NGGPS: update on ensemble prediction

Tom Hamill (ESRL / PSD) and Yuejian Zhu (NCEP / EMC)

Ensemble team objective

"Produce more skillful, reliable, and detailed probabilistic guidance of high-impact weather variables by developing and rapidly implementing improved ensemble prediction methods."

High-level strategy for ensemble development

- Improve methods for defining initial conditions in ensemble prediction systems, including for the coupled state (in conjunction with NGGPS data assimilation team).
- Develop methods to accurately quantify model uncertainty in ensemble prediction systems, including coupled systems (including NGGPS parameterization team).
- Develop system improvements that will generate reliable and maximally skillful guidance to lead times of +30 days and beyond (leverages NGGPS coupled model development).

Focus area 1: improving methods for defining initial conditions in (coupled) ensemble prediction systems.

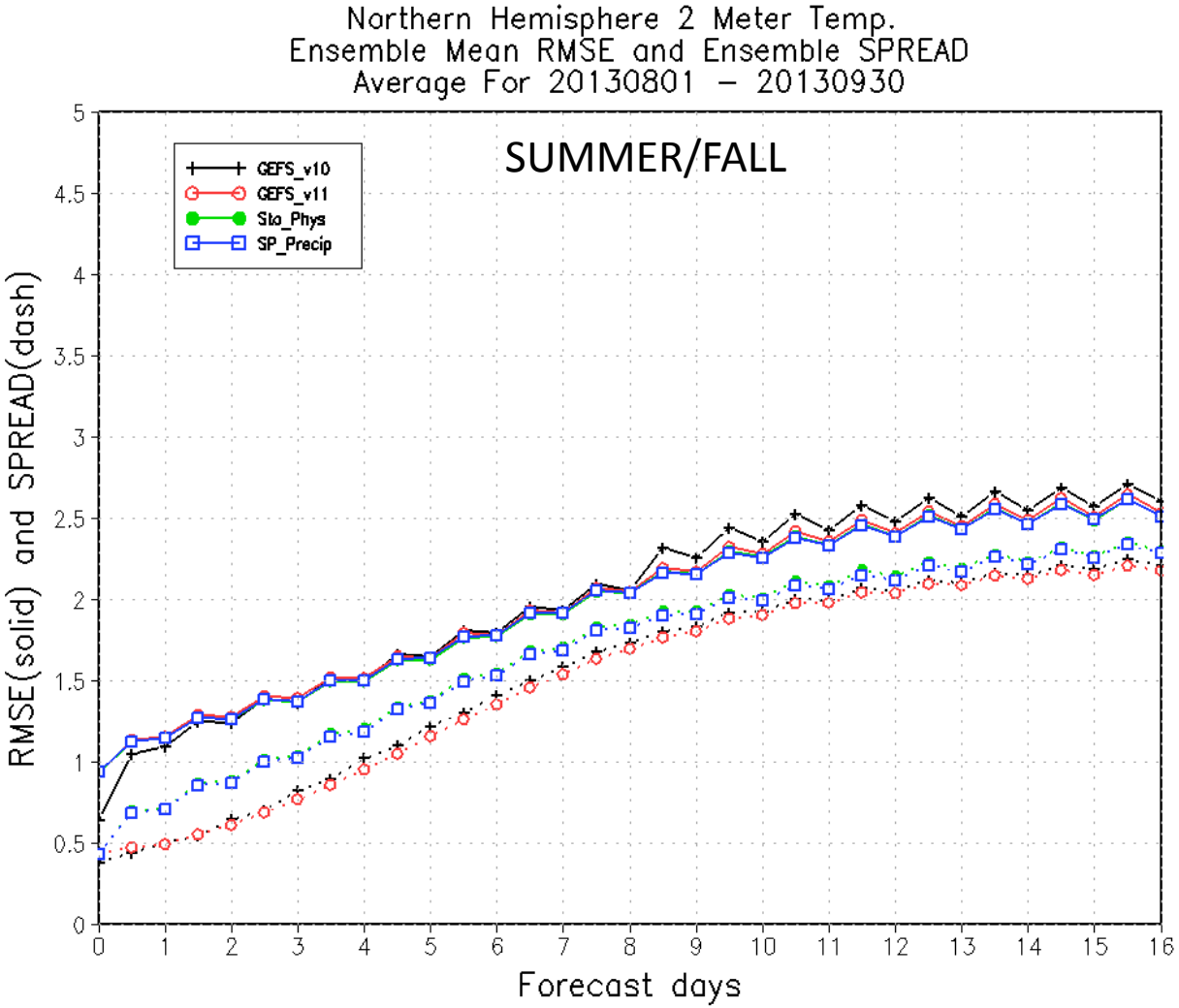
- Improved methods for initializing ensembles of the atmospheric state (largely performed via NGGPS DA team, 4D-En-Var development).
- Initializing land, ocean, sea-ice ensembles:
  - Some early work performed under project for land-surface uncertainty estimation performed by ESRL/PSD (described later).
  - Concerted effort on this should wait till coupled NGGPS system is in place.
  - Physically based coupled initialization (coupled DA) is a frontier research area, suitable for NGGPS grants (once coupled system in place) and OAR research.

Focus area 2: Develop methods to accurately quantify model uncertainty in ensemble prediction systems, including coupled systems.

- Early work (via NWS Sandy Supp., HIWPP, NGGPS): new suite of stochastic parameterizations (ESRL/PSD and EMC; see upcoming slides). Expect this to be in GEFS v.12.
- Early work on land-atmosphere interactions and uncertainty at ESRL/PSD. Shows small but positive impact.
- Follow-on work: physically based stochastic parameterizations (with NGGPS parameterization team) for:
  - deep convection.
  - microphysics.
  - dry dynamics.
  - other parameterizations, TBD.
- Model uncertainty in coupled state:
  - Pre-requisite: NGGPS coupled system in place.
  - Could have bridge technology of using CFS v2 SSTs, LIM forecasts.
  - Basic research needed, suitable for NGGPS external grants (once coupled system in place) and OAR research.

Some results for new suite of stochastic parameterizations

- Stochastically perturbed physical tendencies (SPPT), stochastic kinetic-energy backscatter (SKEB) and stochastic boundary layer RH (SHUM) now extensively tested between ESRL/PSD and EMC.
- Major improvements over GEFS v. 11, especially in tropics and summer hemisphere.

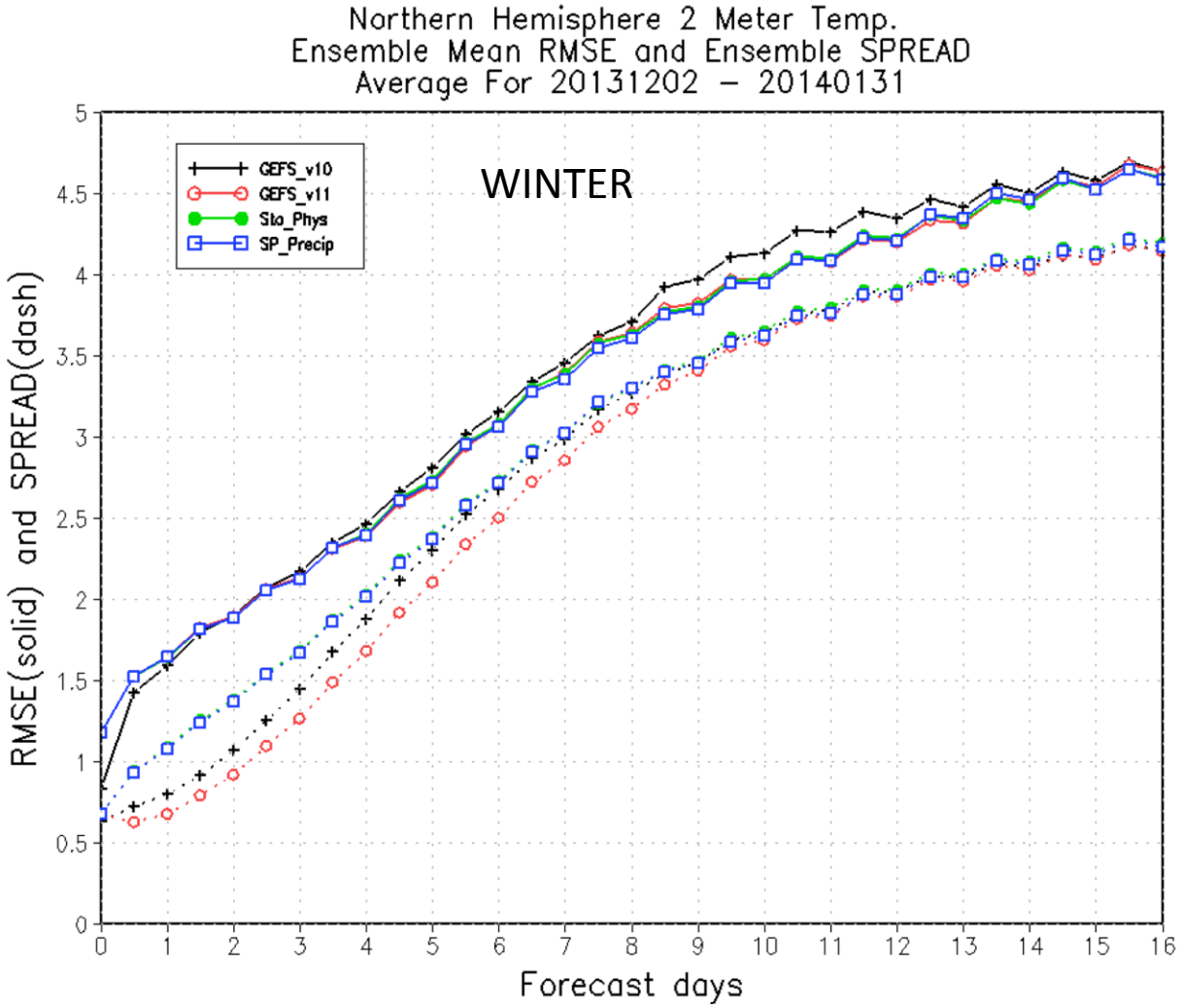


At left: RMS error (solid curves) and spread (dashed) for four configurations, GEFS v. 10 (2012-2015), GEFS v. 11 (2015-current), new stochastic physics without closure of precipitation budget, new stochastic physics including closure of precipitation budget.

