NOAA's Subseasonal Experiment

U.S. DEPARTMENT OF COMMERCE
National Oceanographic and Atmospheric Administration
Silver Spring, MD

Table of Contents

Executive Summary	2
I. Status of SubX	3
Description	3
Operational use	5
Research use	7
II. Current Review Objectives	8
Objectives and Analysis	8
III. Potential Future for SubX	15
Support Program	15
Governance	15
Protocol Suggestions	16
Potential Goals	17
Appendix A: SubX Citations	19
Paper Submitted, In Press or Appeared	19
Presentations	21
Appendix B: Supplemental Output Variables	25
Appendix C: Review Agenda and Speakers	27

Cover image by Tom Hamilton, submitted to NOAA Weather in Focus Contest 2015

Executive Summary

NOAA's Climate Program Office (CPO) originated the Subseasonal Experiment (SubX) in FY16, with an initial plan for two years of work. The project focused on improving subseasonal predictions, largely targeting the Week 3-4 outlooks from NOAA's Climate Prediction Center (CPC), and providing a research and experimental forecast dataset for the community to explore sources of subseasonal predictability and quantify the associated prediction skill in leading modeling systems. Contributing modeling centers are NCEP, NCAR, NRL, ESRL, NASA, Environment and Climate Change Canada (ECCC). Contributing academic partners are from multiple organizations, including George Mason University (GMU), Rosenstiel School of Marine and Atmospheric Science (RSMAS), IRI (International Research Institute), Florida International University (FIU), University of Texas (UT) Austin, NOAA/CPC, and NOAA's Environmental Modeling Center (EMC).

Based on the favorable results and outcomes in the first two years, namely from research contributions from external (academic, nongovernmental organizations, and other federal partners) and internal contributors, the project was continued for each of FY18 and FY19. For these continuation years, Subseasonal to Seasonal (S2S) Program funding was provided to RSMAS for further transfer to GMU and IRI; with other partners supporting their efforts internally.

The project developed and tested the delivery of a weekly-updating subseasonal forecast capability, including a mix of operational and research models, to inform CPC's Week 3-4 Outlooks. External partners (e.g. 4682 unique users) served via the IRI database were able to download 74.7 terabytes of data. The models were evaluated both within NOAA and externally by the broader research community, showing the benefit of the multi-model construct to increase skill and decrease uncertainty.

The panel recommends that SubX be continued for an additional period of development, assessment and experimentation that also informs ongoing operations. The future work should further strengthen the linkages between operational and research modeling centers, advance development of methods to optimize information extraction from the multi-model output, support exploring more efficient pathways to model enhancements (indirectly supporting other major modeling initiatives), as well as document and analyze use cases to ensure NOAA is meeting the needs of the fullest set of stakeholders. During the continuation period, the governing body should develop and scrutinize approaches for optimizing the operational component for maximum decision-support impact, as well as help forge the underlying research and research-to-operations components to identify optimum pathways for improvements. The panel also recommends that the program be moved out from the Climate Testbed (CTB), and a governing body be established under a multi-agency program office such as the National Earth System Prediction Capability (National ESPC) or the Office of the Federal Coordinator for Meteorology (OFCM). Realignment could facilitate better linkages to other proven programs that have delivered benefits from research to operations.

I. Status of SubX

Description

The SubX Project consisted of five 24-month grants supported through the NOAA CPO beginning in FY16. The SubX proposals also included co-Principal Investigators (co-PIs) and collaborators from NCEP, with one project having an NCEP lead-PI. Maintenance funding was provided for FY18 and FY19 to keep the experimental forecast data flowing. For these continuation years, funding was provided to RSMAS for further transfer to GMU and IRI; other partners supported their efforts internally. SubX was originally funded by NOAA's Climate Program Office (CPO), but in FY19 SubX was transferred to NOAA's Office of Weather and Air Quality (OWAQ) as part of NOAA research restructuring under the The Weather Forecasting and Research Innovation Act of 2017, Public Law 115–25 (hereafter, "The Weather Act").

The Subseasonal Experiment developed by CPO largely focused on model predictions for the Week 3-4 time period, to support nascent prediction products for that lead time. The minimum data outputs extended to 32 days lead time, but some models have output available to 45 days[1].

Table 1: Models contributing to the SubX project

MODEL			
CFSv2	NCEP	4/day	45
GEFS	NCEP	20	35
GEM	ECCC	4	32
GEOS-5	NASA/GMAO	10	45
ESPC	Navy	4	45
CCSM4	NCAR	3 or 4/day	45
FIM-HYCOM	NOAA/ESRL	4/week	32

An integral part of the ensemble database was the retrospective reforecast dataset required of each participating member covering at least the years 1999-2015, which is required to be updated with every model upgrade. Data are available for download from the IRI data library[2].

The datasets were delivered in multiple phases to alleviate strain on the participating modeling centers. The requested variables were prioritized and divided into 3 groupings. Data in priority 1 (Table 1) were

[1] http://cola.gmu.edu/kpegion/subx/data/modelinfo.html
[2] http://iridl.ldeo.columbia.edu

chosen to support CPC operational outlooks, while other variables were generated for initial research (Table 2).

