

## Final Progress Report

November 29, 2018

Submitted by: Betsy Weatherhead, U. Colorado at Boulder

[Betsy.Weatherhead@Colorado.edu](mailto:Betsy.Weatherhead@Colorado.edu) (303) 497 6653

### Grant Information

<b>CFDA Number:</b>	11.468	<b>Award Period:</b>	09/01/2016 - 08/31/2018	<b>Program Office:</b>	NWS Office of Science Technology (OST)
<b>Program Officer:</b>	Christopher Hedge	<b>Program Officer Phone:</b>	301-427-9242	<b>Program Officer Email:</b>	Christopher.Hedge@noaa.gov
<b>Grants Specialist:</b>	Rachel E. Miller	<b>Grants Specialist Phone:</b>	301-628-1332 EXT-	<b>Grants Specialist Email:</b>	rachel.e.miller@noaa.gov
<b>Total Federal Funding:</b>	\$400,000.00	<b>Total Non Federal Funding:</b>	\$0.00	<b>Multi-Year:</b>	Yes
<b>Organization Name:</b>	REGENTS OF THE UNIVERSITY OF COLORADO, THE	<b>ASAP Recipient:</b>	Yes	<b>High Risk Recipient:</b>	No
<b>SF-425 Frequency:</b>	Semi-Annual Cash Flow with Final Full Report	<b>Progress Report Frequency:</b>	Semi-Annually	<b>Final Progress Report:</b>	Comprehensive - a interim report is required
<b>Final Reports Due On:</b>	11/29/2018	<b>Project Title:</b>	Critical Comparison and Evaluation of Skill Scores in Support of NGGPS		
<b><u>PIs - PDs:</u></b>	Elizabeth Weatherhead	<b>Closeout Date:</b>	N/A		

### Critical Comparison and Evaluation of Skill Scores in Support of NGGPS

**Final Report**

November 29, 2018

**U. Colorado Speed Type: 13009363**

National Weather Service  
National Oceanic and Atmospheric Administration, Department of Commerce  
Funding Opportunity #: NOAA-NWS-NWSPO-2016-2004713  
CFDA # 11.468, Applied Meteorological Research

**Critical Comparison and Evaluation of Skill Scores in Support of NGGPS**

**PI:** Dr. Betsy Weatherhead  
University of Colorado at Boulder  
Cooperative Institute for Research in Environmental Sciences (CIRES)  
UCB 219, CIRES Building Room 318  
Boulder, CO 80309  
(303) 497-6653  
Betsy.Weatherhead@Colorado.edu

**Institutional Representative:** Marcos Carvalho  
Proposal Analyst, Office of Contracts and Grants  
3100 Marine Street, Room 479  
Boulder, CO 80303-1058  
(303) 735-6992  
[marcos.carvalho@colorado.edu](mailto:marcos.carvalho@colorado.edu)

**Critical Comparison and Evaluation of Skill Scores in Support of NGGPS and HFIP.**

Elizabeth C. Weatherhead, University of Colorado at Boulder  
Proposed Budget: \$400,000; Final Costs: 310,313.75  
Proposed Budget Period: September 1, 2016 to August 31, 2018  
Actual Budget Period for PI: September 1, 2016 to February 28, 2018  
To support collaboration with NOAA's NGGPS Program

This project is in response to Funding Opportunity #: NOAA-NWS-NWSPO-2016-2004713.

## **Final Progress Report Summary**

***No additional work was completed during the final reporting period due to the departure of Dr. Weatherhead.***

Work continued at an accelerated pace on addressing verification metrics for different geographic areas and for non-standard parameters such as waves, land surface, hydrology and sea ice. Dr. Weatherhead, the PI, retired on February 28, 2018 and transferred the remainder of the work to the sub-contractor, the National Center for Atmospheric Research, where the work is continuing. NCAR's Developmental Testbed Center (DTC), under the direction of Dr. Tara Jensen, is continuing the remainder of the tasks.. This project, in part, evaluates the techniques incorporated in MET to see if they are state of the art statistical techniques for evaluation. Progress in the five areas of focused work are:

1) Descriptive Forecast Verification;

Report has been finalized on how to summarize verification statistics for forecast model evaluation. Key features of the report: all descriptive statistics are autocorrelated in a manner consistent within a geographic region and parameter type.

2) Decision Support Statistics for Forecast Verification;

Report has been finalized on how to evaluate decision support statistics for model comparison. Key features of the report: pairwise results can be evaluated making use of autocorrelated error components; number of runs needed depends directly on the size of the verification to be evaluated.

3) Key Metrics for Non Atmospheric Parameters;

Report has finalized on how to evaluate hydrological, land-use and wave forecasts. Results indicate that high spatial terrain techniques are needed for most non-atmospheric (hydrological, land-use and wave) forecasts.

4) Descriptive Verification;

Efforts are finalized on descriptive verification statistics for hydrological, land-use and wave forecasts. These build from what is currently being carried on in NOAA and results were presented to NWS at recent meeting.

5) Decision Support for Non Atmospheric Parameters;

Efforts are at mid-stage for decision support statistics for non atmospheric parameters. This effort is building from the descriptive verification statistics. The literature review is complete and initial data have been downloaded and examined.

## **Communication of Results and Deliverables**

Presentation at TIES-GRASPA' 27<sup>th</sup> Annual Conference of the International Environmentrics Society Joint with Biennial GRASPA Conference, July 24-27, 2017. Presentation: Evaluation of forecasts with imperfect data.

Presentation at the 7th International Verification Methods Workshop. Presentation: Identification of Small Improvements in Forecast Verification. May 3-11, 2017.  
Science Conference Verification on Extreme Events, May 8-9.

NGGPS Science Team Meeting

Analysis summary: Hydrological Forecast Evaluation

Analysis summary: Land Surface Forecast Evaluation

Publication: Identification of small forecast Improvements, presented at AMS Annual Meeting, 2018

Publication: Planning observing systems of the future, presented at the AGU Fall Meeting, 2017

Participated in and presented research on wave verification at the CLIVAR meeting on "Sea Level Rise and Coastal Impacts." Presentation: evaluation of wave and hydro verification models.



# CRITICAL COMPARISON AND EVALUATION OF SKILL SCORES IN SUPPORT OF NGGPS

Principal Investigator: Weatherhead, Elizabeth

Financial Report as of June, 2018

Start Date 2016-09-01 End Date: 2018-08-31

Fully funded. Sub to UCAR is project 1554500. Betsy has retired and chose to not name a different PI. CU will relinquish the award to NOAA.

