NGGPS ensemble prediction and post-processing team reports

Tom Hamill, Yuejian Zhu, and others

ESRL, Physical Sciences Division

tom.hamill@noaa.gov (303) 497-3060

EPS development: NGGPS major objectives

(1) Develop and implement improved methods for **initializing ensemble predictions**, including the initialization of the coupled environmental state (ocean, atmosphere, land, sea ice, and so forth).

(2) Develop methods to accurately **quantify model uncertainty** in ensemble prediction systems.

(3) Develop ensemble prediction system improvements that will facilitate the generation of reliable and maximally skillful guidance to **lead times of + 30 days and beyond**.

These all contribute to making more skillful and reliable probabilistic forecasts for high-impact weather at lead times of concern to NOAA and its customers.

EPS development: NGGPS major objectives

(1) Develop and implement improved methods for **initializing ensemble predictions**, including the initialization of the coupled

See Jeff Whitaker's presentation also. Work supported via HIWPP, SS, and NGGPS.

(2) Develop methods to accurately **quantify model uncertainty** in ensemble prediction systems.

(3) Develop ensemble prediction system improvements that will facilitate the generation of reliable and maximally skillful guidance to **lead times of + 30 days and beyond**.

These all contribute to making more skillful and reliable probabilistic forecasts for high-impact weather at lead times of concern to NOAA and its customers.

Major objective 1: improved initialization of ensembles

- 4D-En-Var to be operational ~ Apr 2016.
- SPPT, SKEB, SHUM stochastic parameterizations used in data assimilation cycle.
- NGGPS desires (strongly) coupled data assimilation and improved and coupled state estimation of land, ocean, sea ice. Should be a focus area
  - WMO workshop on this, likely Toulouse or Toronto, fall 2016.
- Ameliorating position errors of coherent features.
  - Daryl Kleist (UMD) has NGGPS grant to work on this.
- Minimizing noise in analyzed state – this limits spread growth.
  - from small sample sizes.
  - from sub-optimal model uncertainty treatments.
  - existing methods like Lynch filter sub-optimal; incremental analysis updates being explored (see NGGPS DA team).
  - Fuqing Zhang (Penn State) has NGGPS grant to address via analysis-error covariance singular vectors.

Model uncertainty activities, Sandy Supplemental (SS), HIWPP, and NGGPS

- SPPT, SKEB, SHUM expected to be ready for GEFSv12. Gratifying results (Jeff Whitaker).
  - Get ready for implementation in GEFS v12, ~ 2 years hence, with configuration settled in ~ 1 year, for production of next-gen. reanalyses and reforecasts.
- Estimating parameter uncertainties associated with the land surface (via HIWPP, Sandy Supplemental, NGGPS).
  - Also, want to be ready for GEFS v12 implementation, as above.

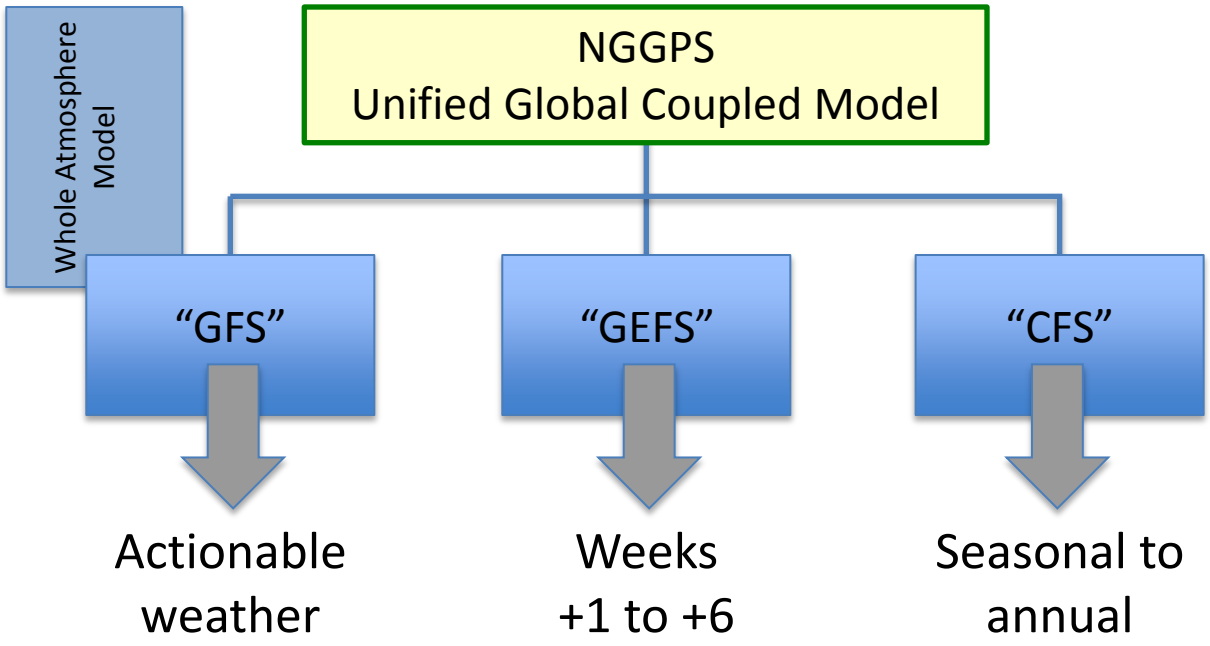
Physically based stochastic parameterization.

- SPPT, SKEB, SHUM are ready for final parallel testing in the GEFS "do the job" but can't all be defended from first principles. We would like to evolve EPSs to a suite that introduce spread in scientifically defensible ways, with a single dycore.

- Important area for NGGPS development; for operations ~5 years hence, need to have university-lab-EMC collaborations, ideally along lines of CPTs, with multi-year funding spanning research and ops.

- ESRL/PSD is putting significant base resources behind this effort, in addition to NGGPS funds.

Ensemble prediction system objective 3: coupled GEFS for sub-seasonal



Update frequency	1 y	2 y	4 y
Length of reanalysis	3 y	20-25 y	1979 - present
Cycles per day	4	1-4	TBD

Extension of forecasts to week +6

- Experiment with GEFS system now, later in context of evolving unified coupled global model system.

- Key science and technical questions.
  - Methods to generate physically consistent atmosphere, ocean, land perturbations.
  - Generating realistic forecast uncertainty in the newly coupled system.
  - Configuration (ensemble size, resolution, reforecast duration/frequency) providing best use of available CPU.
  - Does prediction system represent low-frequency modes of variability that may have predictable skill at 3-6 weeks (MJO, Blocking / AO, ENSO)

Study of GEFS forecasts beyond week 4
(c/o Yuejian Zhu and team)

- Type 1: Un-coupled – atmospheric model only (control)
  - SST relaxation from analysis to climatology
    - Short and medium range skill benchmark
    - Experimental runs readily available
    - Bias correction procedures

- Type 2: One-way forcing – consider SST's impact
  - RTG analysis forcing every 24 hours (AMIP like run)
  - Predicted SST from CFSv2
  - Predicted SST anomaly from CFSv2 with bias correction
  - Perturbed SST – various methods

- Type 3: Two-way coupling
  - Suitably coupled GFS model
  - Coupling with ocean (start at day 0, 5, or 10?)
  - Coupled perturbations