Table 2: Priority I fields provided by the modeling centers

VARIABLE		
Geopotential Height	500, 200 hPa	Average at 0,6,12,18Z
Zonal Velocity	850, 200 hPa	Average at 0,6,12,18Z
Meridional Velocity	850, 200 hPa	Average at 0,6,12,18Z
Temperature	2m	Daily average
Skin Temperature	surface	Daily average
Outgoing Longwave Radiation	Top of Atmosphere	Accumulated every 24 hr
Precipitation	surface	Accumulated every 24 hr

Table 3: Priority II fields are available to inform research needs.

VARIABLE	LEVEL	FREQUENCY
Specific Humidity	850 hPa	Daily average
Vertical velocity (omega)	500 hPa	Average at 0,6,12,18Z
Zonal velocity	100 hPa	Average at 0,6,12,18Z
Meridional velocity	100 hPa	Average at 0,6,12,18Z
Wind u component	10 m	Average at 0,6,12,18Z
Wind v component	10 m	Average at 0,6,12,18Z
Daily maximum Temperature	2 m	24 hour maximum
Daily minimum Temperature	2 m	24 hour minimum
Dewpoint temperature	2 m	Daily average
Sensible Heat Flux	surface	Accumulated every 24 hrs
Latent Heat Flux	surface	Accumulated every 24 hrs

Zonal stress	surface	Daily average
Meridional stress	surface	Daily average
Mean Sea Level Pressure	Sea level	Average at 0,6,12,18Z
Snow water equivalent	N/A	Accumulated every 24 hrs
Net radiation	surface	Accumulated every 24 hrs
Snow Density	surface	Daily average
Snow cover	surface	Daily Average
Vertically integrated soil moisture	N/A	Daily Average
Sea ice concentration	surface	Daily average
Convective Available Potential Energy	N/A	Daily average

Operational use

Both realtime and reforecast datasets were downloaded and used at CPC for the creation of multiple tools as part of the forecaster-in-the-loop process to generate the Week 3-4 outlooks for temperature and precipitation. The SubX tools complement the suite of tools derived from the operational models that were already being ingested at NCEP, and provide more information about potential outcomes during the target forecast period. The model output fields for 500 hPa geopotential heights, 2-m temperature, and precipitation were all routinely plotted and made available to NCEP forecasters. Figures made available to forecasters for the discussion and subsequent production consisted of anomalies for the variables plotted for each model, probabilities of being above or below normal (defined as mean for temperature and median for precipitation), e.g. Fig.1.

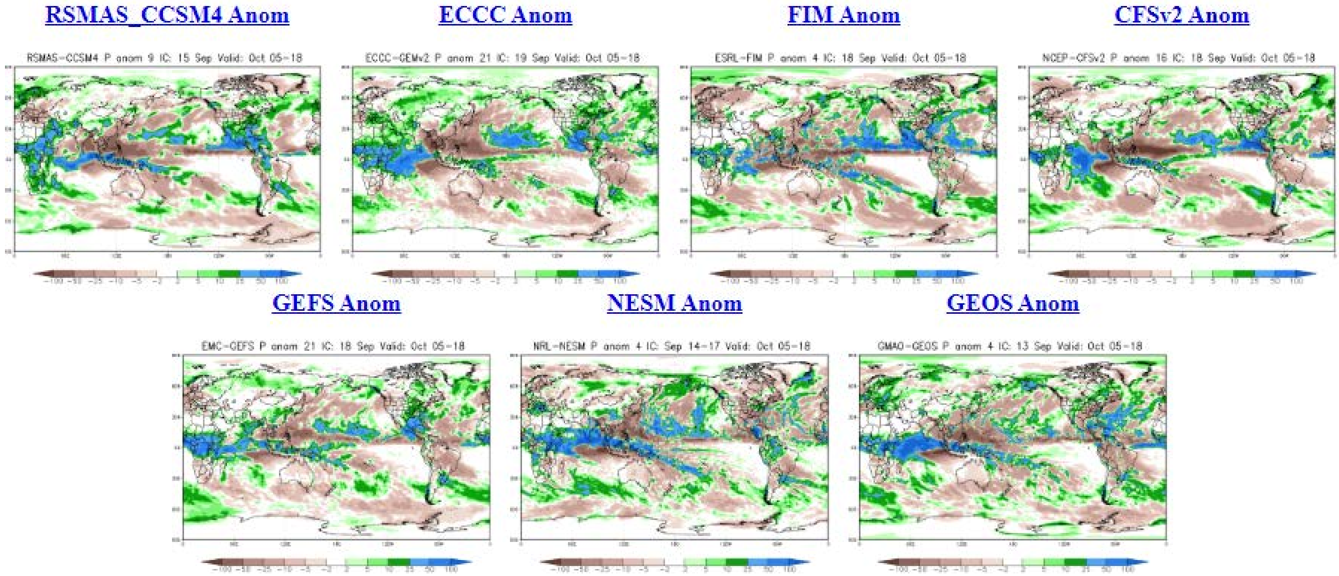

Figure 1. Representative precipitation anomalies for each SubX participating model. Source: NCEP

The anomalies and probabilities were constructed versus each model's own hindcast, which demanded the calculation of a historical multi-model mean as a reference for the multi-model mean realt-time plots. The initial plots and presentation format followed closely to those available for the North American Multi-Model Ensemble (NMME), which was valuable as many of the forecasters that create the official Week 3-4 outlooks at CPC, also use NMME data when they are responsible for the longer range monthly and seasonal outlooks (Fig. 2).