NAT'L. OCEANIC AND ATMOSPHERIC ADMINISTR

Award: NA16NWS4680016

Project: 1554499

Speedtype(s): 13009363

F&A Rate: 26%

Department: 10086

## To Date Actuals

Account Description	Total Funding	Monthly Actuals	To Date Actuals	% To Date	Future Expenses		Total Balance
					Salary, Travel, Operating	Other Projections	
400000 - 418399 -- SALARIES AND WAGES	\$139,025	\$0	\$101,240	72.8%	\$0	\$0	\$37,785
418400 - 439999 -- BEN	\$54,624	\$0	\$37,760	69.1%	\$0	\$0	\$16,864
450000 - 699999 -- OPERATING EXPENSES	\$10,813	\$0	\$8,542	79.0%	\$0	\$0	\$2,271
700000 - 709999 -- TRAVEL	\$24,509	\$0	\$20,313	82.9%	\$0	\$0	\$4,196
810000 - 812499 -- FIXED ASSET COSTS	\$5,000	\$0	\$5,157	103.1%	\$0	\$0	(\$157)
950000 - 952499 -- FACILITIES & ADMIN COSTS	\$59,532	\$0	\$43,642	73.3%	\$0	\$0	\$15,890
<b>Total</b>	<b>\$293,503</b>	<b>\$0</b>	<b>\$216,654</b>	<b>73.8%</b>	<b>\$0</b>	<b>\$0</b>	<b>\$76,849</b>

Anticipated available balance (for direct costs):

\$60,991

## Payroll Projections (12 months)

EmpID	Employee Name	Budget	To Date Payroll	7-18	8-18	9-18	10-18	11-18	12-18	1-19	2-19	3-19	4-19	5-19	6-19	Total	Remaining Balance	Benefit Cost
Speedtype: 13009363																		
108259	WEATHERHEAD, ELIZABETH C	\$139,025	\$101,240													\$0	\$37,785	
Speedtype: 99999999																		
	0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0
<b>Total</b>		<b>\$139,025</b>	<b>\$101,240</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$37,785</b>	<b>\$0</b>

## Travel Projections

EmpID	Employee Name	Final Destination	Speedtype	Start Date	End Date	Total
Total						\$0

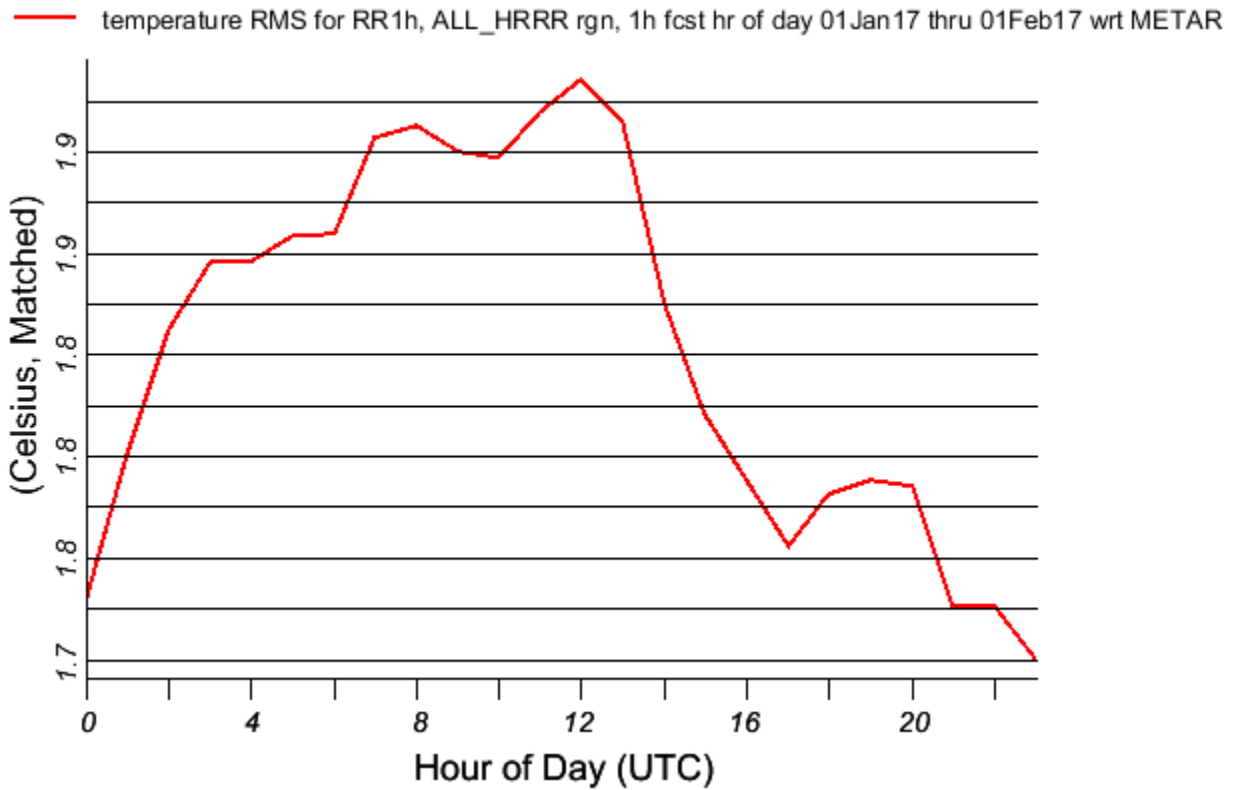
## Other Projections

Budget Category	Description	Amount	Benefit	F&A	Total
		\$0	\$0	\$0	\$0
<b>Total</b>		<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>

