use as case studies in the NWP simulations
- July 2016 − Sept, 2016: incorporate the new VIIRS GVF data set into coupled model runs
- Oct 2016 − February, 2017: demonstrate impacts and utility of the high resolution, daily
  rolling weekly VIIRS GVF data set by comparing the surface energy budget using VIIRS
  GVF to that using the current AVHRR-derived GVF monthly climatology.
- Oct 2016 − February, 2017: quantify the reduction of errors in temperature, humidity, wind-speed forecasts, and the level of improvement in precipitation scores
- March 2017 − April, 2017: summarize results.

LST- Task2

- June − August, 2016: evaluate LST from NCEP NAM
- September − November, 2016: evaluate LST from NCEP GFS
- December, 2016 − February, 2017: evaluate LST from NLDAS
- March − April, 2017: summarize results.
- By April 30, 2017, the gridded VIIRS LST product can be utilized to assess the LST-simulated from NOAA NCEP weather forecast models at regional and global scales.

- **Anticipated collaborating organizations:** JCSDA, NESDIS/STAR, NWS/NCEP/EMC
- **Priority (Low, medium, high):** High
- **Duration:** 05/01/2015-04/30/2017
- **Points of contact:** Dr. Ivan Csiszar, ivan.csiszar@noaa.gov

5.12.4 Long-Term Objectives

The longer term goals for Years 3-5 (through FY2019) include refining and testing overarching performance targets as well as conducting phased testing for advanced system prototypes for NGGPS. Round I projects will be completed within 2 years, by May 2017 and all results evaluated for inclusion in final system requirements and service improvement objectives. The work will continue supporting collaborative phased transition testing for both internal and external investigators, covering progression of component, system, and forecast service testing progressing toward operational readiness testing with increasing emphasis on forecast service impacts testing.

Based on following upgrade schedule for GFS, GEFS:

FY15Q2: GFS upgraded to ~13 km resolution with numerous physics and data assimilation changes
FY15Q4: GEFS upgrades resolution and configuration
FY16Q1: GFS data assimilation goes to 4Dvar-like
FY16Q4 (?): Another GFS upgrade.

5.12.4 Milestones and Outcomes for Long-Term Objectives Beyond Five Years

Forecast services are based on accurate, timely and consistent objective guidance from NGGPS (and regular upgrades), that include impact-based probabilistic guidance developed from model ensemble information. NOAA testbeds and proving grounds continue to facilitate testing to meet objective and subjective targets, including impacts on workflow/workforce and on intended benefits from advanced forecast information/decision support .

5.13 Targets

Objective Testing Targets

- QPF Objective Skill Goal: Current Day 2 threat score for 1" threshold is attained at Day 3 by 2018
- Aviation weather skill goal: Improve the numerical prediction of instrument meteorological conditions (IMC) for the NGGPS by 25% over the baseline skill (to be determined) of GFS upgrade implemented in Q2FY15: T1534L4
- Severe Convection: Extend current severe weather forecast service skill by 24 hours (Goal FY18; e.g. current Day 1 skill at Day 2; Day 3 skill at Day 4)
- Fire Weather: Extend current fire weather forecast service skill by 24 hours (Goal FY18; e.g. current Day 1 skill at Day 2; Day 3 skill at Day 4)
- Test criteria and performance targets for weeks 2-4 are being developed.

Service Targets and Subjective Targets

Precipitation

- QPF Service Goal: Improving extreme event prediction from Day 0 to Day 10. For inclusion in AO: This entails improvements in warm season mesoscale prediction in the short range, as well as extending skill out in time (day 6, 7, ...10)
- QPF Effective, novel ways to synthesize information for the forecaster to communicate forecast impacts.
- Sample Performance Measure: Decision managers using week 2 QPF information for planning.