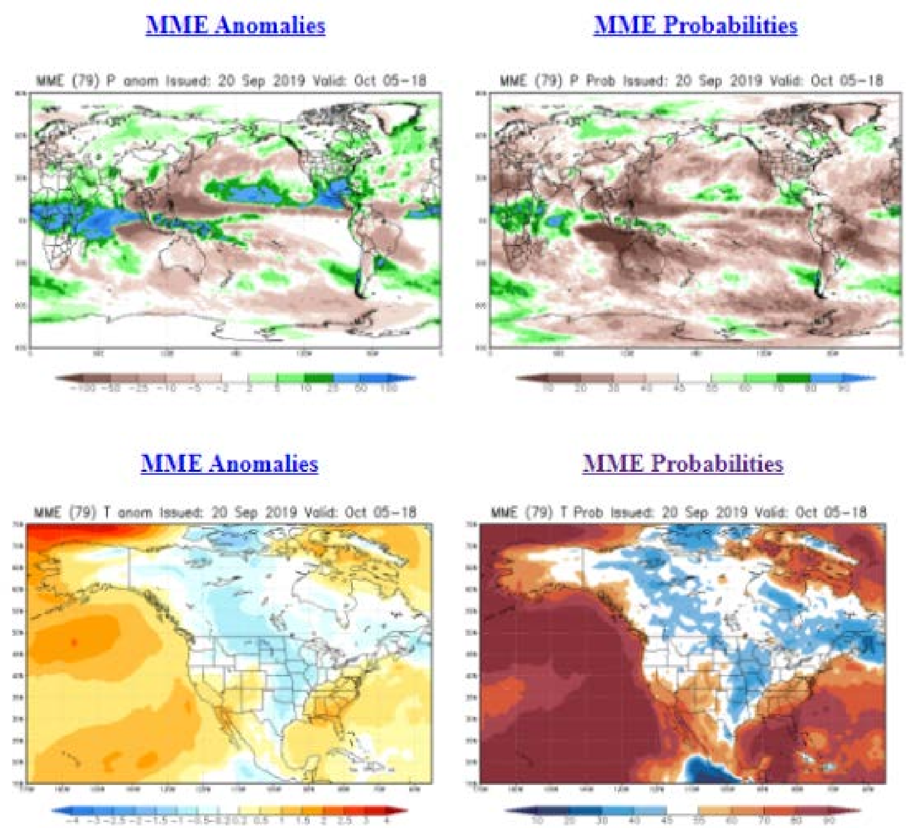

Figure 2. Combined precipitation and temperature probabilities and anomalies. Source: NCEP

During the forecast production process at CPC, SubX model output was called up and used to inform the official forecast. Investigators from some of the contributing modeling centers participated in the weekly forecast discussion, garnering feedback from the operations and contributing to the discussion, a live instance of the R2O2R loop. Furthermore, research indicates that tropical cyclone tracks can be evident in SubX data, and many times the patterns in forecast precipitation were reflected in the official outlooks. So while a direct tropical cyclone forecast from SubX was not formally incorporated into the outlooks, such phenomena were highlighted by the model suite, having a direct influence on the official outlooks.

Research use

A list of citations and presentations appears at the end of this review (Appendix A). As of the time of the review, at least 23 papers had been published or submitted, and 43 presentations given. Given that published peer-reviewed literature is the principal metric of scientific contribution, this has been strong return on investment for CPO's Modeling, Analysis, Predictions and Projections (MAPP) program, particularly as the SubX data set will continue to enable further research for years to come.

SubX research has demonstrated the value of parallel datasets for assessing model performance and prediction skill in bulk (Guan, 2018b; LaJoie, SubX review presentation) and for case study events (Wang, 2018); informed research including ocean eddies and their role in predictability, tropical air-sea interactions (Janiga, 2018), water quality (Ross, submitted, 2019), land-air interactions (DeAngelis, SubX review presentation), and coastal flood prediction (Pegion, SubX review presentation). Investigations along these lines allowed researchers to delineate the best practices within the SubX model suite by cross-comparing results with specific forecast system attributes (data assimilation, coupled vs uncoupled through the integration, varying ocean initial conditions, etc.). Comparing and contrasting the different methods employed across the models, without the coordinated nature of the SubX experiment, would be a much more difficult, if not impossible task. SubX allowed for a much more concentrated and robust analysis, potentially enhancing the flow of research to operations by highlighting the positive attributes that should be singled out for transition.

Furthermore, the SubX dataset is also leveraged in funded projects from CPO's FY18 CTB call (RLs 5-8), will be used in OWAQ's funded FY19 S2S research call (RLs 2-4), and potentially will be used in OWAQ's FY20 Climate Test Bed call (RLs 5-8). This structured and curated dataset will provide the field with years of valuable data to examine, and yet there is even more potential given the degrees of freedom that existed in the SubX project.