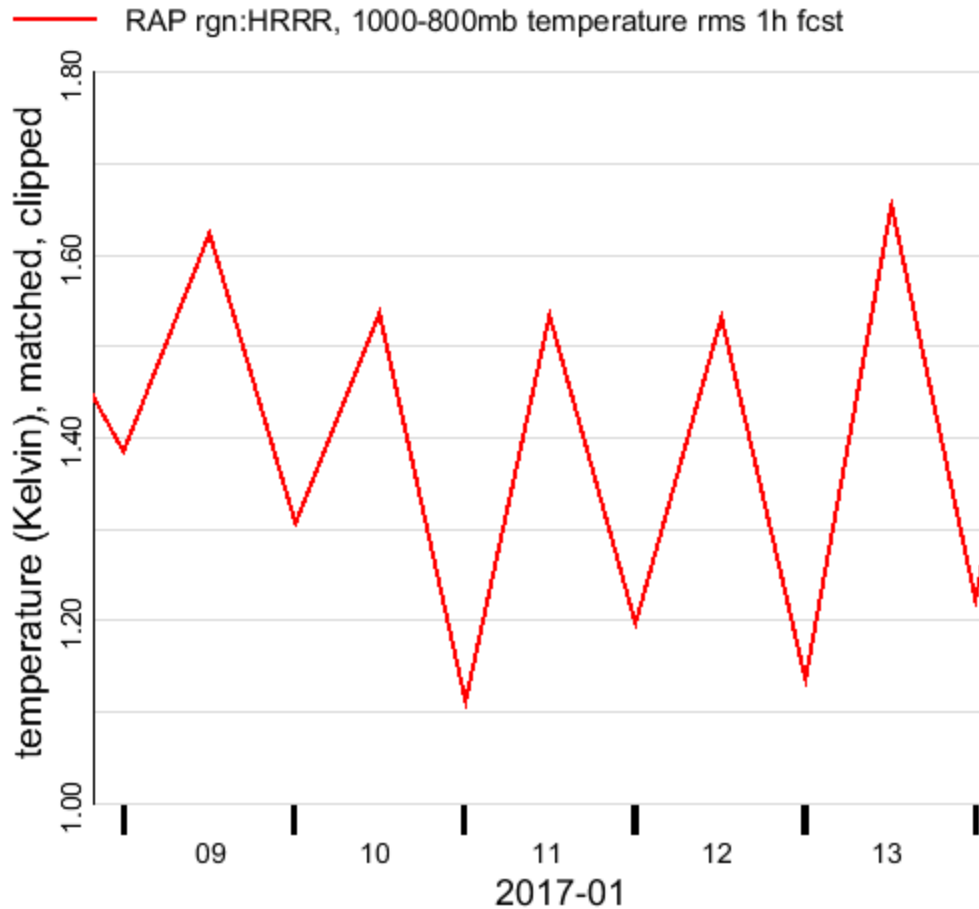
## 1. Key scientific accomplishments

### Identification of need for more expansive tests (AR1)

Verification statistics can now be more robust and resultant error bars likely smaller because of the new understanding on the diurnal aspects of the errors in forecasting. The key scientific result is fundamentally a merging of what has been known about forecasting for some time, that forecast errors are dependent on time of day with the appropriate statistics that respect that.



Caption: A plot of the RMS errors for temperature as a function of time of day for a single forecast run. This diurnal signature is characteristic of the model; the general tendency is seen for most days.



Caption: Forecast verification for temperature RMSE for five days in 2017. For these twelve hour sequential forecasts we see the negative autocorrelation (a high forecast error is more likely followed by a lower forecast error).

The result of this negative autocorrelation in forecast error means that we can be more confident the average results for the accuracy of the forecast than standard statistics might indicate. In this case, because of the negative autocorrelation, the forecast error is smaller than if autocorrelation is ignored.

We can also see from this plot that the variability in the daytime error may be smaller than for the nighttime errors, meaning that we may want to separate these forecast skill characteristics and develop separate metrics, each of which will likely have smaller error bars.

### Identification of the needs for Hydrology

## Non-atmospheric parameters: Hydrological Models

NGGPS Verification

**Attendees:** Brian Cosgrove, Tara Jensen and Betsy Weatherhead (on the phone)

**Perspective:** Hydrological models are increasingly important to NOAA and the country as a whole. A number of parts of NOAA are involved in developing and testing hydrological models. **Key input to GFS:**

hydrological modeling is an integrated part of GFS, with humidity, precipitation, soil moisture and surface fluxes being used as model inputs via data assimilation.

**Goals for hydrological models:** The societal relevance of hydrological forecasting can not be underestimated. Both extremes in hydrological forecasts (droughts and floods), as well as standard hydrological forecasts are important. The hydrological efforts in NOAA are currently focused on North America which has a strong focus on the CONUS. There are currently four modes of hydrological forecasting in NOAA:

Near real time: hourly analysis with the HRRR.

Short range—every hour powered by Rap.

Medium—powered by GFS, run once per day, up to ten day lead time.

Long—powered by CFS, (16 members) up to 30 day lead time.

**Timescales for verification:** Hydrological forecasts on short term (0-24 hours) through to seasonal and climatic forecasts are important nationally and internationally. All twelve months of the year are important for examination, as are diurnal cycles of precipitation. Ideally, we'd like to capture some droughts and floods—perhaps this can be achieved with retrospective forecasts?

**Key parameters to examine:** Precipitation and temperature will be the two most important parameters to look at from the global models. At a secondary level of importance, solar, wind speed, humidity, net long-wave radiation will have direct effects on hydrological forecasting.

#### **Evaluation approaches:**

Water Resource Evaluation Service (WRES) is a program, led by Mark Fresch, within NOAA to evaluate all water predictions. WRES is still in the planning stages. They plan to use R based software, open source. Betsy is happy to work with R programming and the WRES team.

HEFS (Hydrological Ensemble Forecast System). It's not a gridded model; it runs at the forecast centers and is a point type model. It currently uses the GEFS and does a broad range of verification. Brian indicated some concern about the future of GEFS. Tara indicated that METS' initial challenge is to make sure that it can reproduce everything that GEFS did.

The importance of proper verification of hydrological forecasting is large and the complexity is enormous. Likely, the best role of Betsy and Tara will be to coordinate the ongoing activities within NOAA's Office of Water Prediction with the NGGPS process. A possible step forward would be to include some of the most important water metrics within MET so that tests on any global model configuration would automatically include some basic evaluations of some key hydrological outputs.

An important, outstanding question is whether verification for NGGPS will include ensemble forecasting. If this is the case, who will do that?



**Data for Verification:** Brian indicated that data could be currently downloaded from: <http://water.noaa.gov/about/nwm>; Betsy has verified that. Brian indicated that hydrological verification often occurs at the basin level—and there are 2.7 million basins that they will consider (roughly 1 square mile), however the 12 CONUS regions would be useful for verification.