While there is still a lack of spread in 2-meter temperatures, there is a notable improvement from the suite of new stochastic physics.

Some results for new suite of stochastic parameterizations

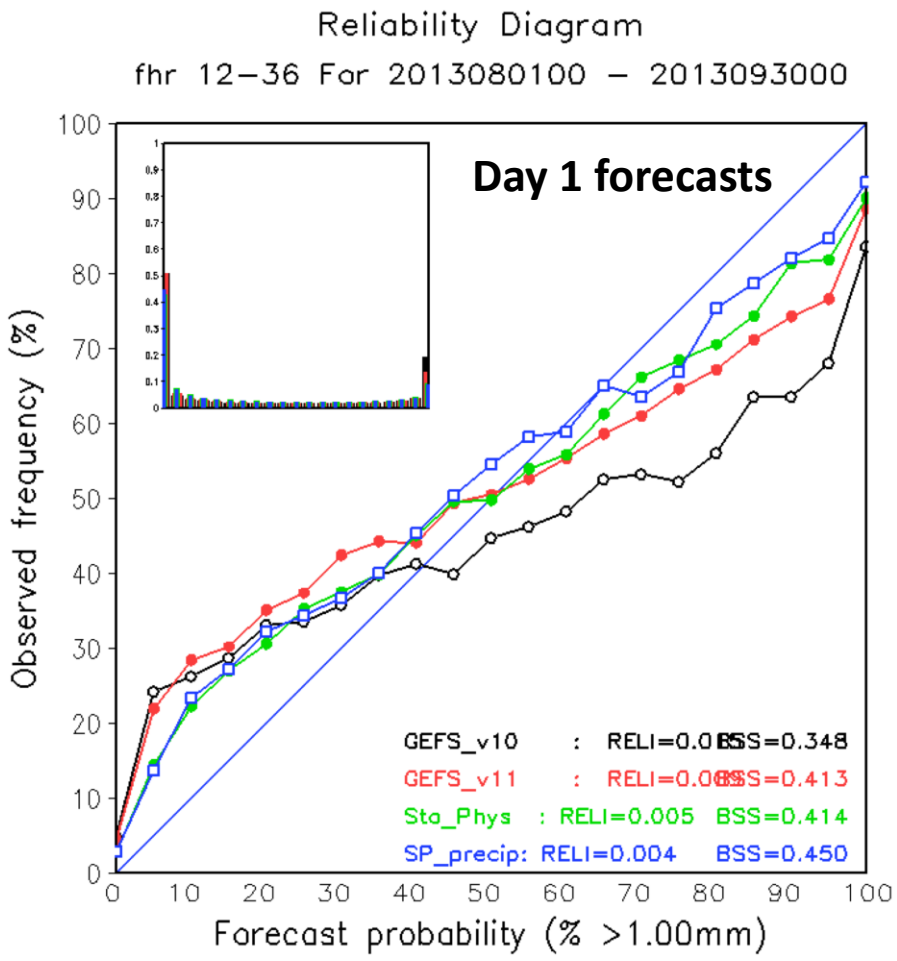
- Stochastically perturbed physical tendencies (SPPT), stochastic kinetic-energy backscatter (SKEB) and stochastic boundary layer RH (SHUM) now extensively tested between ESRL/PSD and EMC.
- Major improvements over GEFS v. 11, especially in tropics and summer hemisphere.

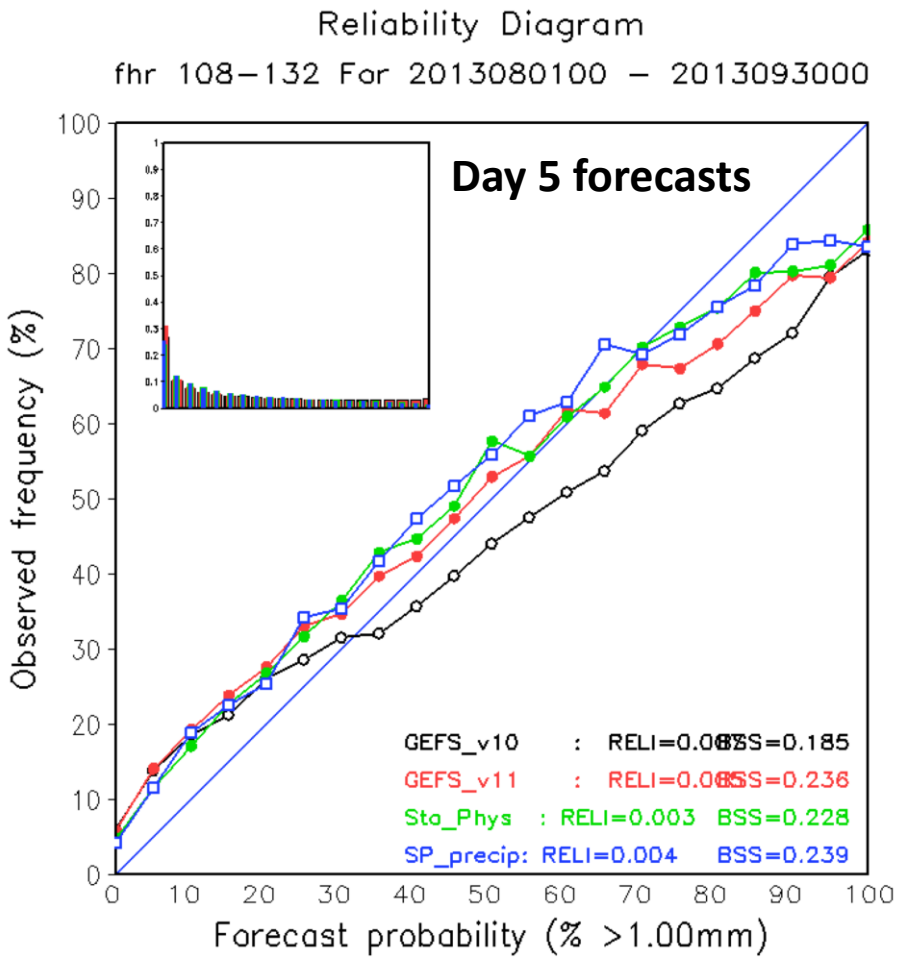


At left: RMS error (solid curves) and spread (dashed) for four configurations, GEFS v. 10 (2012-2015), GEFS v. 11 (2015-current), new stochastic physics without closure of precipitation budget, new stochastic physics including closure of precipation budget.

While there is still a lack of spread in 2-meter temperatures, there is a notable improvement from the suite of new stochastic physics.

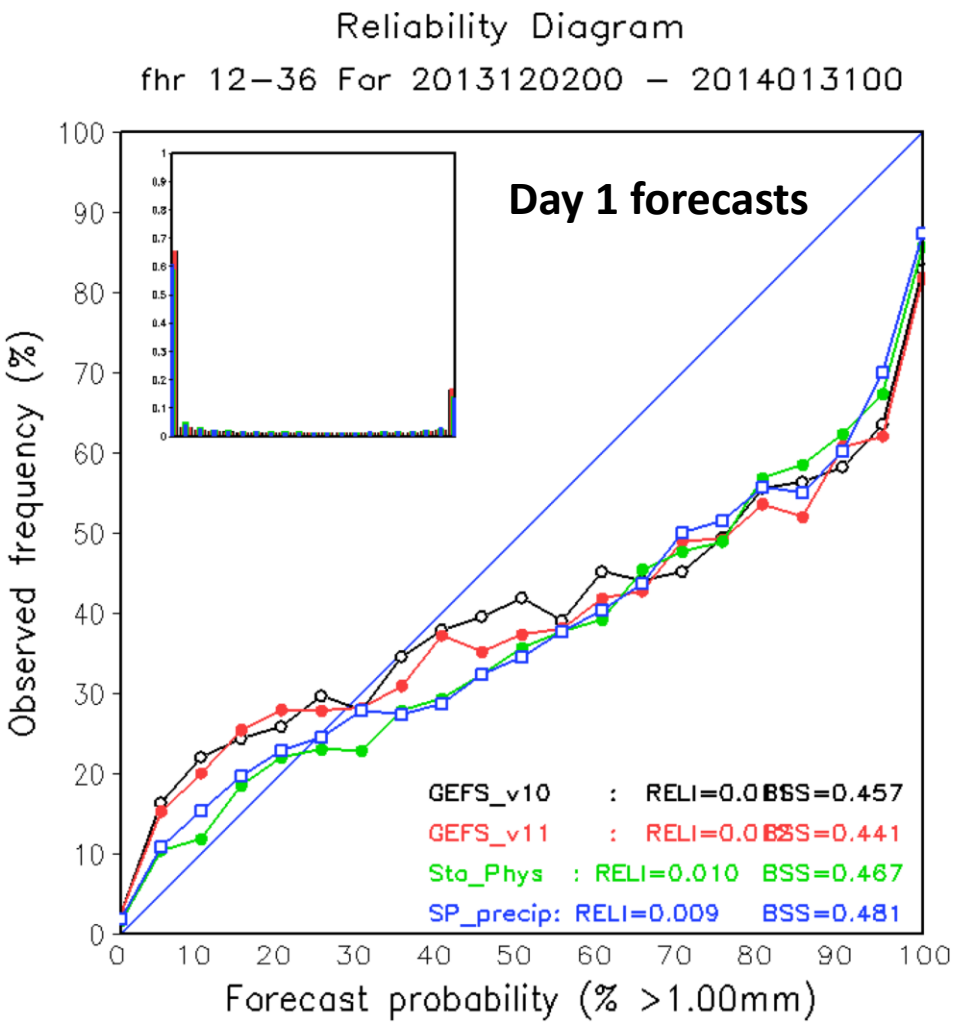
Precipitation reliability for new stochastic parameterizations.