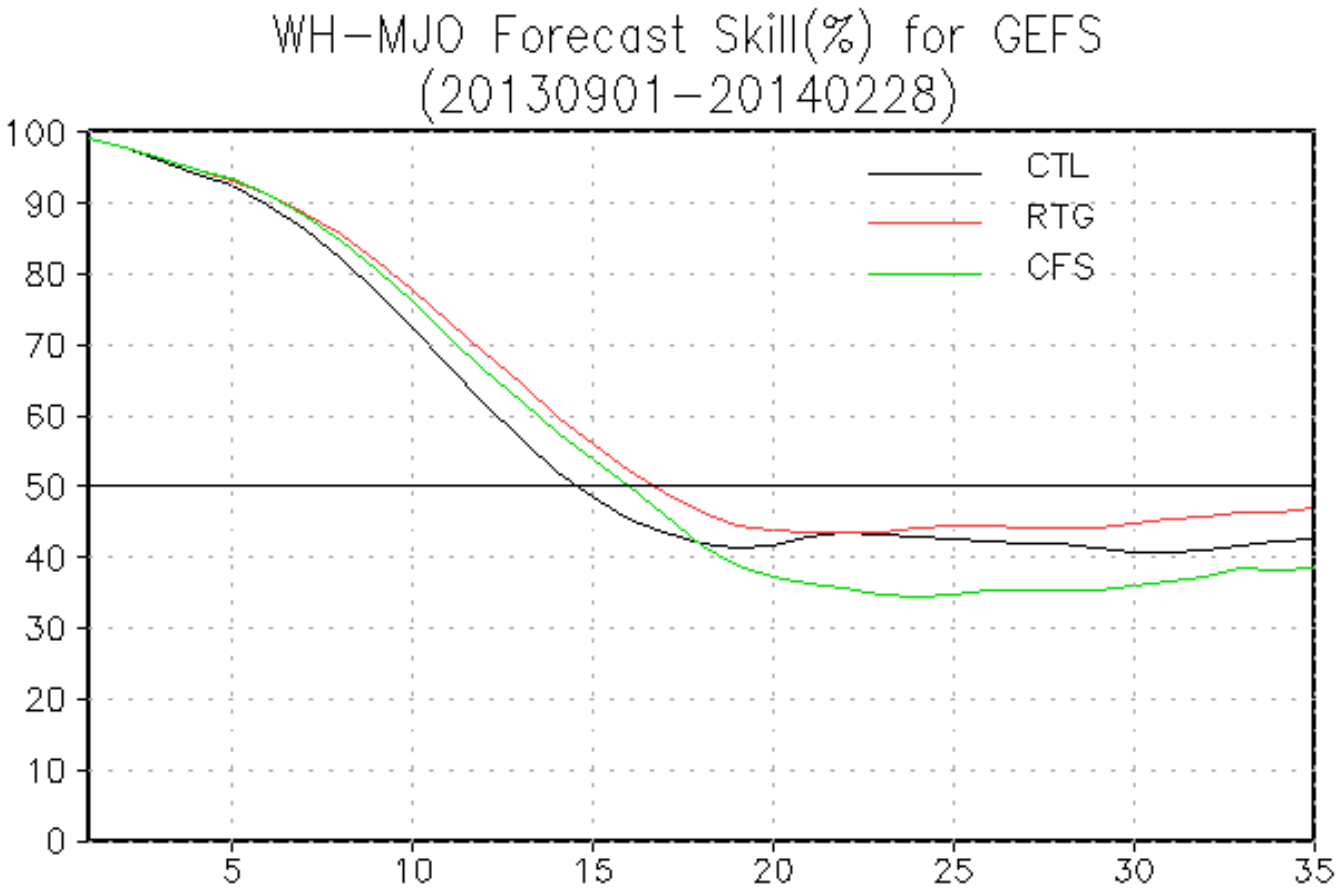
Latest GEFS subs-seasonal experiments

(GEFS V11)

- Extended 2013-2014 winter season (Sep 1 2013 – Feb 28 2014).

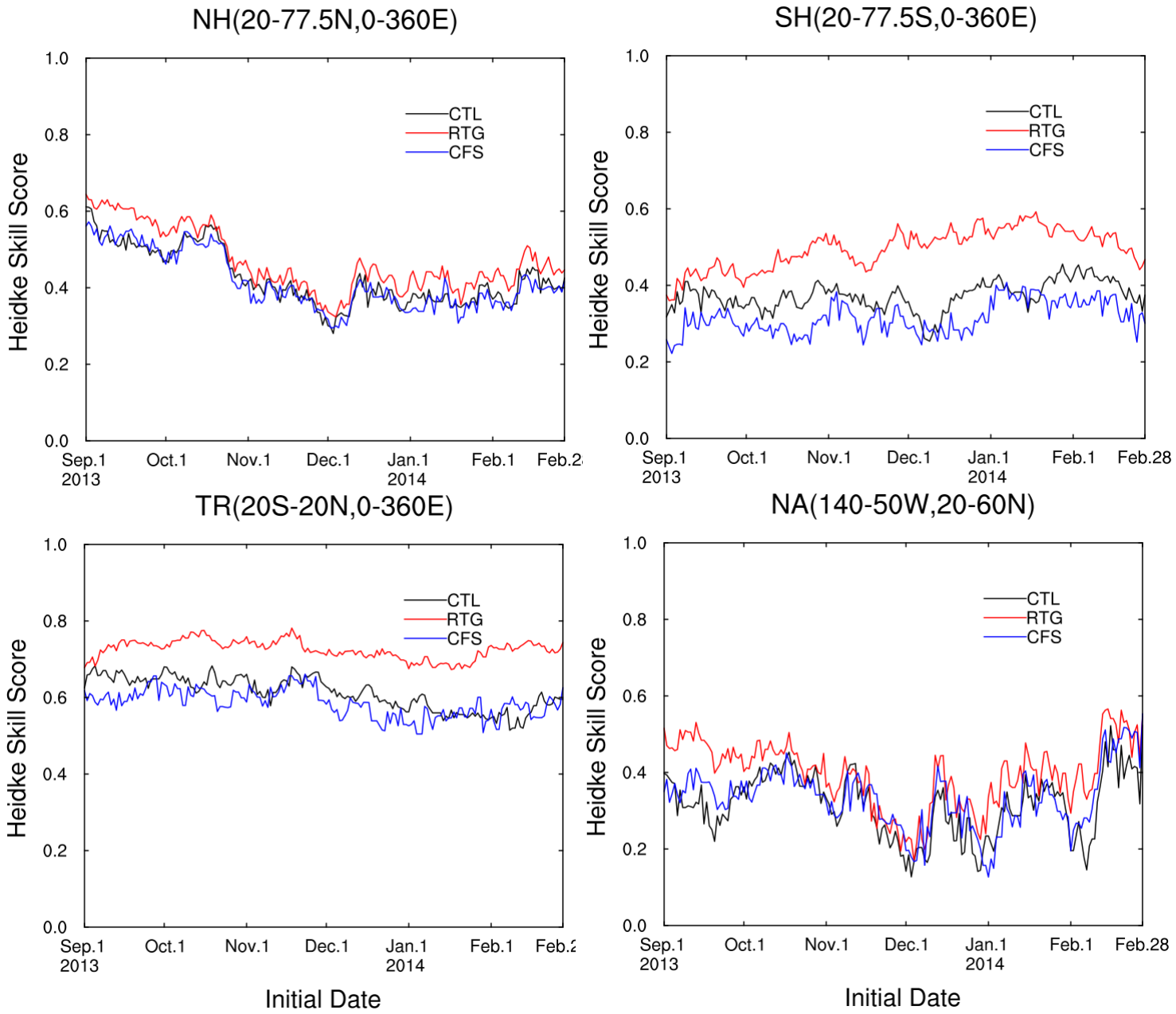
- Four (or more) experiments will be studied:
  - Control (CTL): analysis SST relaxes to climatology; 6 months have been finished [done]
  - Optimum (RTG): realistic SST forcing every 24 hours (AMIP like); 6 months finished
  - Forcing (CFS): CFSv2 predicted SST forcing every 24 hours; 6 months finished
  - Forcing (CFS): CFVv2 predicted SST anomaly with bias correction; experiments just starting
  - Full coupling: coupling with MOM4 – on going