II. Current Review Objectives

The SubX project review adhered to the terms of the CPO MAPP program – NCEP CTB partnership governance documents. The post award review process was delayed by OAR's reorganization in response to the Weather Act, but has been executed in the form of the 19 August 2019 review that prompts this document. Appendix C contains the review agenda.

The SubX web page[3] lists six objectives for the project (listed and discussed below), which include basic elements (1-4) as originally proposed in the FY2016 CTB proposal (NA16OAR4310141), plus two additional elements (5-6) instituted with initiation of the project. These serve as fair metrics for evaluating the value and success of the first years of SubX and are addressed point-by-point below, with a brief assessment of whether each has been achieved. Additional metrics were presented by funded investigators and other CPC partners, demonstrating the improvements gained through the multi-model ensemble.

The SubX team successfully demonstrated an updating subseasonal prediction capability based on a combination of operational and research models. In addition, the various academic partners demonstrated the ability to maintain a pseudo-operational workload and product delivery to keep up with data cycling and to troubleshoot inevitable communications issues. This updating dataset represented a substantial workload and differed from the existing NMME by having weekly, versus monthly, initialization and data transfer.

Objectives and Analysis

1. *Collecting and serving data both internally at CPC for use by operational forecasters and for the external community via the IRI data library*

The SubX project has provided value to CPC operational outlooks at subseasonal time scales, contributing to the weekly forecast decision process, reducing the area of "equal chances" forecasts for temperature and precipitation in weeks 3 and 4, and contributing week-to-week stability to forecasts. The SubX project has provided value to the participating modeling centers themselves as described in presentations given by representatives of most of the participating models. The SubX project has provided value to the research community, particularly to members of the MAPP S2S task force outside the SubX CTB project – access statistics from the IRI data library show continued growing use of SubX data.

Presentations by SubX principals and discussion during the review meeting demonstrated that SubX forecasts are collected and served in a timely manner on a firm weekly schedule designed to synchronize with CPC issuance of outlooks on Fridays for weeks 3 and 4. All hindcast and near-real-time forecast data are made available to the community through the IRI data library; SubX and IRI

[3] https://cpo.noaa.gov/Meet-the-Divisions/Earth-System-Science-and-Modeling/MAPP/Research-to-Operations-and-Applications/Subseasonal-Experiment

personnel have developed and provided software and documentation to facilitate both access to and the use of the data.

Additionally, there are other users of the pseudo-operational forecasts, of which Dr. Pegion highlighted two examples during the reivew. A real-time coastal Integrated Flood Forecast System (iFLOOD) uses SubX forecast data to drive a water level and wave height model; its forecasts are compared to operational models for the Atlantic coast with particular focus on the Mid Atlantic for lead times up to a month (Fig.3).