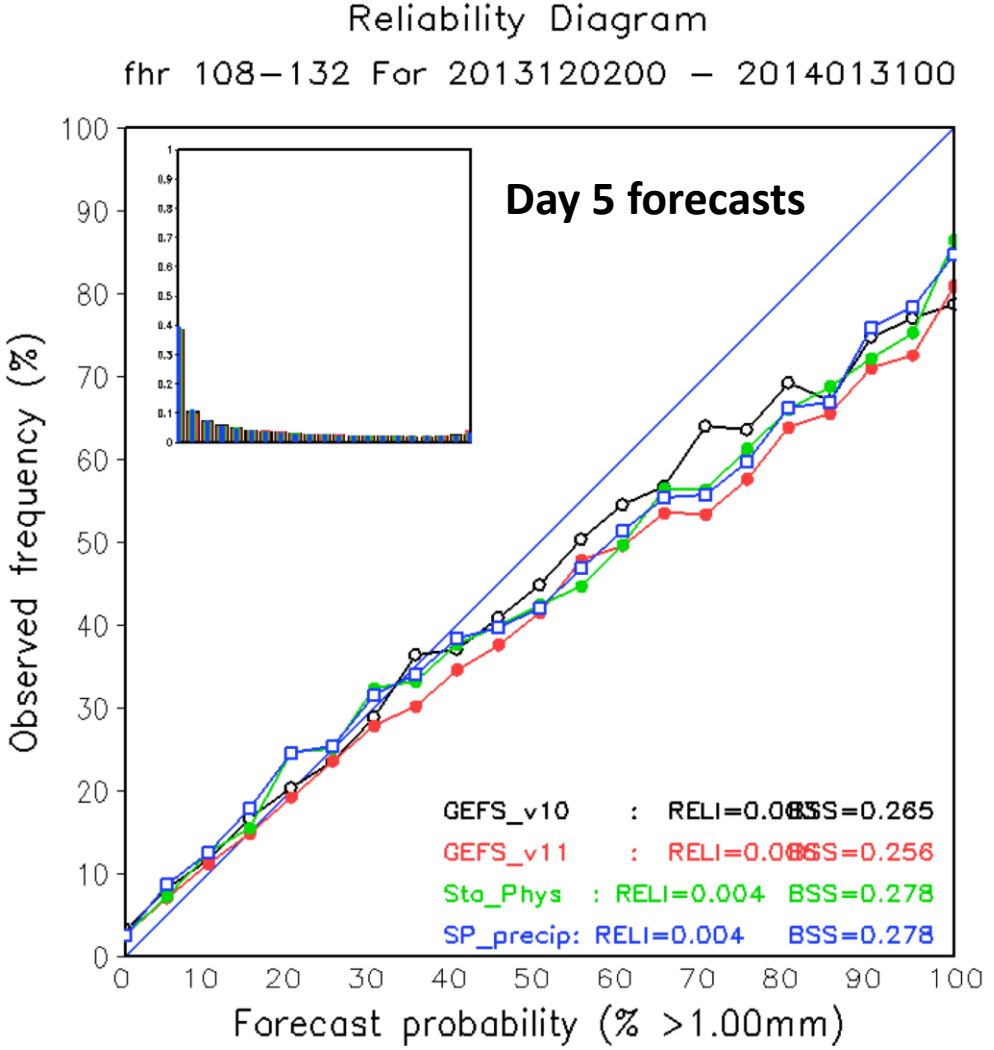




Here, for > 1 mm event and **summer/fall** testing, there is substantially improved reliability and skill with new suite of stochastic parameterizations. Tested against CCPA analyses at 1-degree resolution over CONUS; less reliability when verifying against finer-grid analyses, over shorter time periods.

Precipitation reliability for new stochastic parameterizations.

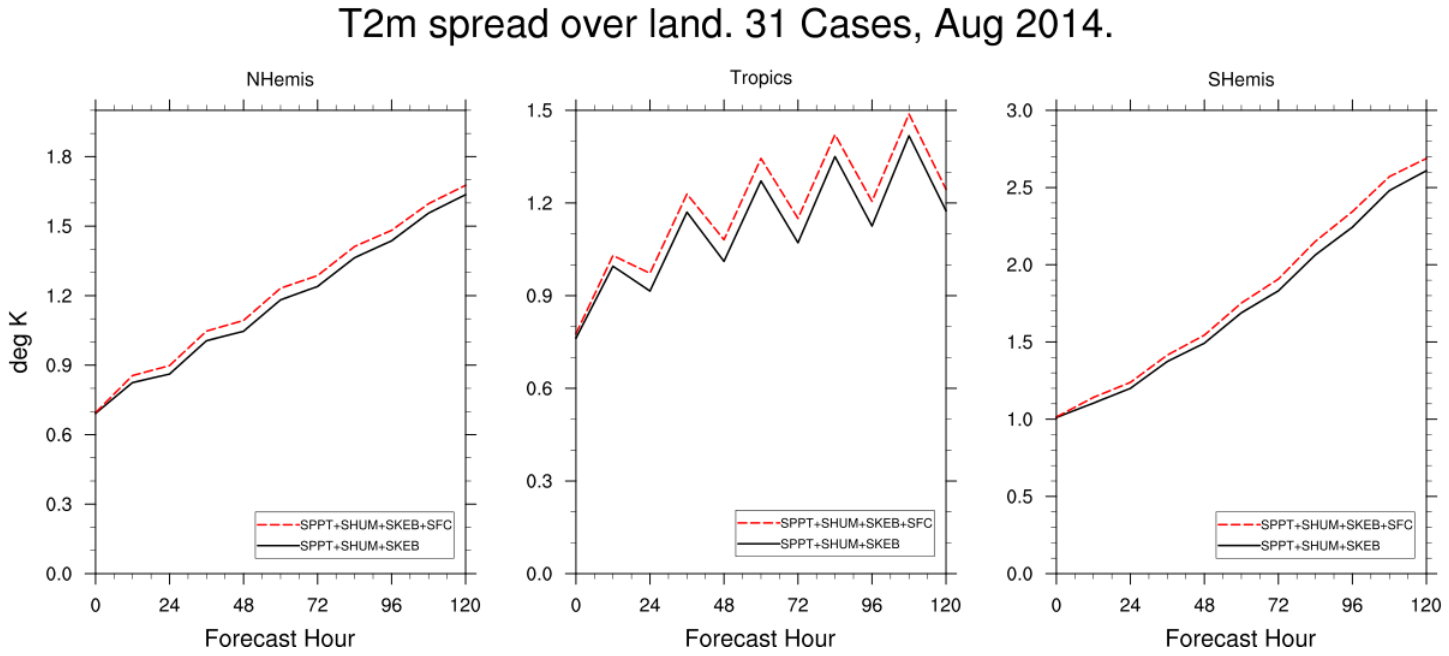




Here, for > 1 mm event and **winter** testing, there is slightly improved reliability and skill with new suite of stochastic parameterizations.

Modeling land-atmosphere uncertainty

- ESRL/PSD (with EMC) has developed methods for perturbing the land-surface parameters to which there is forecast sensitivity (albedo, surface roughness, leaf area index, soil hydraulic conductivity, etc.).
  - Small positive impact on near-surface temperature spread (big impact wasn't expected).



- ESRL/PSD currently working on methodology for perturbing soil moisture (possible to include soil temperature) consistent with uncertainty as determined from differences between NOAH and MOSAIC soil moisture analysis differences.
- Also, EMC experimenting to perturb soil temperature and soil moisture directly.