MJO RMM1 &2 forecast skill for GEFS test configurations

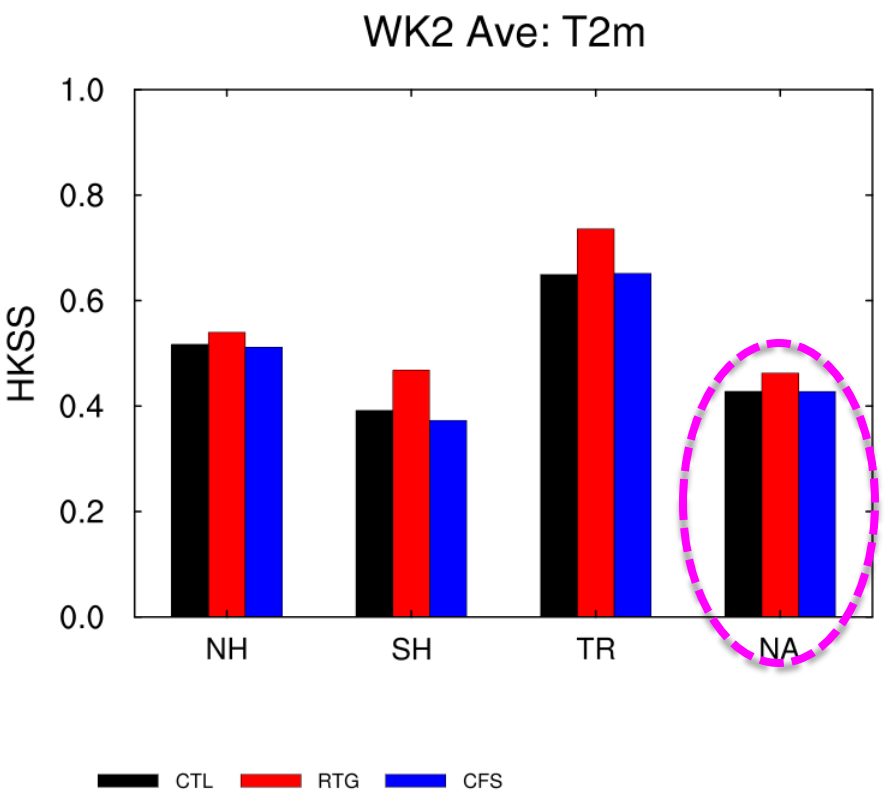


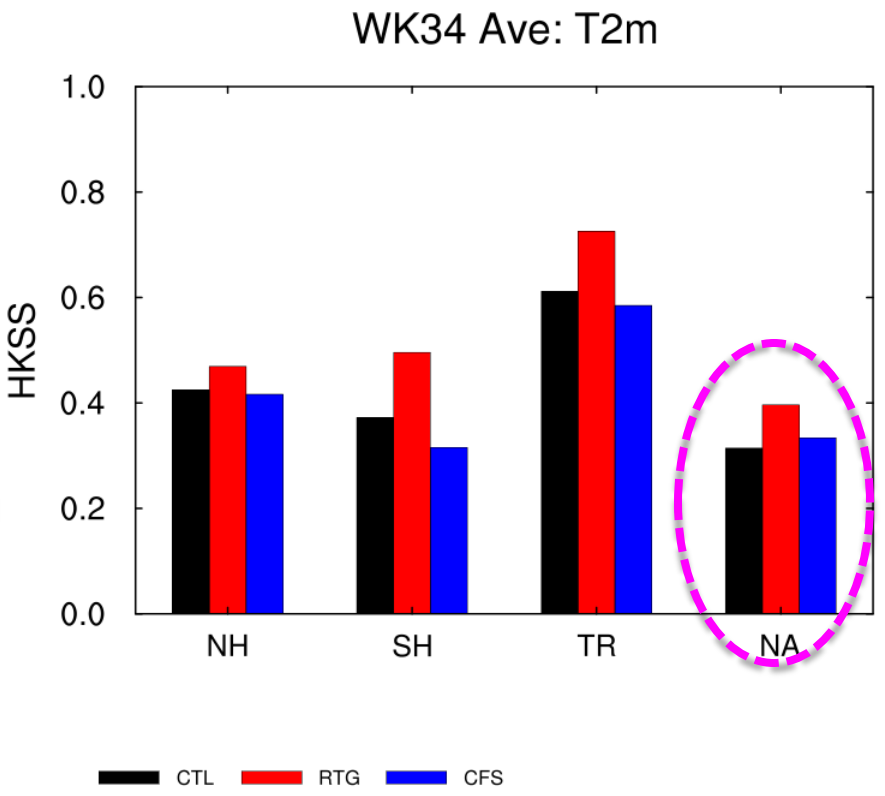
1. RTG-SST (AMIP) runs have highest MJO forecast skill
2. CTL-SST runs have less MJO forecast skills for first 2 weeks
3. CFSv2 predictive SST forcing runs: improved through week 2, worse thereafter; sample size?

Combined Heidke Skill Score for week3-4 averaged T2m



Combined Heidke Skill Score for T2m (WK2 vs. WK3&4)





T2m for Northern American:

RTG (AMIP) runs have best skills

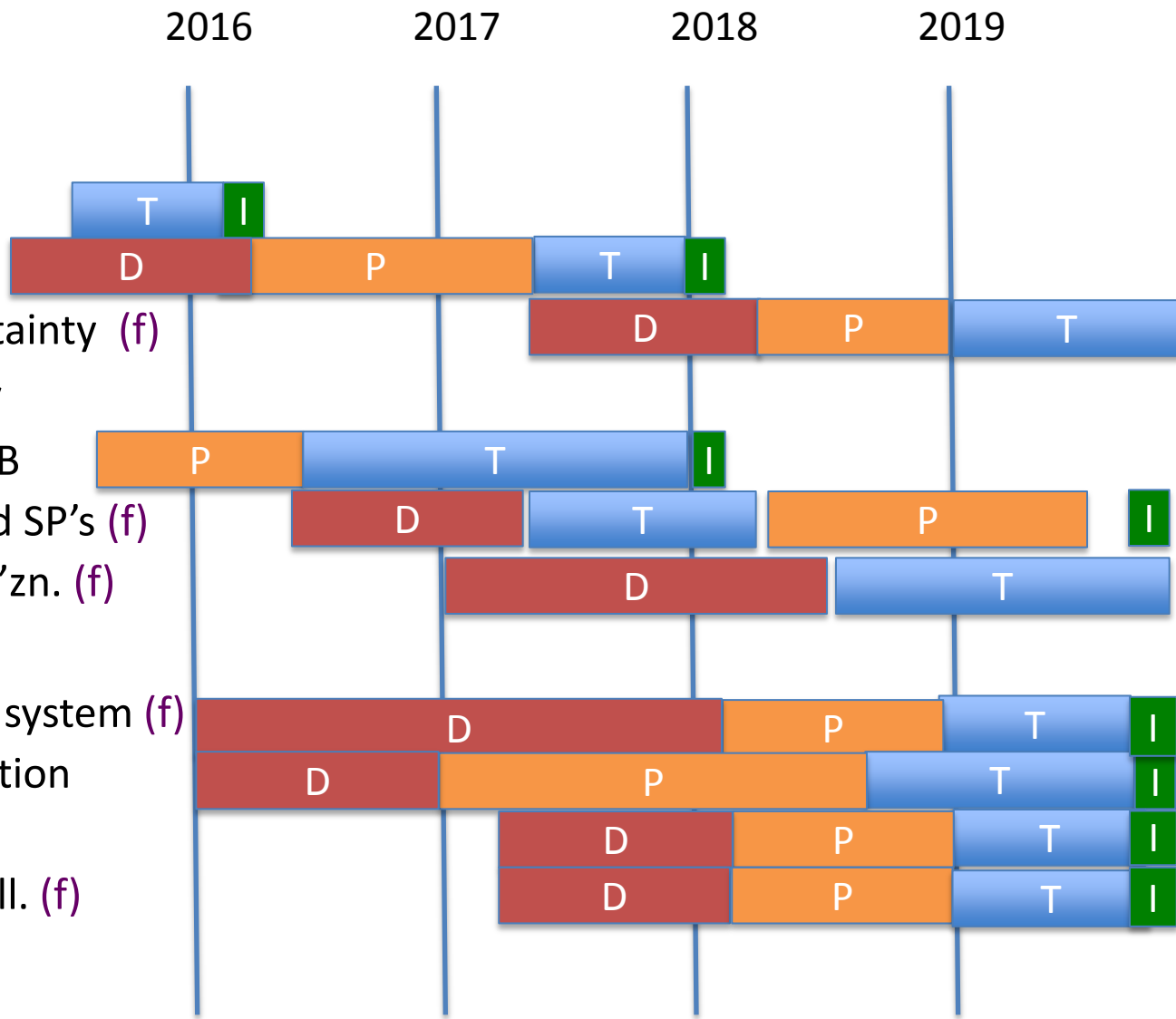
Week-2: The skill is very similar to with/without CFS predictive SST

Week-3&4: CFS predictive SST is slightly better than without (Control)

A tentative GANTT chart for EPS development

(f) = pending funding

- EPS development
  - 4D-En-Var
  - Land surface
  - Ocean, ice uncertainty (f)
- Model uncertainty
  - SPPT, SHUM, SKEB
  - Initial phys. based SP's (f)
  - Stochastic param'zn. (f)
- Monthly forecasts
  - Coupled forecast system (f)
  - Coupled initialization
  - Configuration (f)
  - Evaluation for skill. (f)



D = development P = Preliminary testing T = pre-operational testing I = implementation

Accomplishments, key issues.