Transforming NOAA Water Prediction	
TODAY	THE FUTURE
Approximately 4000 forecast locations at points	Approximately 2,700,000 forecast stream reaches
Forecast river flow/stage	Forecast all hydrologic parameters which define the water budget
Driven by large catchment “lumped” modeling	Driven by high resolution Earth System modeling
Average basin size greater than 420 square miles	Average basin size ~1 square mile
Impact-based forecasts at selected points	Predictions linked with detailed local infrastructure data to communicate street level impacts
13 River Forecast Centers developing separate versions of the same regional model	NOAA, academia, and federal partners developing/evolving same national, community-based model
For the hydrology community, the implementation of the NWM and the leap ahead capability it provides parallels the implementation of mesoscale atmospheric models in the 1970s (i.e., model resolution substantially greater than available observational network)	

[More specific forecasts can be understood in terms of time frames, with the National Water Center identifying](#)

## General Framework:

The NWM is run in four configurations:

1. Analysis and assimilation: snapshot of current hydrologic conditions
2. Short-Range: 15-hour deterministic (single value) forecast
3. Medium-Range: 10-day deterministic (single value) forecast
4. Long-Range: 30-day ensemble forecast

## Analysis and Assimilation:

The Analysis and Assimilation configuration cycles hourly and produces a real-time analysis of the current streamflow and other surface and near-surface hydrologic states across the contiguous United States (CONUS). This configuration also produces a single model restart file which is used to initialize the 15 hour, 10 day and 30 day forecast simulations. Meteorological forcing data are drawn from the MRMS Gauge-adjusted and Radar-only observed precipitation products along with short-range RAP and HRRR.

## Forecast Ranges

### Short Range

Forced with meteorological data from the HRRR and RAP models, the Short Range Forecast configuration cycles hourly and produces hourly deterministic forecasts of streamflow and hydrologic

states out to 15 hours. The model is initialized with a restart file from the Analysis and Assimilation configuration and does not cycle on its own states.

#### Medium Range

The Medium Range Forecast configuration is executed once per day, is forced with GFS model output and extends out to 10 days. It produces 3-hourly deterministic output and is initialized with the restart file from the Analysis and Assimilation configuration.

#### Long Range

The Long Range Forecast cycles four times per day (i.e. every 6 hours) and produces a daily 16-member 30-day ensemble forecast. There are 4 ensemble members in each cycle of this forecast configuration, each forced with a different CFS forecast member. It produces 6-hourly streamflow and daily land surface output, and, as with the other forecast configurations, is initialized with a common restart file from the Analysis and Assimilation configuration.

#### Output:

All NWM output will be stored in NetCDF format in one of three file types:

1. 1km gridded NetCDF (land surface variables and forcing)
2. 250m gridded NetCDF (ponded water depth and depth to soil saturation)
3. Point-type NetCDF (stream routing and reservoir variables)

The two gridded files cover a rectangular domain stretching beyond the CONUS roughly from 19N to 58N, while the point NetCDF files contain model output from the CONUS and hydrologically contributing areas.

Analysis & Assimilation	Short-Range	Medium-Range	Long-Range
<b>Cycling Frequency</b>			
Hourly	Hourly	Daily at 06Z	Daily Ens (16 mem)
<b>Forecast Duration</b>			
- 3 hrs	0-15 hours	0-10 days	0-30 days
<b>Forecast Latency</b> (latency of external forcing data accounts for most of delay)			
1 hour	1 hour 45 mins	6 hours	19 hours
<b>Meteorological Forcing</b>			
MRMS blend/ HRRR/ RAP bkgnd.	Downscaled HRRR/RAP blend	Downscaled GFS	Downscaled & bias- corrected CFS
<b>Spatial Discretization &amp; Routing</b>			
1km/250m/NHDPlus Reach	1km/250m/NHDPlus Reach	1km/250m/NHDPlus Reach	1 km/NHDPlus Reach
Assimilation of USGS Obs			
Reservoirs (1260 water bodies parameterized with level pool scheme)			

#### Important documents to support this work:

- Cosgrove, Brian A., Dag Lohmann, Kenneth E. Mitchell, Paul R. Houser, Eric F. Wood, John C. Schaake, Alan Robock et al. "Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project." *Journal of Geophysical Research: Atmospheres* 108, no. D22 (2003).
- Krzysztofowicz, Roman. "The case for probabilistic forecasting in hydrology." *Journal of hydrology* 249, no. 1 (2001): 2-9.
- Schaake, John C., Qingyun Duan, Victor Koren, Kenneth E. Mitchell, Paul R. Houser, Eric F. Wood, Alan Robock et al. "An intercomparison of soil moisture fields in the North American Land Data Assimilation System (NLDAS)." *Journal of Geophysical Research: Atmospheres* 109, no. D1 (2004).
- Weatherhead, Elizabeth C., Amy J. Stevermer, and Barry E. Schwartz. "Detecting environmental changes and trends." *Physics and Chemistry of the Earth, Parts A/B/C* 27, no. 6 (2002): 399-403.
- Xia, Youlong, Kenneth Mitchell, Michael Ek, Justin Sheffield, Brian Cosgrove, Eric Wood, Lifeng Luo et al. "Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products." *Journal of Geophysical Research: Atmospheres* 117, no. D3 (2012).

## Non-atmospheric parameters: Land Surface Models

## NGGPS Verification

## Discussion Summary

December, 2016

**Attendees:** Michael Ek, Tara Jensen and Betsy Weatherhead (on the phone)

**Perspective:** Land surface models are used to estimate impact of weather to the cryosphere, soil and vegetation globally. Changes to the NCEP Global model will directly impact land surface models which derive critical input from the global models. Michael Ek validates land surface models.

**Key parameters from GFS:** temperature, incoming solar, longwave radiation, humidity and pressure, albedo, soil type and vegetation types are also used, but these are not inputs from the GFS.

**Key input to GFS:** Land surface models provide heat, moisture, momentum and drag that influences surface layer turbulence. Output from the land models goes into the GFS, CFS, mesoscale model (Tania Schmiernova has the rough land model working with RAP/HRRR).

**Key observation sets for verification:** Fluxnet is a network of networks which is used internationally for verification of land surface models. One key parameters is thermal conductivity that depends on soil type and soil moisture content.

**Goals for land surface models:** Calculating drag, surface layer turbulence and heat, moisture and momentum fluxes as well as thermal and hydraulic movement are areas of active research.

**Timescales for verification:** Monthly averaged diurnal cycles of all critical parameters. This could be adequately achieved with four months of verification (e.g. July, October, January, April).