Focus area 3: ensemble systems to +30 days and beyond

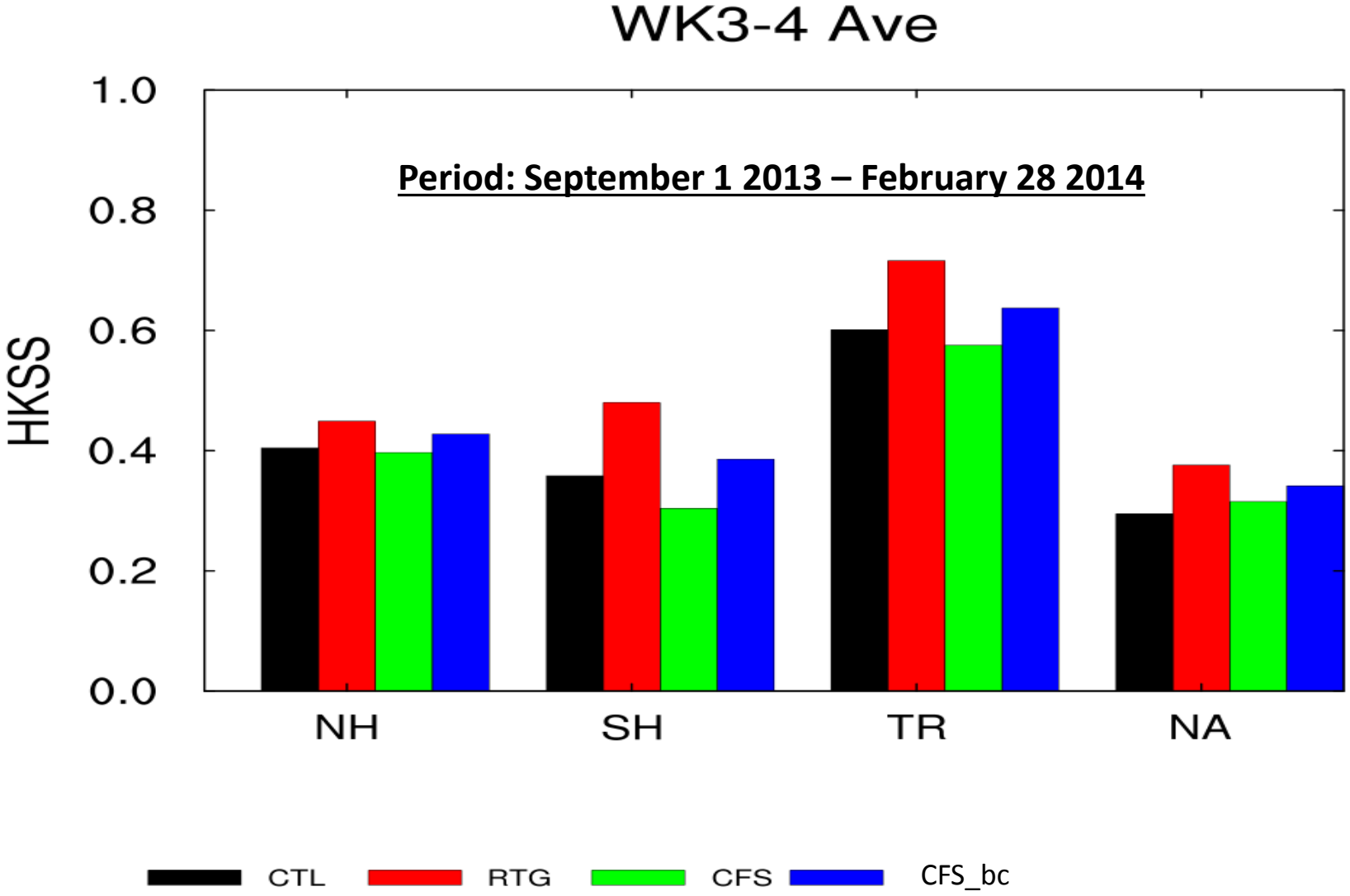
- Short-term:
  - What do we do about effect of changes in ocean state over 30 days?
    - Coupled NGGPS system still in development.
    - Explore transplantation of bias-corrected CFS v2 SSTs (EMC).
    - Explore Linear-Inverse Model (LIM) statistical forecasts of tropical SSTs (unfunded, but preliminary testing in ESRL/PSD via other funds). EMC (with PSD) is planning to set up experiment to compare LIM to exist approach.
- Longer term: explorations with NGGPS coupled system.
  - We are likely to learn that there are challenges in modeling the uncertainty that one state component has on the other.
  - Again, more basic research in character, suitable for more grants.

Exploring GEFS extension forecast to 35 days

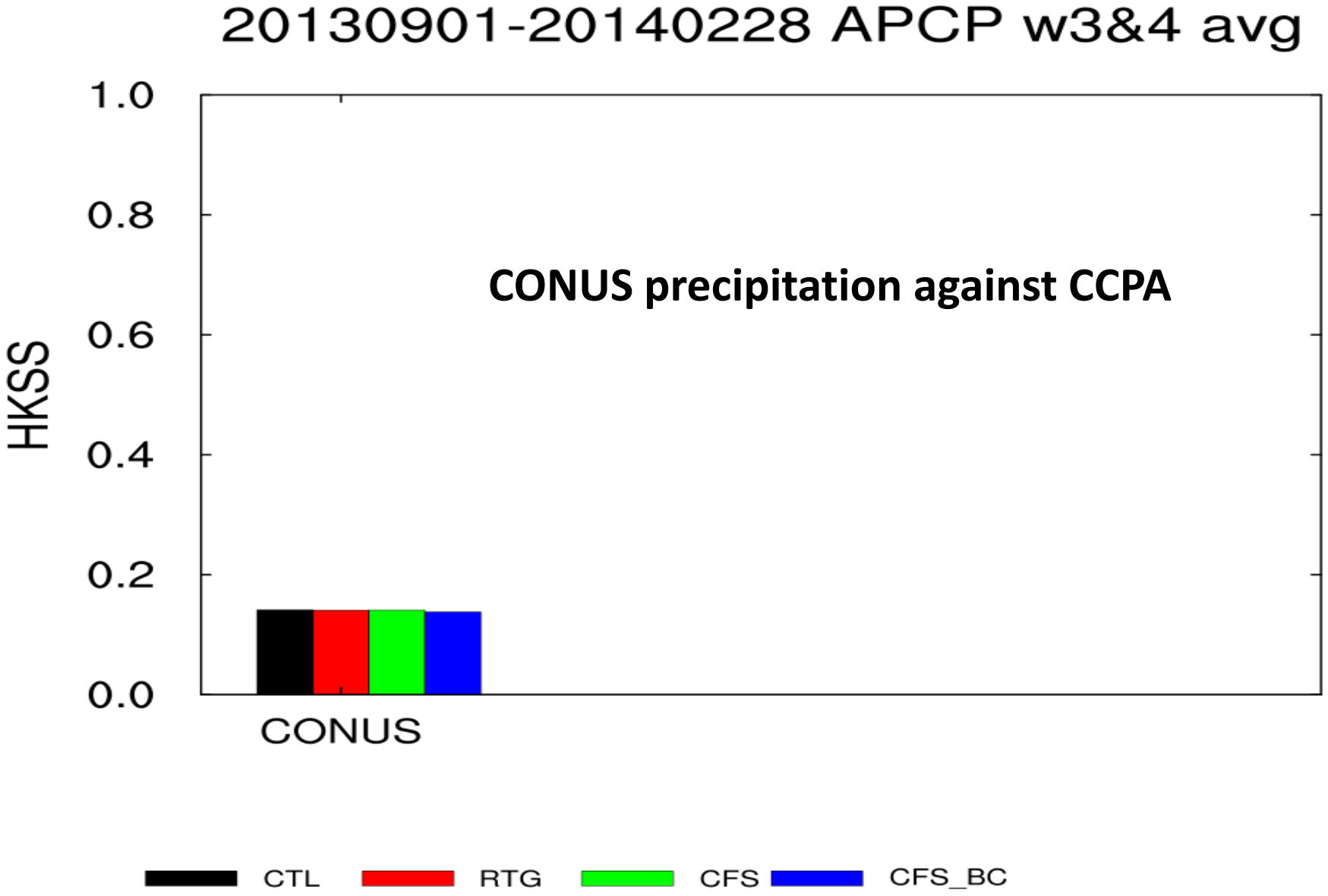
(using GEFSv11)

- NCEP GEFS v11, based on GFS (version 12.0.0 - 2014) used for this study.
- Experiment - extended 2013-2014 winter season (Sept. 1 2013 – Feb. 28 2014).
- Four(or more) experiments studied to date:
  - Control (CTL): analysis SST relaxes to climatology
  - Optimum (RTG): realistic SST forcing every 24 hours (AMIP like)
  - Forcing (CFS): CFSv2 predicted SST forcing every 24 hours
  - Forcing (CFS_bc): CFSv2 bias corrected predicted SST anomaly
- Preliminary evaluations for raw forecasts (next few slides)

Combined Heidke Skill Score for T2m (WK3&4)