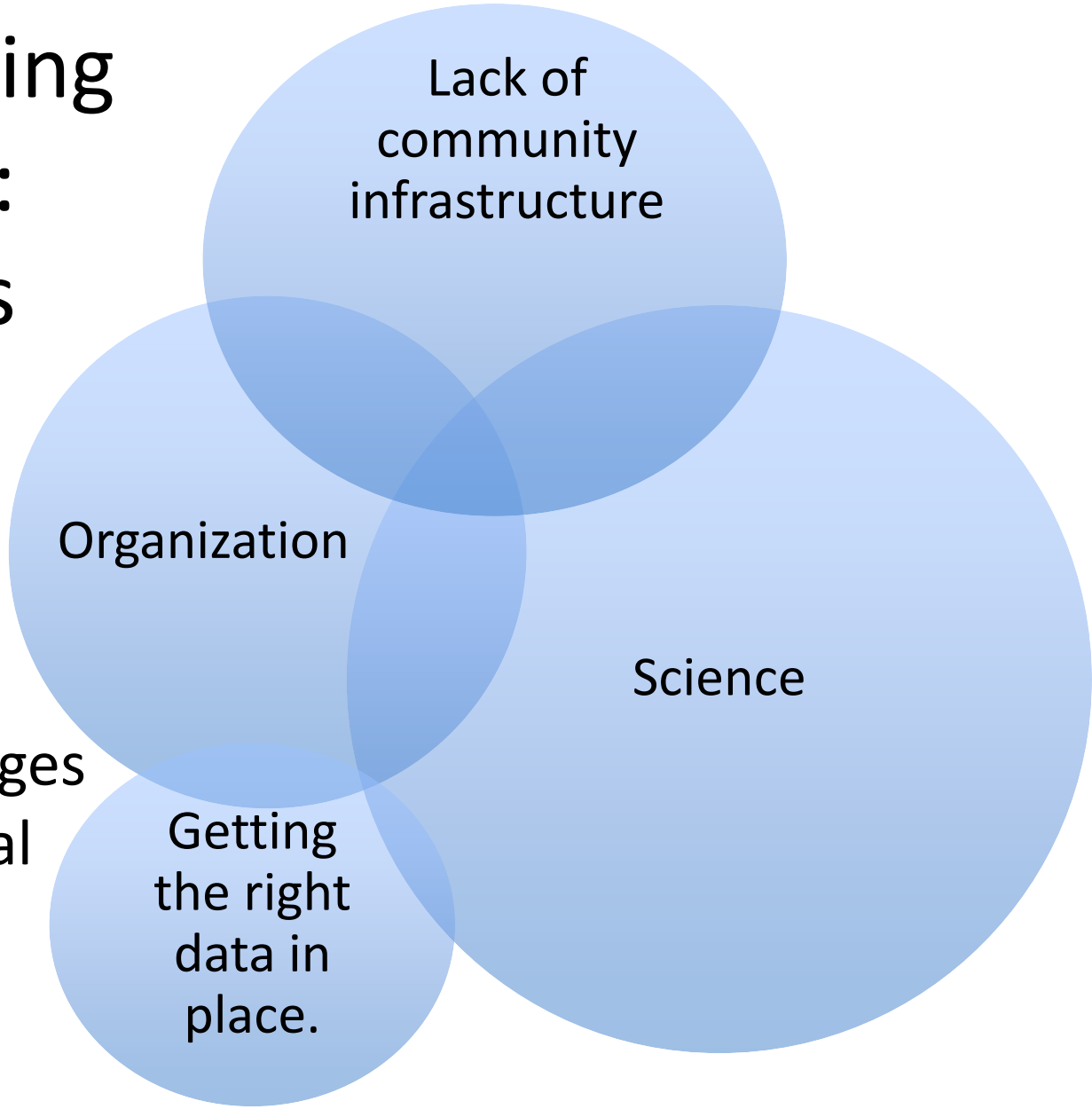
- Accomplishments: Ensemble initialization improved via 4D-En-Var, model uncertainty through SPPT, SHUM, SKEB upcoming. Land-state uncertainty parameterization in the pipeline too.
- Key issues:
  - HPC and storage for R&D*;* WCOSS in better state than theia.
  - Managing EPS implementations in an era of regularly produced reanalysis/reforecast.
    - System configuration decided prior to R/R computation, *~1 year hence*.
  - Academic-sector NGGPS grants.
    - Funded projects don't always project onto the critical NGGPS needs.
    - Ought to be able to either:
      - Require modification of work plans to align more with operational needs.
      - Hold back funding if none of the proposals address NGGPS needs.

Part 2:
post-processing.

NGGPS post-processing objectives

- Conduct post-processing workshop [done; upcoming briefing of NWS senior staff, writing implementation plan(s)]
- Regularly generate supporting data sets, reanalysis/reforecast.
  - With NGGPS and CPO funding, R&D on a next-gen NGGPS global reanalysis is underway.
  - Should include high-resolution reanalyses from a markedly improved RTMA or similar system.
- Improve post-processing algorithms for National Blend.
- Develop post-processing techniques specific to the forecast problems of longer-lead forecasts (weeks 2-4).

Post-processing workshop: challenges



Substantial overlap; some science challenges are also organizational challenges.

Science challenges

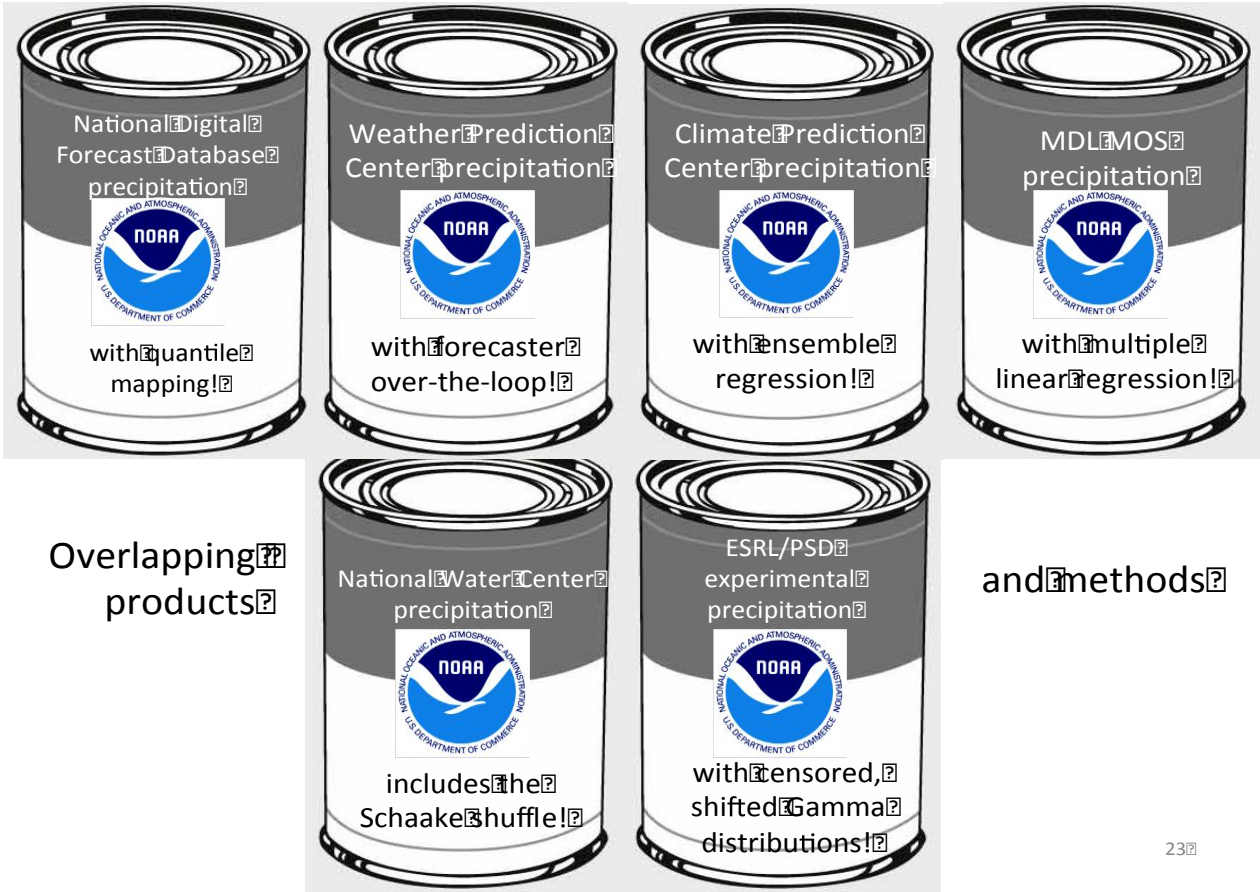
- The methodology behind our current products isn't consistently statistically rigorous.
  - Better products are possible with input of professional statisticians.
  - Preferred methodologies may change as training data is improved and lengthened.
- NOAA has multiple overlapping products produced with a variety of methods, and we haven't carefully evaluated strengths/weaknesses.
- Need reforecasts of high quality and statistical consistency.
- Need observation / reanalysis training data of high quality.
  - If analyses are to be used as surrogate for truth, they must be unbiased and low in error, else product quality suffers.