**Key parameters to examine:** gradients of wind, humidity and temperature in the boundary layer are key metrics for determining the success of a model. The most relevant GFS output is 2m wind, temperature and humidity, so boundary layer and radiation schemes are necessary for verification.

Mike would like to have a hierarchy of model testing up to everything being coupled with FV-3.

Possibly, soil moisture and soil temperature will be the two direct parameters to evaluate:

$$\frac{\partial \Theta}{\partial t} = \frac{\partial K}{\partial t} + \frac{\partial}{\partial \theta} \left( \frac{\partial \Theta}{\partial \theta} \right) + \dots$$

Where  $D_{\theta}$  is the soil moisture diffusivity and  $K_{\theta}$  is the hydraulic conductivity are non-linear functions of soil moisture and soil type.

And for soil temperature:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left( K_T \frac{\partial \theta}{\partial z} \right)$$

Where  $C_T$  is the thermal heat capacity and  $K_T$  is the soil thermal conductivity (not functions of soil moisture and soil type).

**Potential metrics for comparison if controlled, comparison runs are made:**

Comparison of 2-metr temperature and humidity using pairwise autocorrelation from GFS and FV-3 throughout the diurnal cycle for four months out of the year sampling each season.

**Important people to work with on verification:** Helin Wei; [Jairui.Dong@noaa.gov](mailto:Jairui.Dong@noaa.gov), Fang Lin knows about Jack Woolen (climate scientist) who wrote some scripts to pull out all radiosondes that are in grasslands, forests, etc. From the land model perspective, it's very important to understand the soil type.

**Important documents to support this work:**

- Chen, Fei, Kenneth Mitchell, John Schaake, Yongkang Xue, Hua-Lu Pan, Victor Koren, Qing Yun Duan, Michael Ek, and Alan Betts. "Modeling of land surface evaporation by four schemes and comparison with FIFE observations." *Journal of Geophysical Research: Atmospheres* 101, no. D3 (1996): 7251-7268.
- Cosby, B. J., G. M. Hornberger, R. B. Clapp, and ToR Ginn. "A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils." *Water Resources Research* 20, no. 6 (1984): 682-690.
- Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley. "Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model." *Journal of Geophysical Research: Atmospheres* 108, no. D22 (2003).
- Johansen, O. "Thermal conductivity of soil and rock." *Frost I Jord* 16 (1975): 13-21.
- Wang, Bin, June-Yi Lee, In-Sik Kang, J. Shukla, C-K. Park, A. Kumar, J. Schemm et al. "Advance and prospectus of seasonal prediction: assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004)." *Climate Dynamics* 33, no. 1 (2009): 93-117.
- Weatherhead, Elizabeth C., Amy J. Stevermer, and Barry E. Schwartz. "Detecting environmental changes and trends." *Physics and Chemistry of the Earth, Parts A/B/C* 27, no. 6 (2002): 399-403.

## 2. Progress on testing, summary of evaluation and/or verification of proposed improvements

Purchased of dedicated computer. The computer specifications were determined by the DTC staff, when the PI's older computer was deemed inadequate.

Uploaded of MET, R and Office software to the project computer.

The PI took training classes on using MET at NCAR. This was before the project computer was available, so the on-line tutorials have had to suffice since then. Fortunately, the on-line tutorials are very good and have been very useful.

### 3. Interactions with scientists at EMC, other NCEP Centers, NOAA labs, and/or NOAA Testbeds

The PI has an office in NOAA's Boulder building that allows for frequent interaction with GSD scientists including Bonny Strong.

Communication has started with Mike Ek, Hendrik Tolman and Arun Chowla.

The PI gave a presentation on NGGPS biweekly calls on Wednesday, January 18 (Appendices C and D)

The PI will be participating in the community meeting at NCWCE April 18-21 and helped advertise the meeting through the American Meteorological Society's Forecast Improvement Group.

#### 4. Progress against milestones/schedules in proposal

The milestones presented in the original proposal are presented here:

[illegible]

The project is slightly behind schedule due to the lack of appropriate data of intercomparisons.. For each of the two sub-projects that were scheduled to begin in the first quarter of this work plan, the work was broken down into additional sub-projects as explained below.

	1	Name	Duration	Start	Finis	2016				2017				2018			
						Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
1		1. Development of Descriptive Statistics using Appropriate Autocorrelation	9m	09/01/2016	05/10/2017												
2		Finalization of project plan and acquisition of data	1m	09/01/2016	09/28/2016												
3		Mid-latitudes	1m	09/29/2016	10/26/2016												
4		Tropics	1m	10/27/2016	11/23/2016												
5		High latitudes	1m	11/24/2016	12/21/2016												
6		Testing of Results	1m	12/22/2016	01/18/2017												
7		Presentation of Results	1m	01/19/2017	02/15/2017												
8		Transition of algorithms to MET	2m	02/16/2017	04/12/2017												
9		Final Report on Descriptive Statistics	1m	04/13/2017	05/10/2017												

The analysis for Mid-latitudes, tropics has begun and are showing that autocorrelation plays a significant role in most results. The impact on the size of the error bars (the standard error on the mean) appears to be much larger than on the estimate of the metric itself.

	1	Name	Duration	Start	Finis	2016				2017				2018			
						Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
19		3. Identification of Key Metrics for Evaluation of Non Atmospheric Products	140d	11/01/2016	05/15/2017												
20		Finalization of Engagement and Metric Selection Plan	1m	11/01/2016	11/28/2016												
21		Land Metrics and acquisition of data	1m	11/29/2016	12/26/2016												
22		Hydrology Metrics and acquisition of data	1m	12/27/2016	01/23/2017												
23		Sea Ice Metrics and acquisition of data	1m	01/24/2017	02/20/2017												
24		Wave Metrics and acquisition of data	1m	02/21/2017	03/20/2017												
25		Interim Report and initial analyses	1m	03/21/2017	04/17/2017												
26		Finalization of recommendations for evaluation	1m	04/18/2017	05/15/2017												

Discussions for identification of key metrics for evaluation of non-atmospheric products are ahead of schedule with significant improvement in understanding for both hydrology and land metrics. Now that the computer is in place, we are working more actively on the acquiring the initial data.

## 5. Any previously unreported changes to the execution of the originally submitted proposal

The project is slightly behind schedule due to a lack of NGGPS data for the verification work of standard parameters. Work on verification of non-atmospheric parameters is progressing on schedule.. As already noted, the dedicated computer was late in arriving. The installation of software was--and continues to be--a notable challenge. This is not completely unexpected as atmospheric computing needs are always significant.