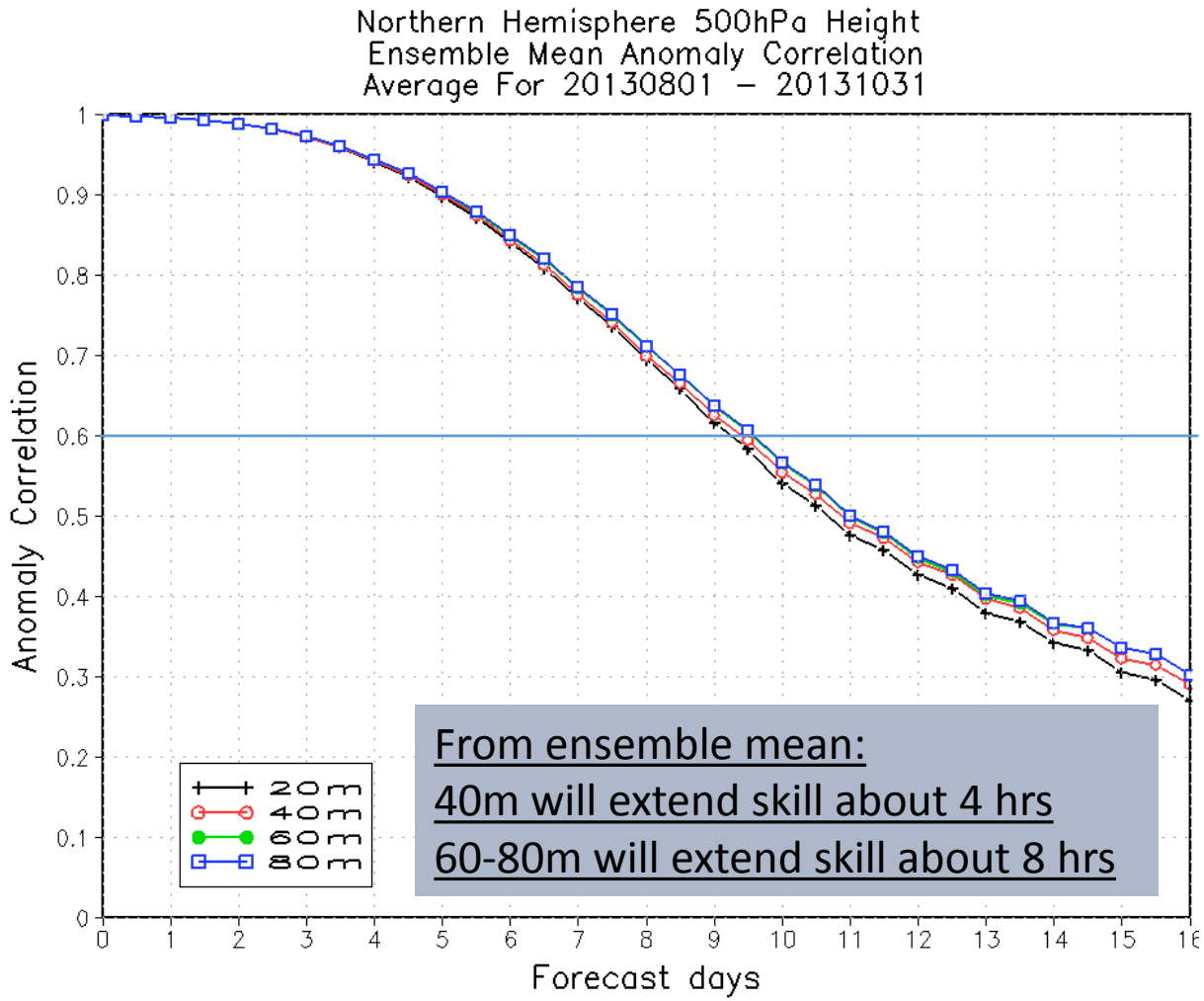
Courtesy of Drs. Li and Melhauser

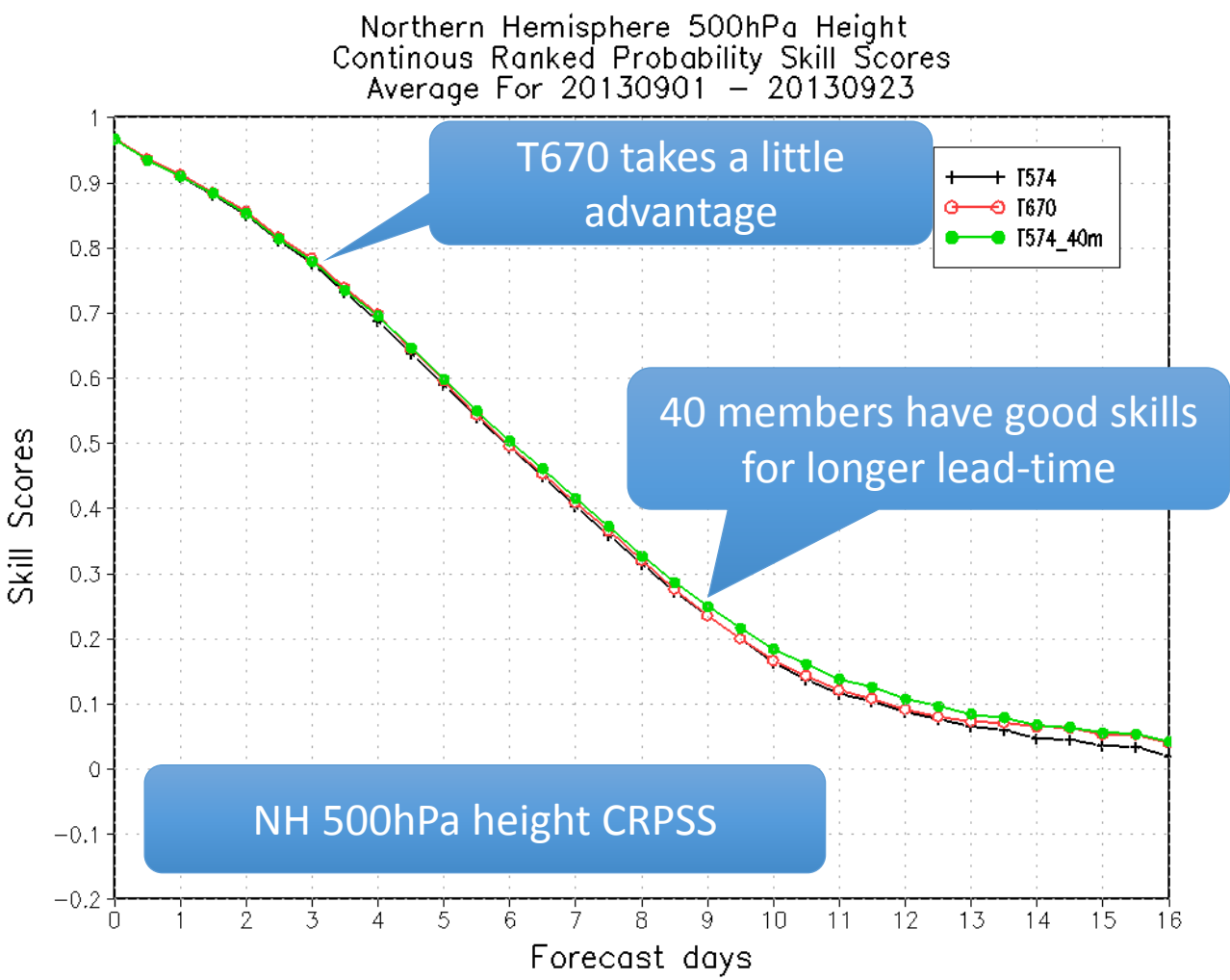

20130901-20140228 APCP w3&4 avg
CONUS precipitation against CCPA
HKSS
1.0
0.8
0.6
0.4
0.2
0.0
CONUS
CTL
RTG
CFS
CFS_BC

Ongoing verification of these experiments

- Surface temperature
  - Against observations
  - Bias correction from GEFS reforecast and CFS hindcast
- Precipitation
  - Bias correction from GEFS reforecast and CFS hindcast

Ensemble configuration test





Scores Card: GEFSv11 21m .vs 41m (August 1 – October 1 2013)

Against NCEP analysis

EMC Global Ensemble Verification Scorecard
Symbol Legend
40m is better than 20m at the two-tailed 95% significance level
No statistically significant difference between 40m and 20m
40m is worse than 20m at the two-tailed 95% significance level
Not statistically relevant
Bootstrap Significant Test
Start Date: 20130801
End Date: 20131031

Green: significant better (95%)

Pink: significant worse (95%)

Grey: In-significant or neutral

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 8	Day 12	Day 16	Day 1	Day 3	Day 5	Day 8	Day 12	Day 16	Day 1	Day 3	Day 5	Day 8	Day 12	Day 16	Day 1	Day 3	Day 5	Day 8	Day 12	Day 16
Anomaly Correlation	Heights	500hPa																								
		1000hPa																								
	Temp	850hPa																								
		2m																								
	U-Wind	250hPa																								
		850hPa																								
		10m																								
	V-Wind	250hPa																								
		850hPa																								
		10m																								
RMSE	Heights	500hPa																								
		1000hPa																								
	Temp	850hPa																								
		2m																								
	U-Wind	250hPa																								
		850hPa																								
		10m																								
	V-Wind	250hPa																								
		850hPa																								
		10m																								
Bias	Heights	500hPa																								
		1000hPa																								
	Temp	850hPa																								
		2m																								
	U-Wind	250hPa																								
		850hPa																								
		10m																								
	V-Wind	250hPa																								
		850hPa																								
		10m																								
CRPSS	Heights	500hPa																								
		1000hPa																								
	Temp	850hPa																								
		2m																								
	U-Wind	250hPa																								
		850hPa																								
		10m																								
	V-Wind	250hPa																								
		850hPa																								
		10m																								
BSS	Heights	500hPa																								
		1000hPa																								
	Temp	850hPa																								
		2m																								
	U-Wind	250hPa																								
		850hPa																								
		10m																								
	V-Wind	250hPa																								
		850hPa																								
		10m																								

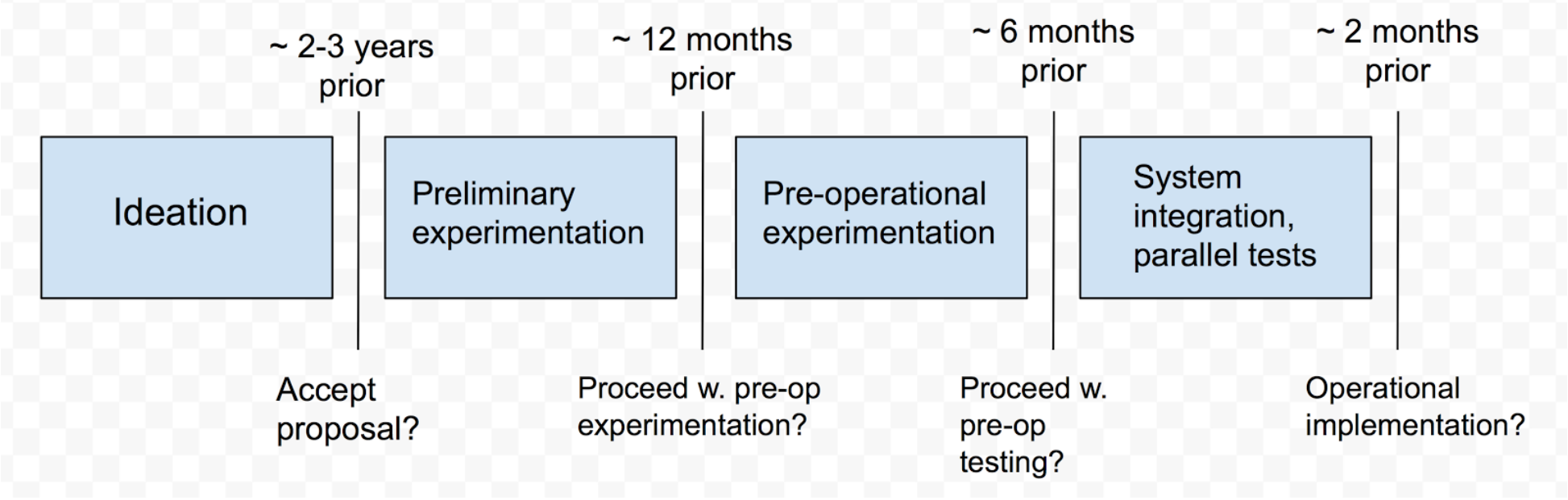
Additional planning and implementation concerns