Organizational challenge

- Existing post-processing infrastructure is complex and hard to maintain.
  - Loss of productivity as we engineer around 20-year old software.
  - Not set up for new era of reforecasts.

Organizational challenge

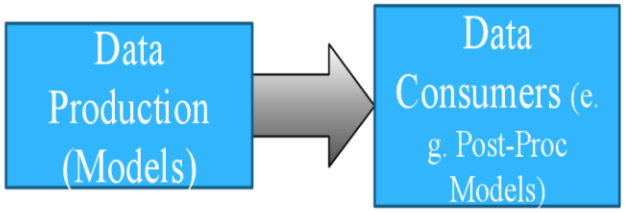
- Parallel product development, using disparate data sets, in several organizations.
  - contributes to lack of seamlessness.



Data challenge

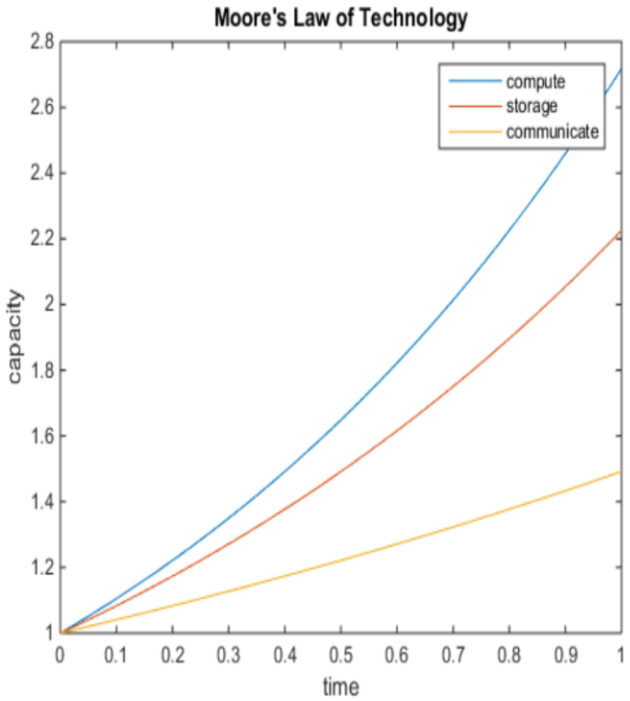
- Performing the product development on different systems from the production of data will cause greater and greater challenges over time, as CPU increases faster than disk and communication.

Our national data flow model



Bring the data to the end uses

Is not scalable and is fundamentally broken



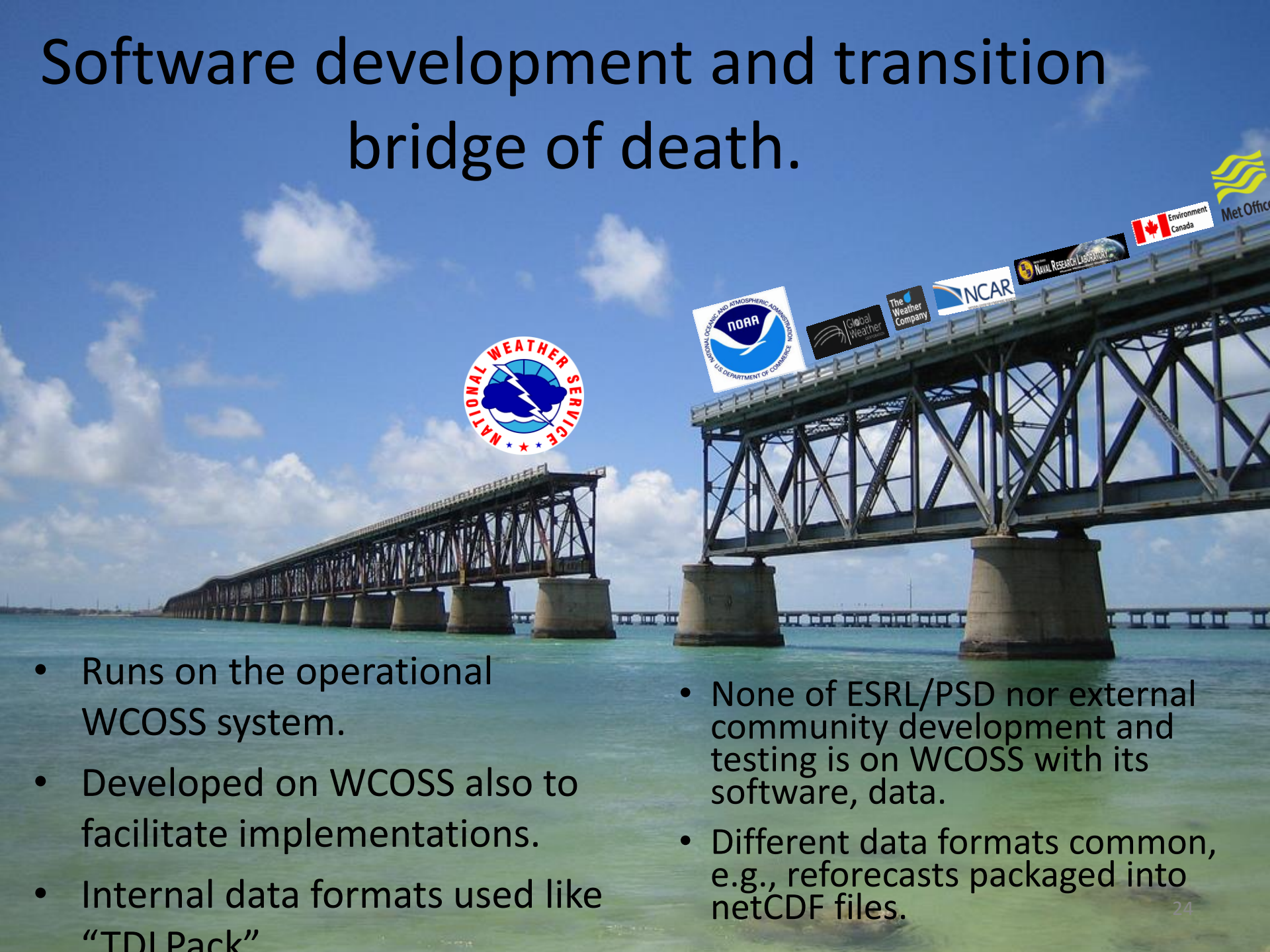
slide c/o Peter Neilley, the Weather Company

Community infrastructure challenge:

little sharing of post-processing software and test data.

(Mostly) unconnected silos of software development; hard to find other's code; sub-standard documentation; unclear policies on use of GitHub and other public repositories; lack of modern version control; training data hard to get.

Software development and transition bridge of death.



- Runs on the operational WCOSS system.
- Developed on WCOSS also to facilitate implementations.
- Internal data formats used like "TDLPack"

- None of ESRL/PSD nor external community development and testing is on WCOSS with its software, data.
- Different data formats common, e.g., reforecasts packaged into netCDF files.

Post-processing workshop findings /recommendations (DRAFT)

- 4 or more statisticians to guide post-process technique development + visiting statistician program.
- Duplication of functionally similar post-processed products across organizations. Hence, perform inter-comparisons of existing techniques, and the use of the best across applications and organizations.
- Reorganization may also be needed to reduce duplication. Options:
  - Continued distributed model but with oversight and governance.
  - New post-processing test bed.
  - More centralization, perhaps in MDL.

Reorganization possibilities: some advantages, disadvantages.