The diurnal aspects of autocorrelation, as discussed in Section 1: Key scientific accomplishments, was not fully appreciated at the time the proposal was written, but has turned out to be an interesting and likely constructive addition to project.



Discussion of non-Gaussian behavior by the NGGPS validation team may result in some additional work, but this would only take place with approval of the program manager.

The discussion of the impact of imperfect observations used for verification was unexpected. The PI wrote a brief white paper on this subject for the NGGPS verification team. The work was not extensive and was within the general scope of the work plan.

## **6. Any outcomes that could be transitioned or offered to operations (previous outcomes can be repeated)**

The issue described in section one of the impacts of negative autocorrelation may be considered for transition to MET. This decision is beyond the scope of this proposal but the PI is currently preparing materials to allow the MET team to consider this transition. The results will have a direct impact on the confidence intervals given for the verification results, so the decision will not likely take place without some careful consideration. This effort may warrant a publication to assure transparency and allow for peer reviewed efforts.

## **7. Budget issues as required in the (separate) financial forms – ok to mention critical information in progress report**

Budget spending is roughly on schedule. Some details of note:

1. The dedicated computer was purchased, although delivery of the computer was delayed due to manufacturer issues.
2. The subcontract to NCAR was established and initial payments started.
3. Travel was light, but is expected to increase with trips to DC and meetings planned for April-July.



## Quality in Weather Verification

NGGPS

Betsy Weatherhead, U. of Colorado and Tara Jensen, NCAR

Verification of weather forecasts relies directly on accurate data to achieve successful, reliable results. While all collected data may be stored and available for use, not all data are of equal quality and usefulness for the verification. Understanding tools available for screening of data quality is important to assure the best use of all available data. Individual datasets used for verification all have documented assessments of known problems and expectations for accuracy, although the evaluation of data and production of quality assessments is not uniform across datasets. Individual observations are screened as part of the ingestion process within MADIS, but these screenings are cursory, looking for very specific issues of reasonableness. Data assimilation fields can be useful for forecast evaluation because they are the result of merged fields that can result in more physically consistent observations. However, data assimilation fields are not immune from problems with data quality.

A simplistic approach is to treat all verification observations as the “truth” and all differences between a forecast and an observation as a “forecast error.” This simplistic approach ignores known and potential problems with data. Kloog et al.(2014) report that satellite derived temperature observations, often used for forecast verification, match surface tempera observations with a MSE of roughly 2° C, although the two sets of observations show high agreement with respect to day-to-day variability ( $R^2 \approx 95\%$ ). Even intercomparisons of available sonde data show agreements in temperature of greater than 0.5 (Bian et al., 2011). Sevruck et al. (2009) report that wind and evaporation can introduce differences of rain between 4 and 6%, while problems with snow or mixed precipitation can be as large as 60% based on wind conditions. Ingleby and Huddleston (2007) report differences in ocean temperature of greater than 1% and note that over 20% of collected data were rejected once strict quality controls were introduced. In light of these and many other published studies on the potential errors in observations commonly used for forecast verification, the question of how to address observational uncertainty becomes more important than ever before.

Schutgens look at the very relevant question of how well point observations can be expected to agree with spatial averages using WRF-CHEM. They conclude that differences can be as large as 30-80% and more for aerosol parameters. Multiple observations in a single gridbox can help reduce these differences, but not likely eliminate the differences; offsets of observations from the gridbox can cause disagreements between models and observations.

A few studies have sampled the impact of observational uncertainties on seasonal forecasts. Junya Hu presented results at last year’s AGU showing that initial sea temperature errors could cause significant forecast errors of El Nino events. Similarly, Rong Fang and WanSuo Duan (2017) showed that initial temperature errors caused Indian dipole events. For both of these studies, very specific regions where the temperature errors would be most impactful.

MADIS as the primary ingest of observations used both for modeling and verification of models has an important role in real-time processing of observations. Galarus et al, (2012) reported on some of the quality checks that are done to assure usefulness of observations before they are used in data assimilations. Verification of weather forecasts and data assimilation both rely heavily on accurate data to achieve their respective goals. Some of the data needs for the two communities are common needs. Eugenia Kalnay and her group from U. MD have been working on developing more sophisticated methods of screening data for quality in the data assimilation process (Chen et al., 2015). Verification of weather forecasts and data assimilation both rely heavily on accurate data to achieve their respective goals. Some of the data needs for the two communities are in common and methods of addressing quality issues may take advantage of efforts from either community.

Given what is currently understood about observational accuracy, it makes sense to develop verification techniques that respect known problems with the data. Data issues can be divided into known, average uncertainties, and problems with individual observations. Both sets of issues will require some level of care to address. For known uncertainties, these uncertainties need to be compared with the biases observed in comparing observations to models. For individual observational problems, improved identification of potentially bad data needs to be adopted to assure verification results will be robust.

- 
- 
- Bian, Jianchun, Hongbin Chen, Holger Vömel, Yunjun Duan, Yuejian Xuan, and Daren Lü. "Intercomparison of humidity and temperature sensors: GTS1, Vaisala RS80, and CFH." *Advances in Atmospheric Sciences* 28, no. 1 (2011): 139-146.
- Chen, T. C., D. Hotta, and E. Kalnay. "The Performance and Feasibility of Ensemble Forecast Sensitivity to Observations-based Proactive Quality Control Scheme." In *AGU Fall Meeting Abstracts*. 2015.
- Feng, Rong, and Wansuo Duan. "IOD-related optimal initial errors and optimal precursors for IOD predictions from reanalysis data." *Science China Earth Sciences* 60, no. 1 (2017): 156-172.
- Galarus, Douglas E., Rafal A. Angryk, and John W. Sheppard. "Automated Weather Sensor Quality Control." In *FLAIRS Conference*, pp. 388-393. 2012.
- Hu, Junya, and Wansuo Duan. "The initial errors that induce a significant" spring predictability barrier" for El Niño events and their implications for target observation: results from an earth system model." In *EGU General Assembly Conference Abstracts*, vol. 18, p. 1409. 2016.
- Ingleby, Bruce, and Matt Huddleston. "Quality control of ocean temperature and salinity profiles—Historical and real-time data." *Journal of Marine Systems* 65, no. 1 (2007): 158-175.
- Kloog, Itai, Francesco Nordio, Brent A. Coull, and Joel Schwartz. "Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the Northeastern USA." *Remote Sensing of Environment* 150 (2014): 132-139.
- Schutgens, Nick AJ, Edward Gryspeerdt, Natalie Weigum, Svetlana Tsyro, Daisuke Goto, Michael Schulz, and Philip Stier. "Will a perfect model agree with perfect observations? The impact of spatial sampling." *Atmospheric Chemistry and Physics* 16, no. 10 (2016): 6335-6353.
- Sevruk, B., M. Ondrás, and B. Chvíla. "The WMO precipitation measurement intercomparisons." *Atmospheric Research* 92, no. 3 (2009): 376-380.