- Full NGGPS ensemble team hasn't been used much since initial draft of plan.
  - Need to update ensemble team members.
- Per UMAC, there is a need for evidence-based decision making in the ensemble prediction component of NGGPS as much as for more visible components (dycore selection).
- Pending NGGPS and EMC management approval, we propose to draft an ensemble team terms of operation (TOO) that defines roles for team in:
  - keeping the NGGPS ensemble plan up to date, consistent with UMAC recs, and recommending developmental priorities.
  - reviewing proposals, providing guidance on which projects to fund.
  - monitoring and evaluating NGGPS ensemble project development.
  - making recommendations in a stage-gate development process (proceed, stop work, modify plan).

A proposed “stage-gate” development process.

- Current ensemble development process is perhaps too informal; test plans not always fully thought out, peer review of results is sporadic, tendency to soldier on even if early results not promising.
- We need an implementation process that is as orderly and speedy as is possible (with constraints such as producing reforecasts). How about the following stage-gate process, overseen by NGGPS ensemble team?

A conceptual model of the development process with R&D stages (blue) and gate (reviews; vertical black lines).



Arguably, many NGGPS teams might adopt this.

NOAA-internal suggested FY17 and beyond milestones and deliverables.

- Early FY17: nail down configuration of GEFS v12 and reanalysis in advance of reanalysis/reforecast production.
  - Possible/likely changes to GEFS system include new suite of stochastic parameterization, NEMS architecture, inheritance of GFS physics changes, model resolution changes, increasing ensemble members, land-surface uncertainty methodology, treatment of SST evolution.
  - Decide what variables to save from reforecasts to meet various user needs.
  - Decide on reanalysis system configuration (streams, resolution, ensemble size, etc.).
- FY17-18: Reanalysis and then reforecast production.
- FY18 and beyond: adaptation of GEFS to new dycore. Development of coupled EPS and treatment of coupled uncertainty in DA and ensemble forecast.

FY17 and beyond milestones and deliverables.

- Early FY17: nail down configuration of GEFS v12 and reanalysis in advance of reanalysis/reforecast production.
  - Possible/likely changes to GEFS system include new suite of stochastic parameterization, NEMS architecture, inheritance of GFS physics changes, model resolution changes, increasing ensemble members, land-surface uncertainty methodology, treatment of SST evolution.
  - Decide what variables to save from reforecasts to meet various user needs.
  - Decide on reanalysis system configuration (streams, resolution, ensemble size, etc.).
- FY17-18: Reanalysis and then reforecast production.
- FY18 and beyond: adaptation of GEFS to new dycore. Development of coupled EPS and treatment of coupled uncertainty in DA and ensemble forecast.

External projects.

- **Accelerating coupled NGGPS developed for predicting week 3 and 4** PI: James Kinter, Center for Ocean-Land-Atmosphere Studies (COLA), George Mason University
- **Development and testing of a multi-model ensemble prediction system for sub-monthly forecasts**. PI: Andy Robertson, IRI/Columbia.
- **Calibration and Evaluation of GEFS Ensemble Forecasts at Weeks 2-4** PI: Ping Liu, SUNY- Stony Brook.

NGGPS COLA Team Summary

- **Major Accomplishment in FY16:**
  - Extensive reforecast analysis (S2S, CFSv2, modified CFSv2 with HCF, EDMF)
  - Tests with HCF convection trigger at ~100-km and ~35-km grids show positive impacts on forecast skill for rainy season onset, total monsoon accumulation and tropical cyclones
  - Installed SP in CFSv2
  - Tested multiple ocean analysis efficacy for coupled prediction
  - Contributed to UGCS design and development
- **Priority Focus for FY17**
  - Shift to UGCS for future development and testing
  - Apply advanced statistical methods to evaluate skill and predictability
- **Key Issue**
  - Readiness of UGCS for scientific evaluation
  - High-end computing resources

Andy Robertson - NGGPS Columbia University

- **Major Accomplishment in FY16:**
  - Archived subset of S2S forecast/re-forecast database in IRI Data Library (ECMWF, NCEP, CMA model subsets).
  - Evaluated ECMWF and NCEP week 1–4 re-forecast performance in weekly-averaged precip, 500hPa geopotential and 2m temperature (anomaly correlation, RME error).
  - Developed extended logistic regression calibration of sub-seasonal precip forecasts for ECMWF, NCEP and CMA models, together with a multi-model combination.

- **Priority Focus for FY17**
  - Improving extended logistic regression model for week 3-4 precipitation forecasts.
  - Diagnostics of ECMWF vs NCEP performance differences over the U.S.
  - Tailoring post-processing codes for NCEP usage.

- **Key Issue**
  - Developing well-calibrated sub-seasonal (week 3-4) probabilistic forecasts over the U.S.

Relevance to the priorities of R2O Initiative

- Area 2b. Service impacts: Weeks 3–4 forecast development
- Area 5. Advances in ensemble development.
- Area 6. Advances in post-processing

Planned Deliverables

- A peer-reviewed manuscript documenting individual S2S model performance (includingCFSv2) over the U.S., with emphasis on weekly averages in weeks 2–4, gridded fields of precipitation and temperature, atmospheric indices such as the NAO and PNA, and diagnostics of predictability.
- A peer-reviewed manuscript describing the development and testing of the new MME methodology for sub-monthly forecasts.
- Implementation of a real-time S2S MME at CPC, built using the most skillful and models that are available to CPC in real time.
- Maintenance and updating of the North American Multi-Model Ensemble (NMME) archive of seasonal forecast model data in the IRI Data Library.

Ping Liu - NGGPS Stony Brook University

- **Major Accomplishment in FY16:**
  - Revised the Real-time Multi-variate MJO index (RMM-r; Liu et al. 2016; MWR).
  - Constructed a web server updating the RMM-r (http://mjo.somas.stonybrook.edu).
  - Developed a new objective algorithm for identifying and tracking persistent maxima of 500-hPa geopotential height (PMZ) including both ridges and blockings (to MWR).
  - Verified the predictability of PMZs in the GEFS V10 Reforecasts from 1985-2015.
  - PMZ events can extend the useful AC skill by more than a day after an EOF approach.
  - Constructed a web server for monitoring the PMZ events over North America and for their predictions by the GEFS (http://mjo.somas.stonybrook.edu/PMZ).

- **Priority Focus for FY17:**
  - Predictability of PMZs and MJOs (RMM-r) impacting the US in the GEFS V10 and V11.
  - Predictability of precipitation and T_{2m} over CONUS associated with PMZs and MJOs.
  - Application of the EOF approach to extend the weeks 2-4 predictability of CONUS precipitation and T_{2m} associated with PMZs and MJOs.

- **Key Issue:**
  - Week 3-4 predictability of CONUS precipitation and T_{2m} associated with PMZs and MJOs because of lacking sufficient reforecasts.

NGGPS ensemble team summary

- **Major Accomplishments in FY16:**
  - Refinement of suite of stochastic parameterizations for GEFS v12.
  - Development of GEFS 35d forecast with evolving SSTs
  - Testing different resolutions, ensemble size for GEFSv12 configuration
  - Other GEFS v12 development (land-surface uncertainty).
- **Priority Focus for FY17:**
  - Finalizing GEFS v12 configuration.
  - Reanalysis/reforecast production
- **Key Issues:**
  - Per UMAC recommendation, setting up a more orderly ensemble development process with stages and gates, overseen by a review team (NGGPS ensemble team is suggested).
  - Adaptation to new NGGPS dycore and coupled system and beginning the exploration of coupled ensemble DA and forecast issues.
  - There is no enough resource (include funding) to support ensemble development, especially for sub-season prediction

Supplementary slides.