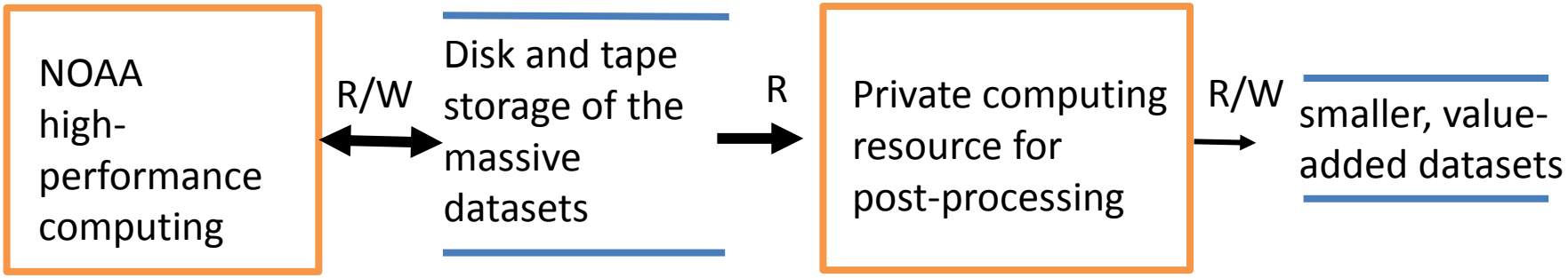
Proposed change	Advantages	Disadvantages
Distributed post-processing development, but with strong post-processing governance and oversight.	• Minimal disruption to existing organizations. • Better coordination, reduction in duplication → more seamless products.	• What if organizational leaders have different vision and priorities from those providing oversight?
Post-processing test bed (from virtual to physical test bed).	• Some positive examples with other testbeds of accelerated R2O.	• Return on past investment in test beds mixed. • Many of greatest post-proc needs not in the middle of R2O funnel, where testbeds flourish. • With community infrastructure and visitor program, importance of physical test bed lessened.
MDL as hub for post-processing development.	• Makes organizational roles more clearly defined, reduces duplication → more seamless products.	• Federal employees hard to move. • New expectations from MDL will require new resources there.

Post-processing workshop findings /recommendations (DRAFT)

- Overhaul of post-processing infrastructure.
  - Regular reanalysis/reforecast production.
  - Organized archival of data.
  - Make NOAA disks readable by private sector?

- Institution of a community infrastructure for post-processing, with modern software management and development practices.
  - Train staff accordingly.

Recommendation: explore possible community storage paradigms, including:



- Create an area where data can be accessed and processed by internal and external collaborators easily, including international collaborators.
- Convenient workaround for storage, I/O bottleneck, especially for corporate partners.
- Risk: can security concerns be addressed?

proposal inspired by Peter Neilley, the Weather Company, following Met Office paradigm.

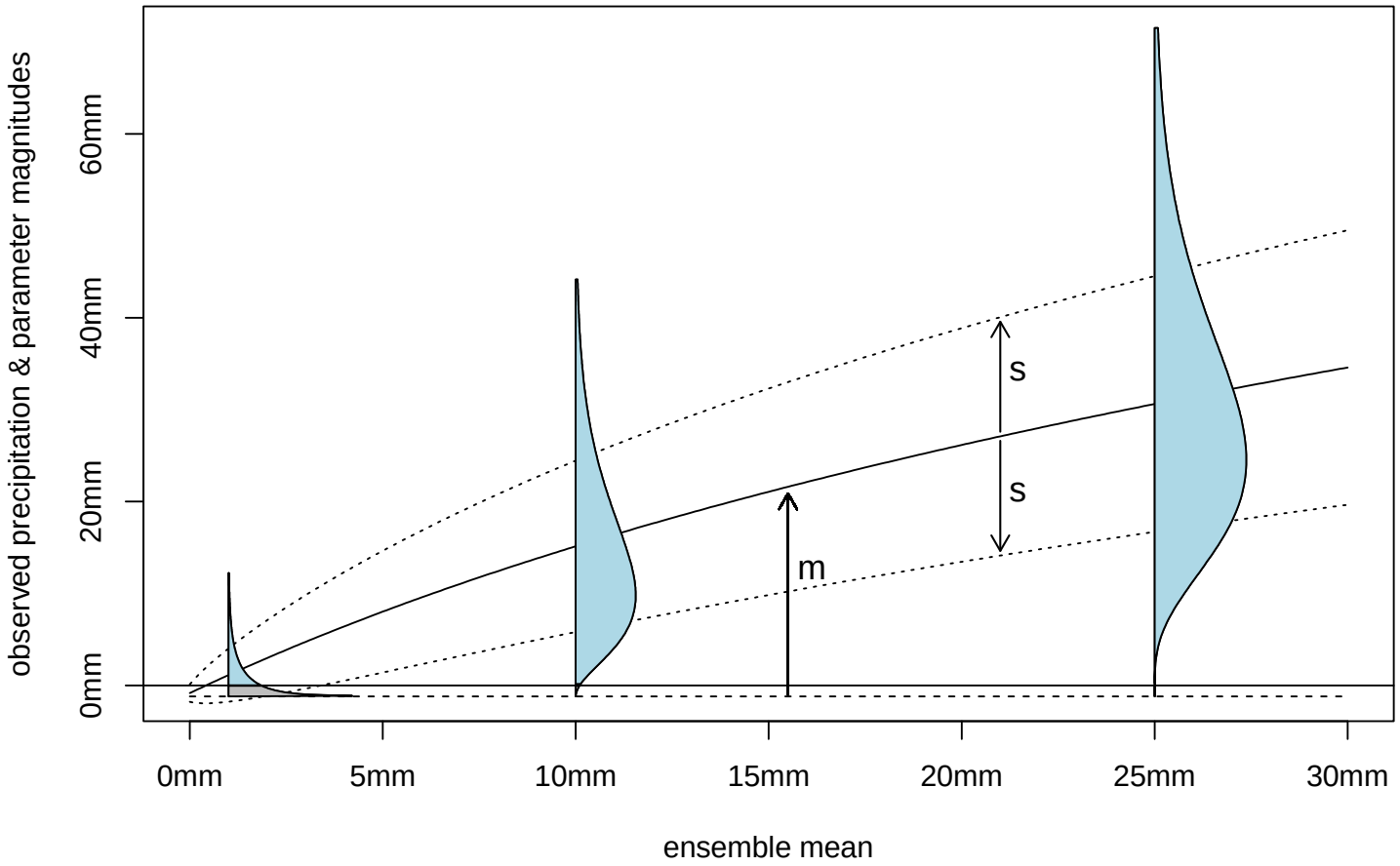
NGGPS post-processing objectives

- Conduct post-processing workshop [done; upcoming briefing of NWS senior staff, implementation plan(s)]
- Regularly generate supporting data sets, reanalysis/reforecast.
- Improve post-processing algorithms for National Blend.
- Develop post-processing techniques specific to the forecast problems of longer-lead forecasts (weeks 2-4).

PP objective #2: regularly generate supporting data sets, reanalysis/reforecast (R/R).

- Reanalysis plans and issues.
  - Produce a next-gen reanalysis/reforecast (R/R) for GEFSv12 (beginning, with NGGPS and CPO support).
  - Future HPC and disk/tape should be sized to include their regular R/R production.
  - Software being developed to facilitate future regular R/R production.
    - Scripts that require minimal human intervention.
    - Re-usable, extendable observations database that support diagnostics.
    - Diagnostic tools.
- Hi-res *surface* reanalysis.
  - Quality of surface analyses: are they low in error, unbiased? Post-processing uses these for training, validation.
  - Are compute cycles allocated for a surface reanalysis?

Post-processing objective #3: improving algorithms



Advanced techniques are in development with Sandy Supplemental and NGGPS funds; more support will be needed for unusual, high-impact variables, support infrastructure.

c/o Michael Scheuerer, ESRL/PSD.