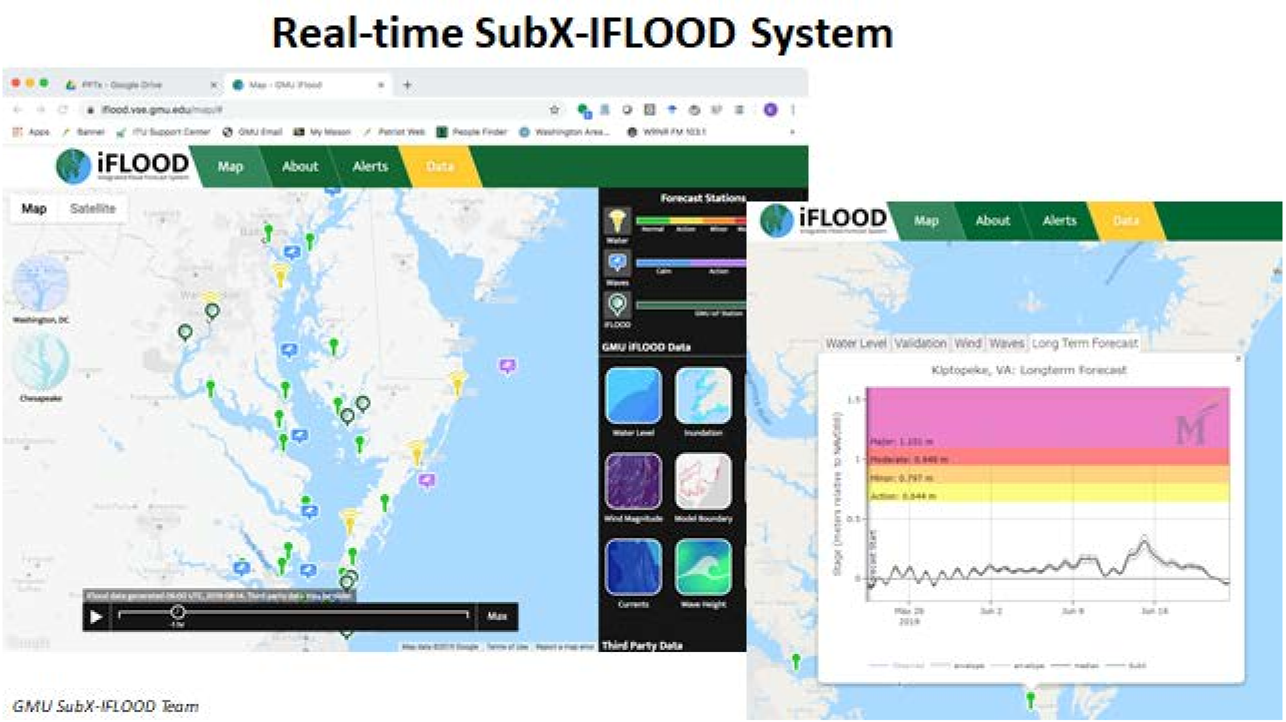

Figure 3. Sample use of SubX output in a pseudo-operational setting. Source: Pegion brief, slide 31.

SubX data is also being used to drive the Chesapeake Bay ROMS Community Model (ChesROMS)[4] estuarine model, which uses the Rutgers Regional Ocean Modeling System (ROMS) over the Chesapeake Bay region to predict sediment transport, water quality, hypoxia and related quantities affecting ecosystem and human health. The potential applications in this time period are just beginning to be explored, and that points to the need to further the research, so eventual operational implementations will reside on firm scientific footing.

Additionally, there are many users in the research community, particularly within the MAPP S2S project, who are using SubX data to investigate diverse topics such as heat wave initiation and maintenance, tropical storm genesis and precipitation, MJO and NAO predictability, air-sea interactions, model resolution impacts on forecast fidelity, and the sources and consequences of model biases (See Appendix A).

[4] http://ches.communitymodeling.org/models/ChesROMS/

Some potential users have stated that because of the uncertain future for SubX, they have been hesitant to put in the effort to incorporate SubX forecasts into their applications. Concern that the quasi-operational forecasts, in particular, could cease in the near future has been a barrier to wider applications for the SubX forecasts.

2. *Providing a baseline verification particularly for the weeks 3-4 temperature and precipitation probability forecasts*

The review presentations demonstrated the focus on temperature and precipitation forecasts. Kirtman showed validation results for weeks 2-4 from SubX multi-model forecasts, particularly how the multi-model forecast usually performs better than any of the individual contributing models (Fig. 4).

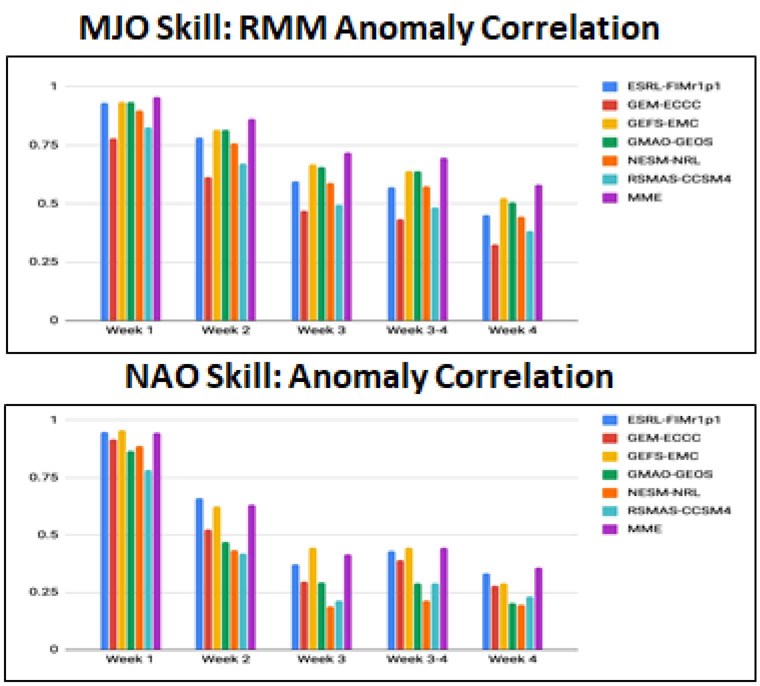

Figure 4. Anomaly correlations for MJO and NAO for different lead-times. Source: Kirtman brief, slide 30.

LaJoie provided independently calculated results showing similar benefit due to the inclusion of multiple models in the ensemble. Metrics shown included Heidke Skill Scores for temperature and precipitation (Fig. 5), RPSS, anomaly correlation, and the DeSole and Tippett Sign Test[5] (Fig. 6).

[5] Delsole, Timothy & Tippett, Michael. (2015). Forecast Comparison Based on Random Walks. Monthly Weather Review. 144. 150904101551007. 10.1175/MWR-D-15-0218.1.