Improved Statistical Post-processing with the **Bayesian Processor of Ensemble (BPE)**

Zoltan Toth (GSD/ESRL), Mark Antolik (MDL/NWS), Roman Krzysztofowicz (Univ. Va), Melissa Petty (CIRA at GSD), Geary Layne (CIRES at GSD), and Malaquias Pena (IMSG at EMC)

- Fuse climatology & multiple forecast guidance
  - Prototype for probabilistic National Blend of Models
- Bayesian & distributional approach
  - Extremes handled well with smaller set of hind-casts
  - All continuous variables calibrated with unified method
- Multiple & versatile output formats
  - Probabilistic NDFD w only factor of 2 storage increase
- Algorithms & codes to be developed & tested
  - Comparison with operational EKDMOS

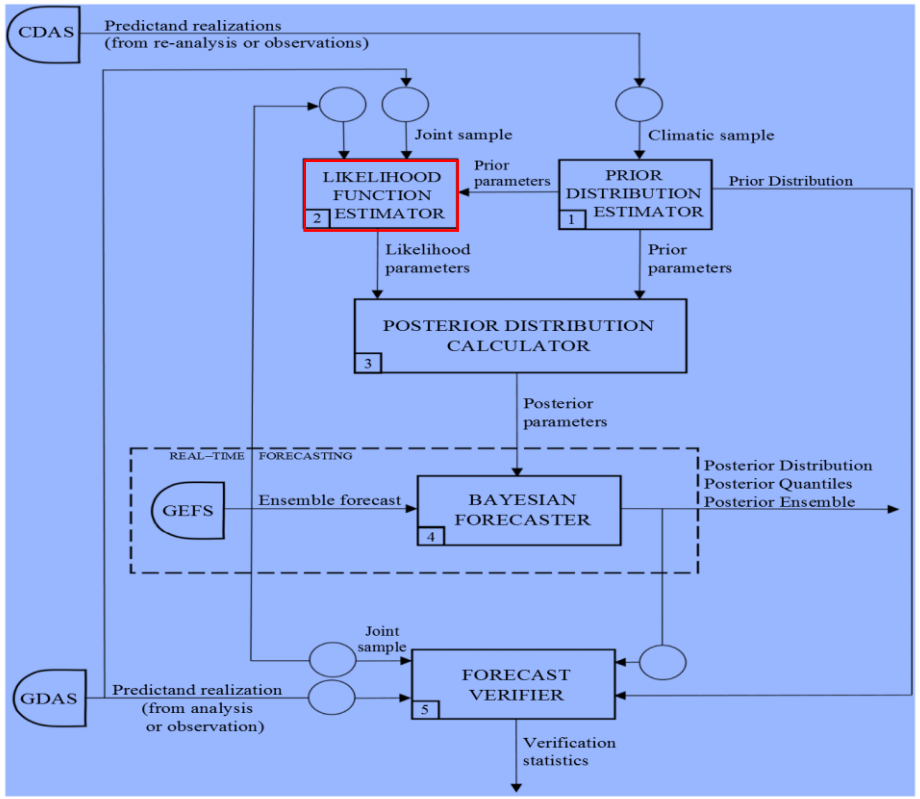


Figure 1. Structure of the BPE — its five components and its couplings with systems supplying data, illustrated for NCEP: Global Ensemble Forecast System (GEFS), Global Data Analysis System (GDAS), and Climate Data Analysis

Contributions to accomplishing major NGGPS goals

- "Reliable weather, water… information"
- "Improv[e] the probabilistic guidance of high impact weather elements by
  - Decreasing… uncertainty and improving product calibration"
- "Cutting edge techniques" made available to NWS to
  - "develop them for operations"

Improved Statistical Post-processing with the **Bayesian Processor of Ensemble (BPE)**, cont.

Toth et al, GSD/ESRL

Status

- Algorithms developed & documented (U. Va, see right)
  - Multiple predictors from multiple controls & one ensemble
- Codes developed (GSD) & tested (GSD & U. Va)
- Transition to MDL (GSD & MDL, by Oct16)
- Enhancement added (U. Va & GSD, by Nov16)
  - Capability to merge ensembles from multiple centers
- Compare with EKDMOS in operational environment
  - Assessment of quality and comp. speed at observation sites
    - Final report by May 2017

Potential use in operational probabilistic NBM

- BPE ready to serve EKDMOS functionality (yr-2, 2017)
- Gridpoint application of BPE on model variables
  - Serve NAEFS functionality (yr-3, 2018 – not yet funded)
- Develop perfect prog application
  - Relate fine scale user variables to calibrated model vars
  - Serve probabilistic NBM functionality (yr-4, 2019)
- Include NAM, HRRR, SREF etc fine scale guidance (yr-5 2020)
  - Expand NBM to short ranges, serve LAPM functionality

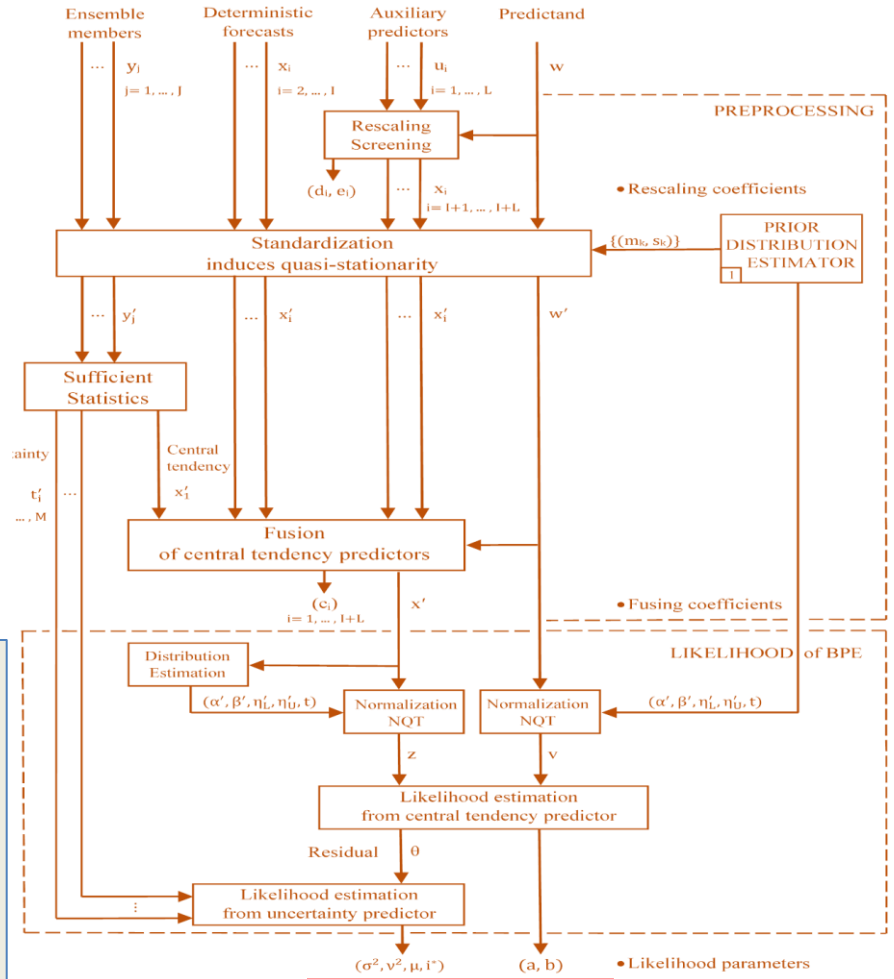
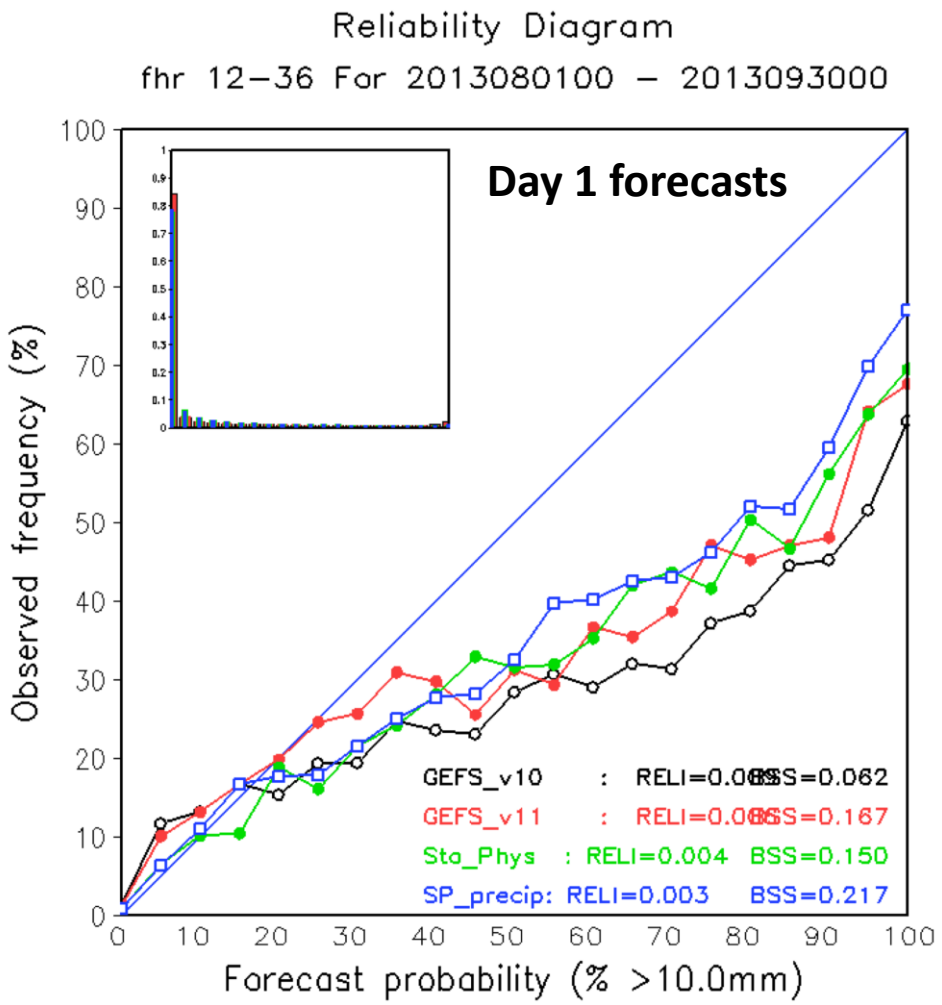
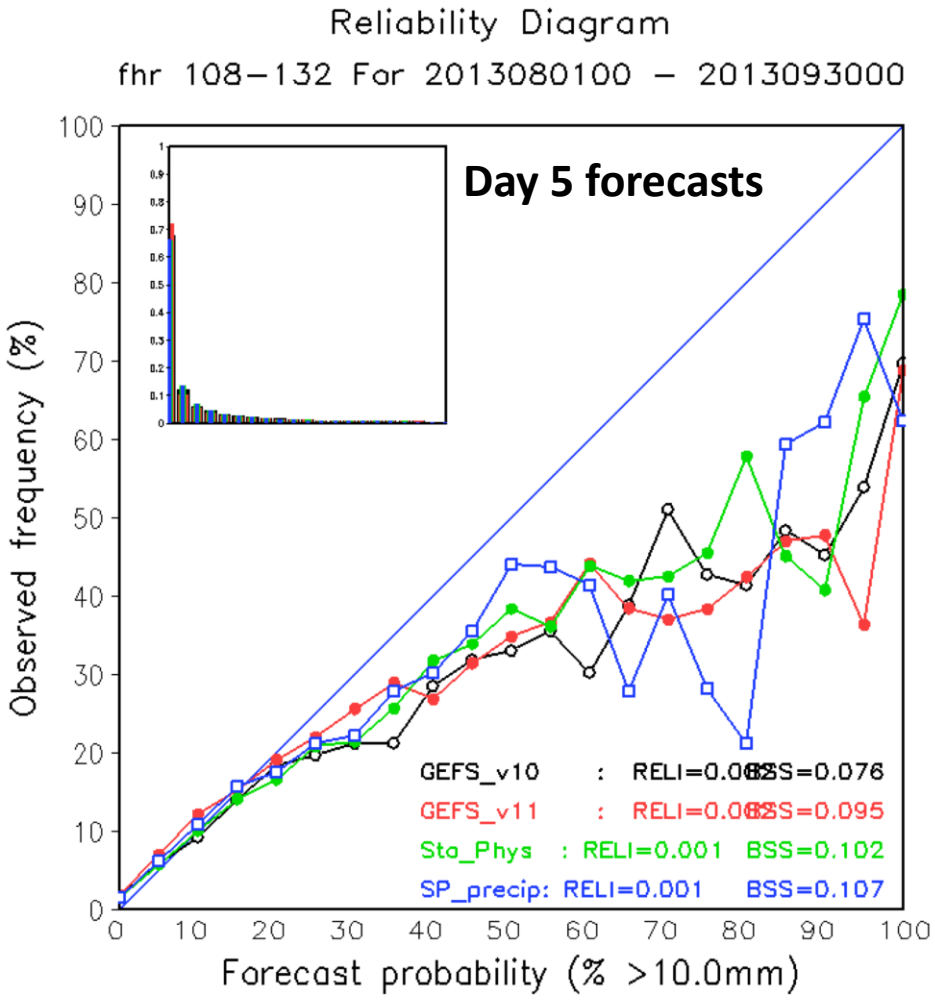