# CRITICAL COMPARISON AND EVALUATION OF SKILL SCORES IN SUPPORT OF NGGPS

Principal Investigator: Weatherhead, Elizabeth

Financial Report as of June, 2018

Start Date 2016-09-01 End Date: 2018-08-31

Fully funded. Sub to UCAR is project 1554500. Betsy has retired and chose to not name a different PI. CU will relinquish the award to NOAA.

NAT'L OCEANIC AND ATMOSPHERIC ADMINSTR

Award: NA16NWS4680016

Project: 1554499

Speedtype(s): 13009363

F&A Rate: 26%

Department: 10086

## To Date Actuals

Account Description	Total Funding	Monthly Actuals	To Date Actuals	% To Date	Future Expenses		Total Balance
					Salary, Travel, Operating	Other Projections	
400000 - 418399 -- SALARIES AND WAGES	\$139,025	\$0	\$101,240	72.8%	\$0	\$0	\$37,785
418400 - 439999 -- BEN	\$54,624	\$0	\$37,760	69.1%	\$0	\$0	\$16,864
450000 - 699999 -- OPERATING EXPENSES	\$10,813	\$0	\$8,542	79.0%	\$0	\$0	\$2,271
700000 - 709999 -- TRAVEL	\$24,509	\$0	\$20,313	82.9%	\$0	\$0	\$4,196
810000 - 812499 -- FIXED ASSET COSTS	\$5,000	\$0	\$5,157	103.1%	\$0	\$0	(\$157)
950000 - 952499 -- FACILITIES & ADMIN COSTS	\$59,532	\$0	\$43,642	73.3%	\$0	\$0	\$15,890
<b>Total</b>	<b>\$293,503</b>	<b>\$0</b>	<b>\$216,654</b>	<b>73.8%</b>	<b>\$0</b>	<b>\$0</b>	<b>\$76,849</b>

Anticipated available balance (for direct costs):

\$60,991

## Payroll Projections (12 months)

EmpID	Employee Name	Budget	To Date Payroll	7-18	8-18	9-18	10-18	11-18	12-18	1-19	2-19	3-19	4-19	5-19	6-19	Total	Remaining Balance	Benefit Cost
Speedtype: 13009363																		
108259	WEATHERHEAD, ELIZABETH C	\$139,025	\$101,240													\$0	\$37,785	
Speedtype: 99999999																		
	0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0	\$0
<b>Total</b>		<b>\$139,025</b>	<b>\$101,240</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$37,785</b>	<b>\$0</b>

## Travel Projections

EmpID	Employee Name	Final Destination	Speedtype	Start Date	End Date	Total
Total						\$0

## Other Projections

Budget Category	Description	Amount	Benefit	F&A	Total
		\$0	\$0	\$0	\$0
<b>Total</b>		<b>\$0</b>	<b>\$0</b>	<b>\$0</b>	<b>\$0</b>

## Original Proposal Summary

The work proposed is to develop the techniques to evaluate and support NWS decisions on NGGPS model development. The results of this work will help feed directly into all 11 areas of the NGGPS Verification and Validation team priorities, with a focus on development of both descriptive and decision support verification and validation techniques using state of the art statistical tools. These tools will be translated into general use capabilities by NCAR/DTC to assure that the tools will be available to all users of the MET software system.

The proposed work and capabilities development will assure that defensible analyses of the models are carried out with respect to the research to operations testing of the selected model development. Likely, by the time a decision is made on this proposal, the next generation global model will be chosen. **Problem:** once the global model is chosen, a large number of decisions will need to be made to develop this model so that it is ready for operational use. **Rationale:** for each of these decisions, the impact on the model's performance and a comparison to performance of competing models will need to be made. Some, but not all, decisions will come under intensive scrutiny because of potential negative impacts to some aspects of the forecast.

The tools developed within this proposal will help defend why specific decisions were made and how these decisions will likely affect the many aspects of the operational product suite.

**Brief Summary of Work:** five separate sub-projects have been addressed to develop, test and transfer to MET capabilities that can result in state of the art statistical verification tools for both atmospheric and non-atmospheric parameters. 1) Descriptive Forecast Verification; 2) Decision Support Statistics for Forecast Verification; 3) Key Metrics for Non Atmospheric Parameters; 4) Descriptive Verification and 5) Decision Support for Non Atmospheric Parameters. For each sub-project, documentation was produced in advance of the work to assure that the results will be useful to NWS' goals and needs.

**Relevance:** this proposal is in direct response to priority e) from the call for proposals: Advances in Verification Methods. The proposed work will add new capabilities (both descriptive and decision-support statistics) within EMC and the Model Evaluation Tools (MET) to allow defensible decisions to be made earlier, with efficient use of computer resources and be reproducible by the community. This proposal also supports the unification initiative through helping define critical measures still needed in MET to address the entirety of the NGGPS system. Particular attention will be paid how measures respond within climatologically diverse regions (i.e. mid-latitudes versus tropics) and which measures are appropriate for four non-atmospheric components (land, hydrology, waves and sea-ice). The most critical non-atmospheric measures will then be added to MET and training on best practices will be provided.