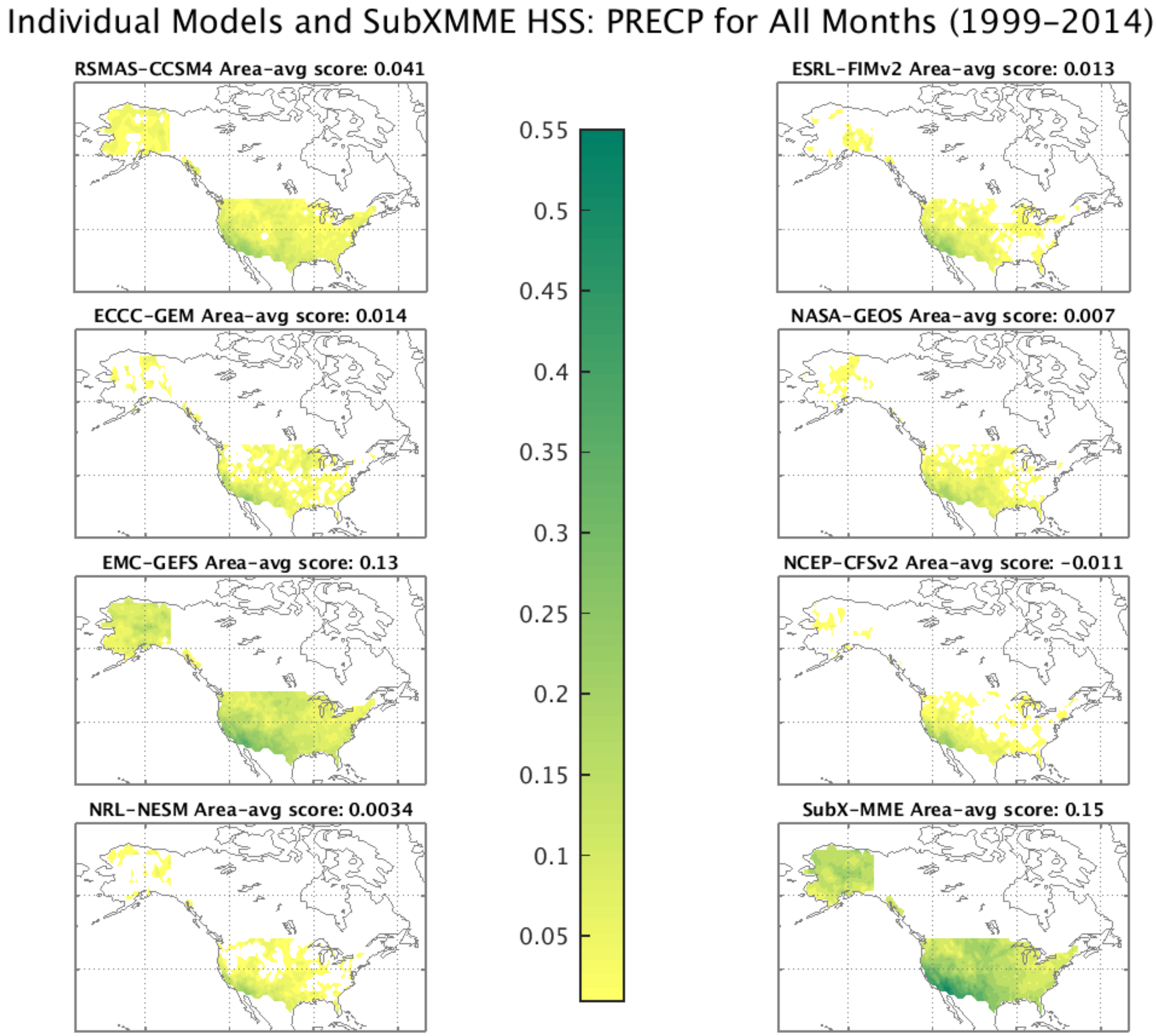

Figure 5. Heidke Skill Scores for US precipitation for SubX models. Source: LaJoie brief, slide 9.

3. *Investigating multi-model combinations including selecting suitable models, optimizing the design of the system, and evaluation of the prediction products.*

The ensemble was constructed from what models were available from the modeling centers and which were able to provide the needed/minimum ensemble members according to the rubric of entrance criteria. The ensemble has not yet been subjected to rigorous study of optimization criteria such what number of models, and what number of members of each model, would yield the best results and for which set of stakeholders. For instance, the optimal ensemble construction might be different for regions overseas than for over the North American continent.

Sensitivity testing of multi-model construction from SubX forecasts at CPC has demonstrated that larger numbers of models contribute to forecast stability (less variation between consecutive forecasts). It was also shown that inclusion of SubX forecasts into the existing suite of NCEP and other nations' operational models demonstrably improves overall multi-model forecast skill. LaJoie and Kirtman both showed “random walk” tests that quantified single and multi-model skill as a function of lead, and pinpointed CFSv2 in the comparisons. Inclusion of models that individually perform poorly in forecasts of specific variables does not seem to degrade multi-model forecasts whereas better performing models clearly boost multi-model skill, consistent with previous findings regarding multi-model predictability.

LaJoie showed results of a targeted CPC verification exercise wherein the impact of SubX multi-model forecasts on skill compared to existing single-model (CFS) and multi-model (CFS+GEFS and CFS+GEFS+ECCC) forecasts was evaluated. In each case, addition of SubX forecast data to each of the combinations of models resulted in net improvement of forecast skill (Fig. 6).