Winter Weather

- Examine alternative objective skill goals, based on analysis of reforecast and soon-to-be available gridded snowfall analysis, and develop recommended forecast skill target, Q4 FY15
- Achieve Winter Storm Warning lead time of 24 hours while maintaining POD of 0.90 by 2018. This lead time is 4 hours longer than the current GPRA goal. For inclusion in AO: This entails improved prediction of explosive cyclogenesis, improved snowfall algorithms, extending skill out in time (day 6-10), and finally novel ways to synthesize information for the forecaster.

Aviation weather

- Advance the spatial and temporal consistency of NWS products and decision support services provided to the aviation community based on increasing global model resolution (e.g. 13km, T1534L64 implemented in Q2FY15) and future higher-resolution windows. A mesoscale global model with higher-resolution windows is a major step toward meeting NTSB Safety Recommendation A-14-17 on achieving consistency across NWS aviation and non-aviation specific statements, forecasts, advisories, and warnings that provide information on hazards important to aviation safety such as low clouds, ceilings, visibility, wind, convection, and precipitation.
- Improve the basis for decision support services re instrument meteorological conditions, with numerical prediction guidance of clouds and visibility, and particularly hydrometeors

Severe weather

- Severe Convection: Develop effective objective measures of severe weather forecast skill appropriate for both initial NGGPS (evolved GFS) and future convection-allowing nests through work with forecast specialists and numerical modelers. Baseline current global forecast system and Storm Prediction Center objective severe weather forecast skill, and refine objective goals for NGGPS project and associated Hazardous Weather Testbed testing (Goal: 4QFY15).
- Severe Convection: Improve Day 2-10 national severe weather and convection sensitive services exploiting service-oriented objective and subjective metrics and iterative testbed experiments, through a focus on NGGPS forecasts of severe storm supporting mesoscale environments, boundary layer thermodynamic structures, and convective precipitation (objective metrics established through projects initiated in FY15).

- Severe Convection: Develop specialized NGGPS information extraction to support severe weather service improvement goals, and approaches for effective communication and use of improved NGGPS predictive skill by NWS forecasters and community decision makers

Fire weather

- Develop effective objective measures of fire weather forecast skill appropriate for both initial NGGPS (evolved GFS) and future convection-allowing nests through work with forecast specialists and numerical modelers. Baseline current global forecast system and Storm Prediction Center objective fire weather forecast skill, and refine objective goals for NGGPS project and associated Hazardous Weather Testbed testing (Goal: 4QFY15).
- Improve Day 2-10 national fire weather services exploiting service-oriented objective and subjective metrics and iterative testbed experiments with a focus on implied fuels, large-scale fire weather conditions, and implicit or explicit forecasts of convection and potential new fire starts by associated lightning (objective metrics established by FY16).
- Develop specialized NGGPS information extraction to support fire weather service improvement goals, and approaches for effective communication and use of improved NGGPS predictive skill by NWS forecasters and fire weather decision makers.

Tropical weather

- Improve tropical cyclone genesis, track, intensity, and size prediction guidance out through seven days.
- Improve probabilistic wind and storm surge guidance for tropical cyclones out through seven and three days, respectively.

Marine weather

- Improve marine (ocean, wave) prediction guidance for the Atlantic Ocean, Gulf of Mexico, Caribbean Sea and Pacific Ocean out through seven days.

Climate outlooks

- Develop skillful extended range model guidance to support current experimental Week 3-4 temperature and precipitation outlooks
- Extend decision support services for extreme events to lead times of Week 3-4: assess potential target performance based on available and appropriate model forecast guidance

5.14 Appendix: Readiness Levels

RL 1	Basic principles observed and reported
RL 2	Technology concept and/or application formulated: Applied research.
RL 3	Analytical and experimental critical function and/or characteristic proof-of-concept
RL 4	Component/subsystem validation in laboratory environment:
RL 5	System/subsystem/component validation in relevant environment:
RL 6	System/subsystem model or prototyping demonstration in a relevant end-to-end environment
RL 7	System prototyping demonstration in an operational environment
RL 8	Actual system completed and "mission qualified" through test and demonstration in an operational environment
RL 9	Actual system "mission proven" through successful mission operations