Figure 2. Structure of the likelihood function estimator of the BPE, version **SMM**, wh processes and fuses **S**ingle ensemble forecast, **M**ultiple deterministic forecas **M**ultiple auxiliary predictors.

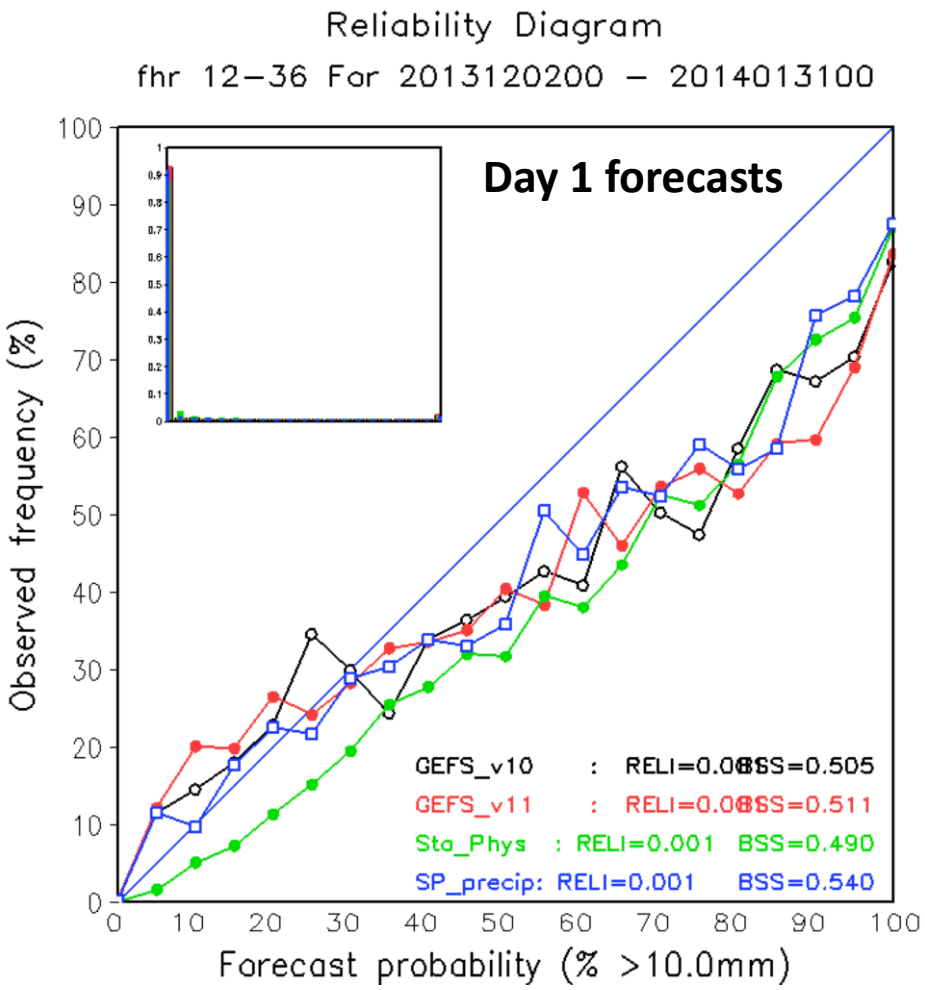
Precipitation reliability for new stochastic parameterizations.

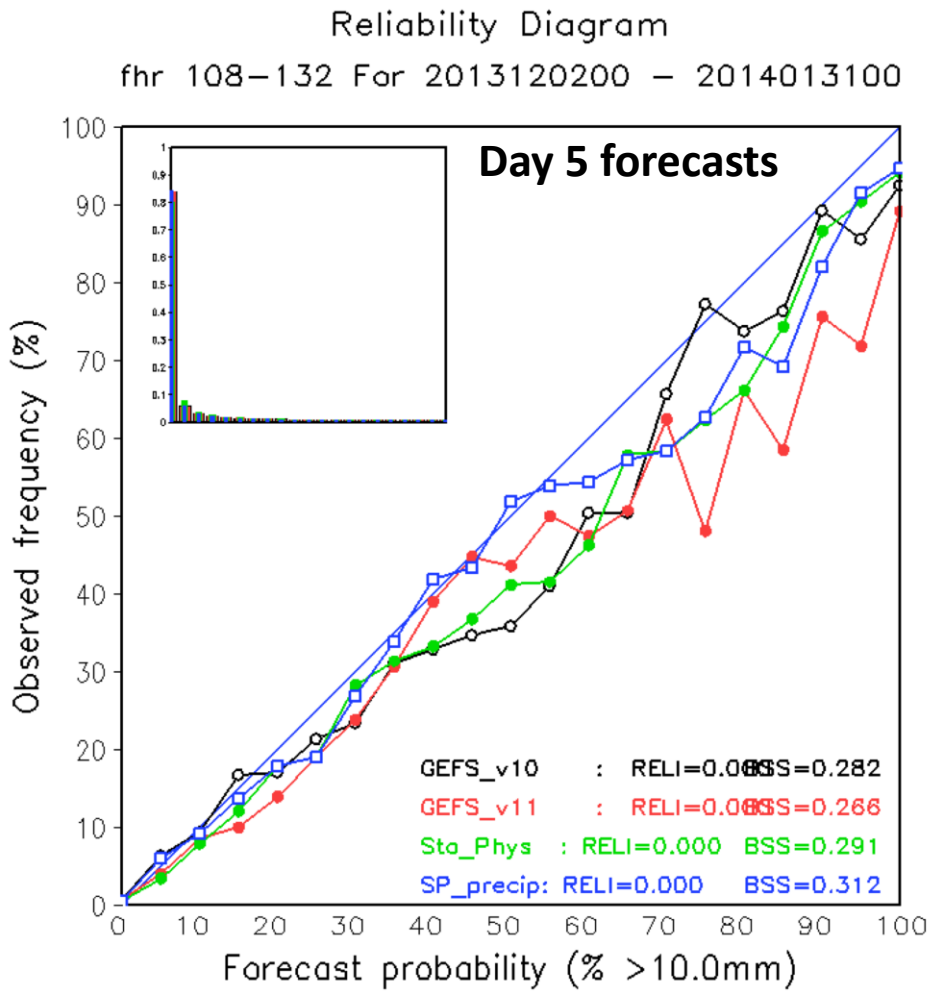




Here, for > 10 mm event and **summer/fall** testing, there is less reliability and skill, but still an improvement with new suite of stochastic parameterizations.

Precipitation reliability for new stochastic parameterizations.





Here, for > 10 mm event and **winter** testing, there is less reliability and skill, but still an improvement with new suite of stochastic parameterizations.