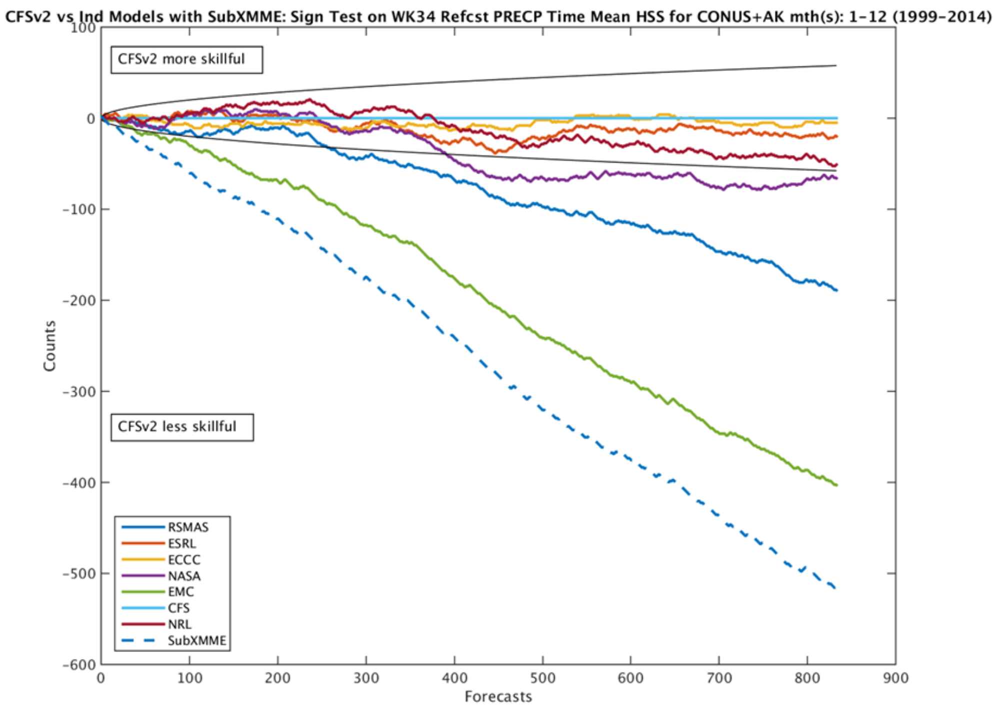

Figure 6. Sign test results for precipitation in SubX models, relative to CFSv2. Source: LaJoie brief, slide 17.

Collins showed results comparing official guidance configurations using ECMWF and JMA forecasts combined with CFS, although terms of data agreements with non-US providers prevented a full suite of comparisons to be shown. Nevertheless, SubX data was shown to contribute in several ways, particularly for calibrated multi-model forecasts. The additional models also add forecast-to-forecast stability and diversity, increasing ensemble spread (Fig. 7).

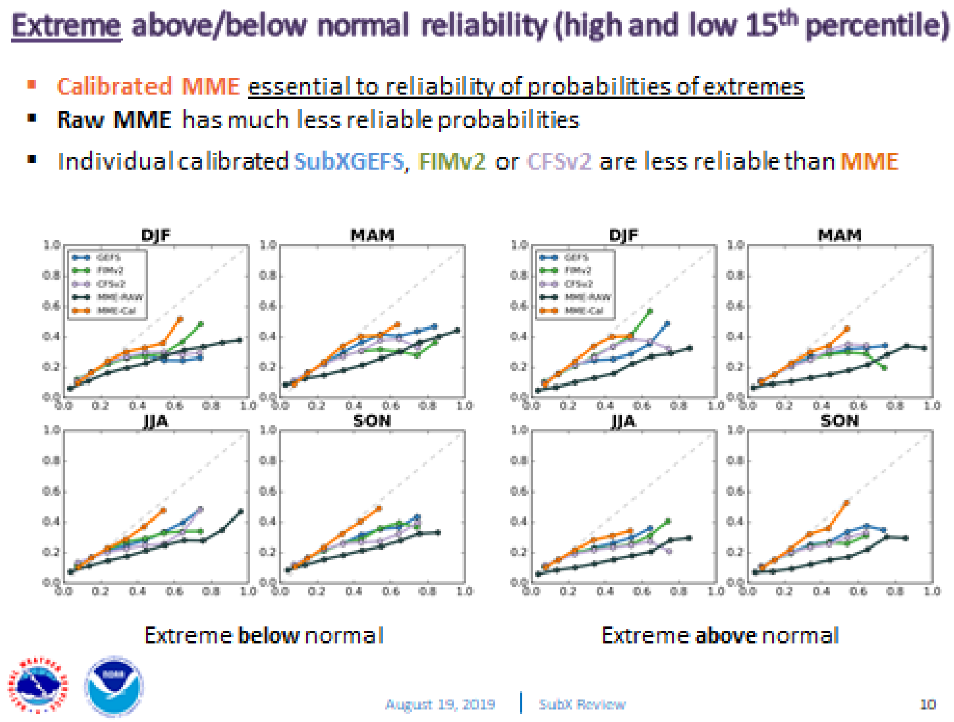

Figure 7. Reliability of prediction of extremes for different seasons using SubX. Source: Collins brief, slide 10.