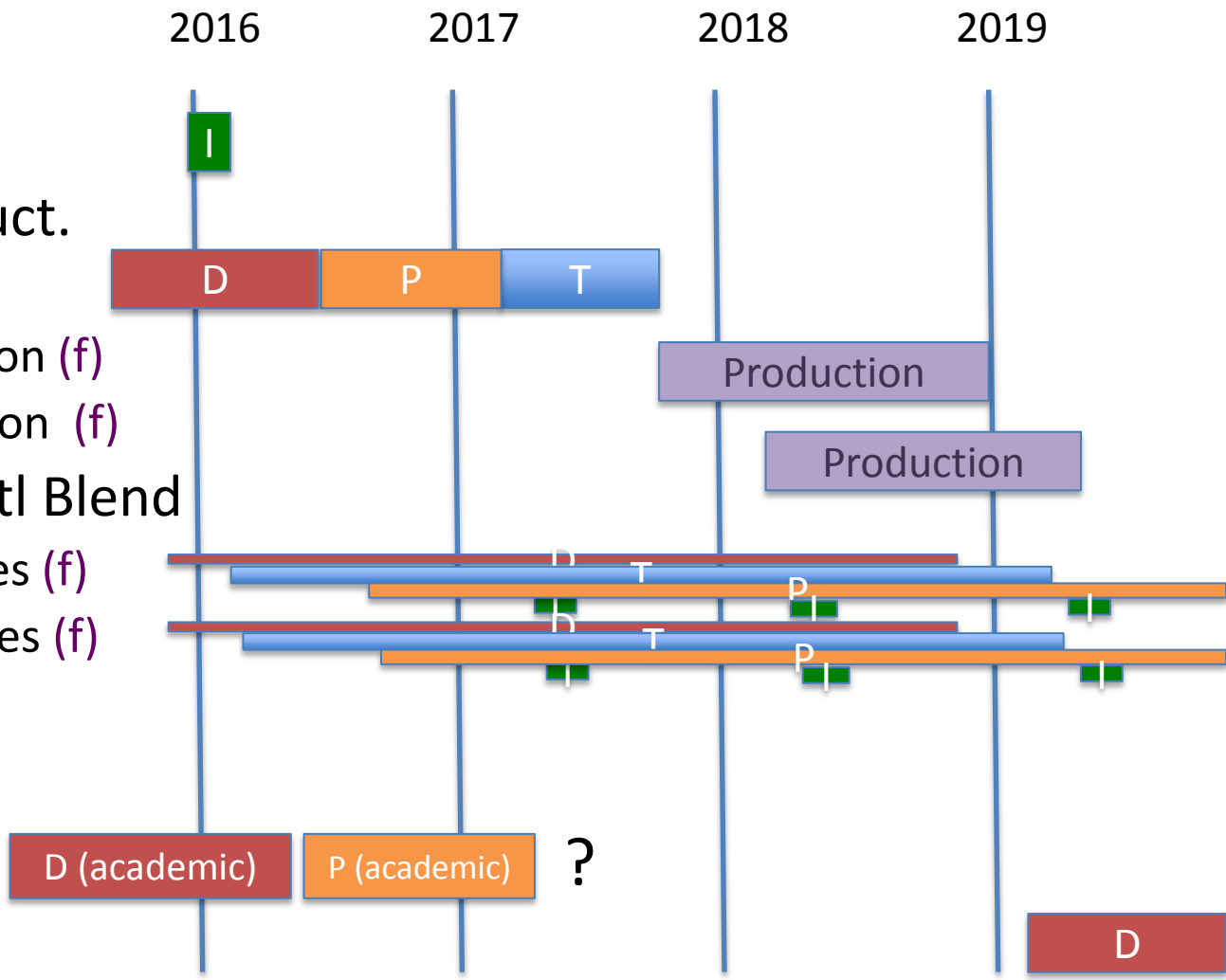
Objective #4: post-processing of weeks +3 to +4

- Initial-condition skill mostly gone, except episodically:
  - ENSO-related circulation changes, MJO, blocking/AO, PNA, etc.
- Small detectable signal buried in large amount of chaotic error, model bias.
- Lengthy reforecasts, stable models needed to tease out what skill there is.
  - Need reforecasts to span multiple climate regimes, ENSO +/-
- New post-processing techniques may be needed, tailored to unique challenges of these time scales.
- Work underway via several NGGPS grants, at CPC, at ESRL/PSD.

A tentative GANTT chart for post-processing

(f) = pending funding

- Workshop
- Supporting infrastruct.
  - Reanalysis prep
  - Reanalysis production (f)
  - Reforecast production (f)
- Improved PP for Natl Blend
  - Improved techniques (f)
  - High-impact variables (f)
- Post-processing for week 3 & beyond
  - Initial work
  - Advanced work



D = development P = Preliminary testing T = pre-operational testing I = implementation

Accomplishments, key issues.

- Accomplishments:
  - Post-processing workshop provided clarity about changes needed; see previous slides.
  - National Blend initial products available over CONUS.
- Issues: as in previous slides:
  - HPC and storage planned for with future R/R.
  - Need for trained statisticians.
  - Streamlining post-processing organizational responsibilities.

Thank you.

- Supplementary slides to follow.

Why is ocean, land, and sea-ice ensemble initialization a priority? In part because surface fields (and precipitation) are under-spread.

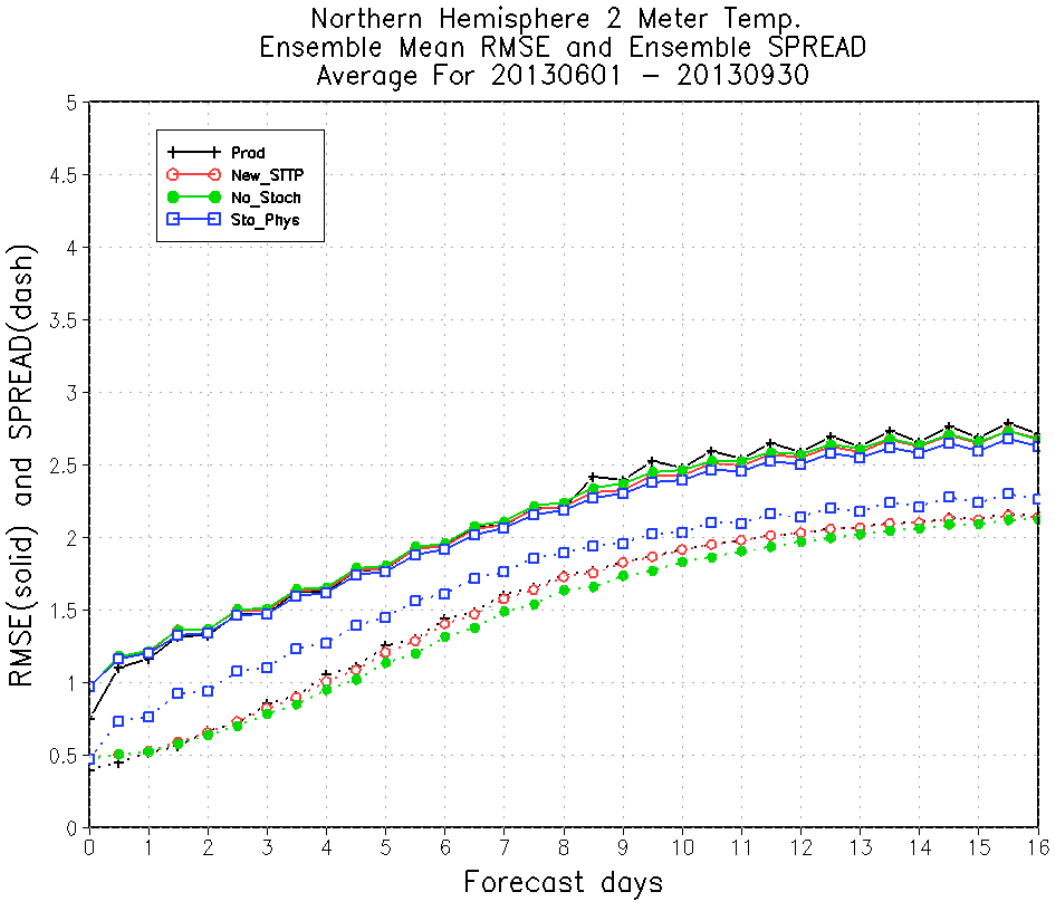


figure c/o Walt Kolczynski, EMC.

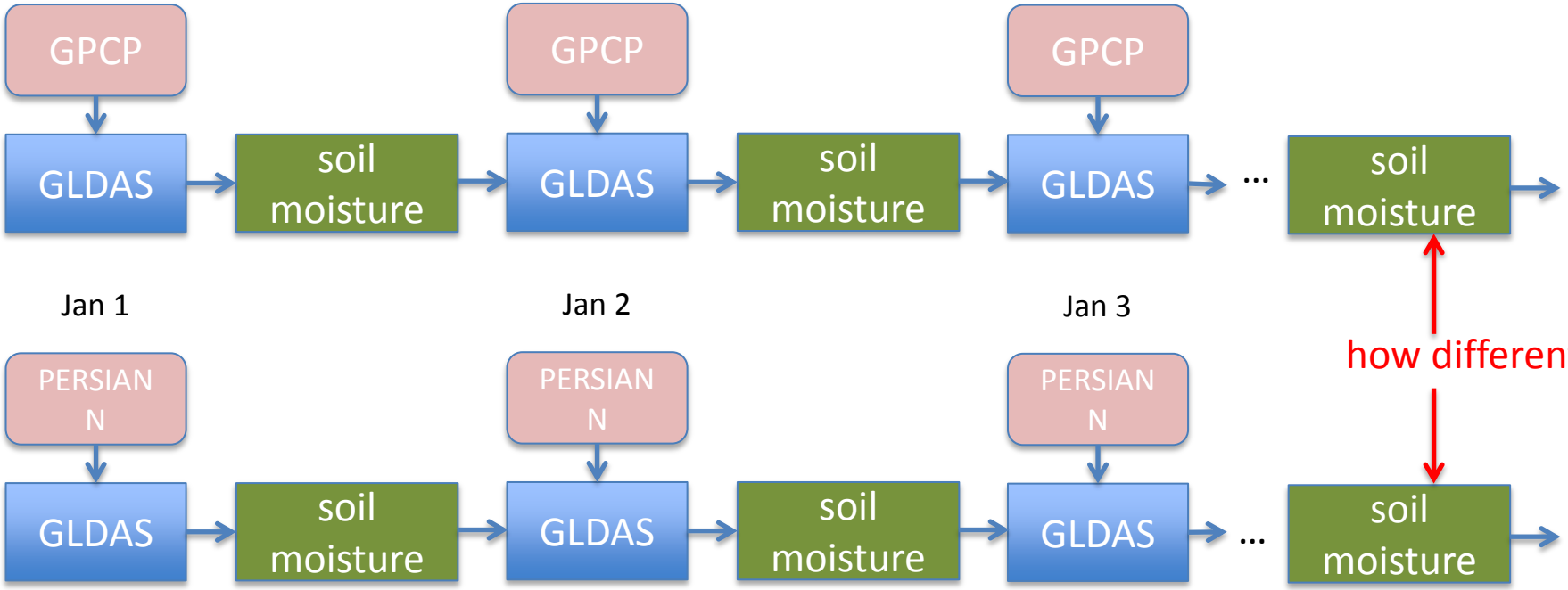
Even with modernized stochastic physics suite (discussion to come), surface temperature is under-spread, leading to unreliable probabilistic forecasts.