FINANCIALDETAILII

rn-Fin

ACCT CODE	BJP IO				SPDTYPE A CTUAS CODE PO	INV10	IN DATE	SI-C	IN*	JRNLO	NLOAH	FY	PD
13009363 - CRITICAL COMPARISON AND EVALUA													
4*1111-a11B-EIPENDmJIES													
4*CIO- RIPPAY													
4-0-0600	106259	WEATHERHEAD	HRPAIROI	JULNAL RANSACTION	61641 6	13019363	JulJU -017	11	G.694		Jul3U -017	"	1
4-0-0600	106259	WEATHERHEAD	HRPAYROLLJULNAL RANSACTION	G26426	13009363	P1J131.1-017	PAY 6.438	PAY1313775	Aug31. X17	"	2		
4-0-0600	106259	WEATHERHEAD	HRPAYROLLJULNAL RANSACTION	6.26426	13009363	5>p31 2017	PAY 6.276		51p29. X17	"	1		
4-0-0600	106259	WEATHERHEAD	HRPAYROLLJULNAL RANSACTION	1.589.55	13009363	Oct26, 2017	PAY 134	MYUJ111	Nw. 3 X17	"	S		
4-0-0600	106259	WEATHERHEAD	HRPAIROI	JULNAL RANSACTION	3,179.11	13009363	Nw 3Q, 2017	11 111	J.11111SZ	Nw 30, X17	"	S	
4-0-0600	106259	WEATHERHEAD	HRPAYROLLJULNAL RANSACTION	3,179.11	13009363			111					
4-0-0600	106259	WEATHERHEAD	HRPAYROLLJULNAL RANSACTION	9.7B	7.8	13009363	c31, 2017 J 31, 2018	11 111	111111S1	Det 29,X17 Jan31, X18	"	1	
4-0-0600	106259	WEATHERHEAD	HRPAIROI	JULNAL RANSACTION	6,994.04	13009363		PAY 6,468			"	7	
4-0-0600-11ASMCFT?PAY.S1:													
411GO-wfACFTPPAV													
401800	106259	WEATHERHEAD	HRPAYROLLJULNAL RANSACTION	SE9.S6	13009363	Oct31. 2017		11 111	11112W1	Oct31,2-017	"	4	
401800-11ASAC.mPAY, " 51													
41119-AU.FFACRINGI-EN													
11111													
4119W9		FACULTYFF RIN GEBELERTS			2,355.36	13010H63		BFG 110	11100J1111	Jun 1. X117	"	1	
		FACULTYFFRINGENERTS			2,355.36	13019363							
41192-09		FACULTYFF RIN GEBELERTS			2,355.36	130-09363		BFG 111	11111	Aug 1, 2-017	"	2	
41192-09		FACULTYFFRINGENERTS			1193.-02	13009363		11 111	11111	51p30 X17	"	1	
41192-09		FACULTYFF RIN GEBELERTS			1,195.35	1300H63		BFG 737	11111	Nw 30, X17	"	S	
41192-09		FACULTYFFRINGENERTS			3,8604	1300H63		BFG 729	11100J1111	Det 31. X17	"	1	
41192-09		FACULTYFFRINGENERTS			2.6	13009363		BFG 674	11111	Jan31, X118	"	7	
41GGI-AU.ITFM:FRIPG- 11111													
41221No-11111:11111:11111													
422176		FACULTYFF RIN GEBELERTS			235.1 6	13009363		BFG 209		Oct31, X17	<018	4	
G211-11111:11111:11111													
48762 - 11111:11111:11111													
487W2	106259	WEATHERHEAD	WEATHERHEAD	11111	65.00	13009363	CL(SRB400	c19, 2017	MATHEEA-001	enE	<01	Jan16 X118	<011 7
MKT/PUB/REPRO SERVICES (487)402 11111													
CARVALHOSEPT- 11111:11111:11111													
PRO SERVICES(437602)11111:11111:11111													
2.4 65.00 13009363													
11111													
D c 6, 2017 <018 6													

1,111 ;?0lew,cgg,cu.W

■■■■

rn.rlnr JW..(I\*(O( ALHI0-1..)SF'M

[illegible]

FY PERIOD 2016 THROUGH 2019 12 (Jun 30, 2019)  
 SPEEDTYPE STATUS A SPEEDTYPE 13009963 - CRITICAL COMPARISON AND EVALUA (ACTIVE)

ACCT CODE	EMPLID	EMP	DESC	ACTUALS	SPTYPE CODE	PO	INVID	INVD DATE	VNDR	SRC	LN#	JRNLID	JRNL DATE	FY	PD
950100			HR PAYROLL JOURNAL TRANSACTION	826.57	13009963					GM	3.004	ENA1336286	Nov 30 2017	2018	5
950100			FACULTY FT FRINGE BENEFITS	466.19	13009963					GM	4.836	ENA1338211	Nov 30 2017	2018	5
950100			WEATHERHEAD AMS PUBLICATION	640.00	13009963					GM	454	ENA1339127	Dec 6 2017	2018	5
950100			WEATHERHEAD PUBLICATION CEN-C	208.00	13009963					GM	496	ENA1339132	Dec 6 2017	2018	5
950100			WEATHERHEAD VALLEY PUBLICATION	468.00	13009963					GM	620	ENA1339127	Dec 6 2017	2018	5
950100			FACULTY FT FRINGE BENEFITS	310.75	13009963					GM	856	ENA1344070	Dec 31 2017	2018	5
950100			HR PAYROLL JOURNAL TRANSACTION	826.57	13009963					GM	6.499	ENA1343431	Dec 31 2017	2018	5
950100			WEB-BASED SOFTWARE TO ALLOW CO	90.48	13009963					GM	523	ENA1345534	Jan 9 2018	2018	7
950100			WEATHERHEAD AGU 1/3/15/17	703.60	13009963					GM	531	ENA1346851	Jan 16 2018	2018	7
950100			HR PAYROLL JOURNAL TRANSACTION	2,479.71	13009963					GM	275	ENA1349446	Jan 31 2018	2018	7
950100			WEATHERHEAD AUSTIN 1/13/18	569.54	13009963					GM	2,240	ENA1349446	Jan 31 2018	2018	7
950100			FACULTY FT FRINGE BENEFITS	952.27	13009963					GM	3,527	ENA1351612	Jan 31 2018	2018	7
950100			WEATHERHEAD AUSTIN 1/5/2018	73.57	13009963					GM	536	ENA1352696	Feb 8 2018	2018	8
950100			HR PAYROLL JOURNAL TRANSACTION	1,018.45	13009963					GM	2,306	ENA1356193	Feb 28 2018	2018	8
950100			FACULTY FT FRINGE BENEFITS	683.74	13009963					GM	4,450	ENA1358526	Feb 28 2018	2018	8
950100 - FACILITIES & ADMIN EXPENSE				19,843.04											
400000 - DEPENDENT				96,162.32											
13009963 - CRITICAL COMPARISON AND EVALUA				96,162.32											
TOTAL				96,162.32											