In near term, we are attempting to remedy this somewhat by:

(1) perturbations to soil moisture.
(2) land-surface parameter perturbation (discussed later).

Work in progress (via Sandy Supplemental): initialization of soil moisture.

- Determine what are realistic soil moisture perturbations by driving land-surface analyses with different precipitation data sets (cycled over many years).

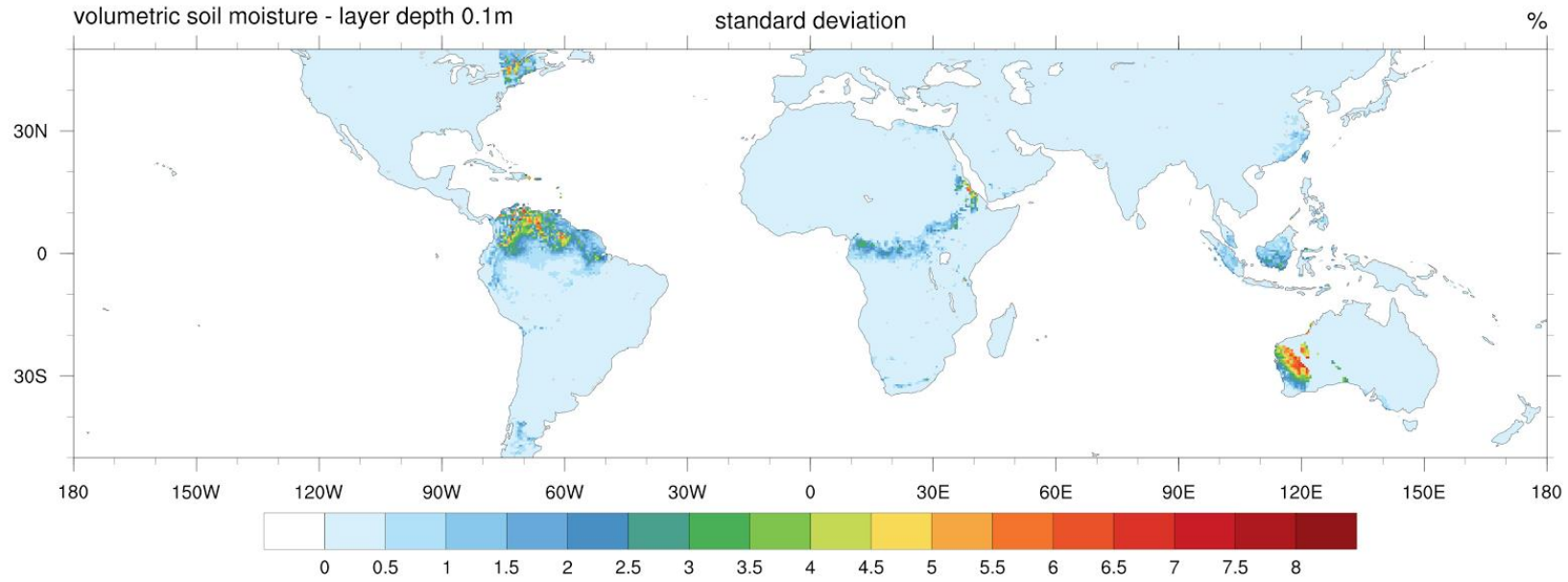


The differences in soil moisture (variance, covariance, etc.) will be used to determine a reasonable perturbation methodology for initial soil moistures.

experiments by Maria Gehne, ESRL/PSD, with help from EMC land-surface team.

Soil moisture differences

- Differences during first month of cycling (still working to extend to several years of cycling).



Mean standard deviation of soil moisture from the four runs forced with different precipitation analyses. The standard deviation is computed at each day and grid point and then averaged over January 2012. PERSIANN, PERSI-CDR, GPCP, CMORPH provide the precipitation forcing.

NGGPS external PI grants in 2015 related to EPS development

- **Development and testing of a multi-model ensemble prediction system for sub-monthly forecasts.** Andrew W. Robertson, PI, Columbia University.
  - **Activities**: Develop and test a multi-model ensemble (MME) prediction system for sub-monthly forecasts (NCEP CFSv2, ECMWF and the Environment Canada model, and other models that become available).
- **Accelerating development of NOAA's next generation global coupled system for week-3 and week-4 weather prediction** Jim Kinter, PI George Mason University.
  - **Activities**: Conduct a series of model development and rigorous testing exercises designed to (1) correct systematic biases; (2) quantify the predictability and skill of weather forecasts for weeks 3-4.
- **An investigation of the skill of week-two extreme temperature and precipitation forecasts at the NCEP WPC.** Lance Bosart, PI, University at Albany, SUNY,
  - **Activities**: Evaluate newly proposed percentile forecast methods, persistent flow anomalies, and NH climate database in context of WPC's development of new forecast formats for Days 8-10. These forecast formats and methodologies for identifying EWEs will be tested in the WPC Hydrometeorological Testbed, and then will be implemented into WPC operations.
- **Exploitation of Ensemble Prediction System Information in support of Atlantic Tropical Cyclogenesis Prediction.** Chris Thorncroft, Pi, University at Albany, SUNY.
  - **Activities**: To ensure that recent and current research concerned with the variability of African easterly waves (AEW) structures and downstream tropical cyclogenesis probability is transferred into operational decision-making at NHC, and to develop and evaluate tools that exploit key information in dynamical ensemble prediction systems in support of tropical cyclogenesis prediction.

NGGPS academic grants in 2015 related to post-processing

- **Development of Ensemble Forecast Approaches to Downscale, Calibrate and Verify Precipitation Forecasts.** Dave Novak, WPC, PI
  - **Activities**: Enhance the skill of high-resolution quantitative precipitation forecasts (QPF) for detection of high-impact events via downscaling, quantile mapping
- **Calibration and Evaluation of GEFS Ensemble Forecasts at Weeks 2-4.** Ping Li, PI, SUNY Stony Brook.
  - **Activities**: Decompose GEFS extended range into a limited number of principal components to calibrate with observations.
- **Probabilistic Forecasts of Precipitation Type and Snowfall Amounts based on Global Ensemble Forecasts.** Tom Hamill, ESRL/PSD.
  - **Activities**: Develop novel experimental post-processing methods for precipitation type and snowfall amount.
- **An Investigation of Reforecasting Applications for Next Generation Aviation Weather Prediction: An Initial Study of Cloud and Visibility Prediction.** Dr. David Bright, NOAA/NWS/NCEP Aviation Weather Center, PI.
  - **Activities**: Utilize NOAA's second-generation Global Ensemble Forecast (GEFS) reforecast dataset, and be the first aviation-based GEFS reforecast study to construct a model climatology and downscaled calibrated prediction of instrument meteorological conditions (IMC).
- **Improved Statistical Post-Processing with the Bayesian Processor of Ensemble (BPE).** Zoltan Toth, PI, NOAA/OAR/ESRL/GSD.
  - **Activities**: Develop scientifically based, comprehensive algorithms and software for use in unified NWS statistical post-processing operations to address both the calibration of prognostic variables and the derivation of additional user variables. Test and demonstrate the algorithms for the calibration of prognostic variables.