Stringent quantification of the relative value of increased ensemble size versus increased model diversity on multi-model forecast fidelity remains unanswered. Other questions of specific relevance to CPC operational forecasting are listed in the document titled, “SubX Evaluation Criteria NOAA FY16 Climate Test Bed”, which provides potential guidance for future pseudo-operational SubX operations.

4. *Enhancing communications between operational forecasts and the model forecast producers*
Access to and allocation of NOAA high-performance computing (HPC) resources can be difficult for academic researchers. SubX, like NMME, provides forecast datasets to a wider research community by leveraging external computing and dissemination resources. This structure helps move research forward while also providing a broader set of data to operational forecast centers.

A direct impact was demonstrated by Collins who related how the Friday weeks 3-4 forecast discussions are informed by SubX forecasts, and specifically what contributions they make to the process of producing weekly outlooks. SubX personnel participate in a weekly forecast discussion, which is part of the official, human-in-the-loop forecast process at the CPC. During the project and the review, recommendations were made on how SubX could further enhance support of CPC operations, which are described below.

The value and impact of land surface initialization was demonstrated by the presentation of Koster in the SubX models, and in many other studies. Koster showed evidence of the role of land surface initialization fidelity on forecast skill – results by DeAngelis and Wang for the 2012 Midwest drought and heatwave as forecast by 4 SubX models at 3-4 week lead (Fig. 8). NCEP lags behind other nations in terms of land surface data assimilation, particularly of satellite soil moisture retrievals. As land surface initialization has been demonstrated to have its greatest impact in the subseasonal time range[6], SubX is an ideal platform to explore this improvement, which also projects on the question of forecasts of opportunity for extremes.

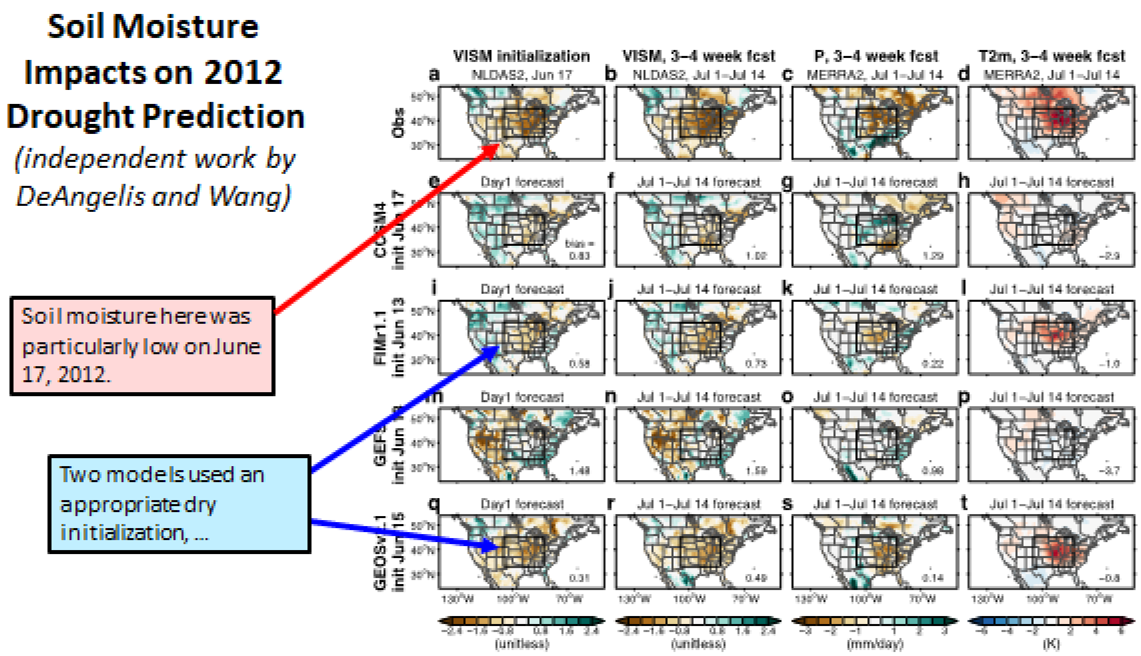

Figure 8. Impacts of land moisture initialization on drought prediction. Source: Koster brief, slide 9.

5. *Evaluating the skill of individual model systems*

The SubX protocol greatly facilitates comparison of participating models, and results were also shown for this exercise. One drawback of the current SubX set-up is that although the submission date/time for forecasts is fixed on Thursday mornings for all models, the initialization dates and protocols vary widely among models, meaning some models submit forecasts initialized closer to the valid period than others, making a fair comparison of model skill difficult.

Meanwhile, centralized calculation of model metrics (now at CPC) ensures fair comparison and evaluation, LaJoie and Pegion collaborated to evaluate the predictive skill of the MJO in the SubX models, with evaluation algorithms being initially coded by Pegion and ported to CPC infrastructure by LaJoie. Continued work in concert with investigators, additional methods to evaluate and compare the model forecasts can be ported to CPC's computational infrastructure.

6. *Participation in the NOAA/MAPP S2S Task Force*

The NOAA/MAPP S2S Task Force has been, from its inception, a combination of the S2S project investigators and SubX investigators. Dr. Kathy Pegion of SubX serves as one of the co-leads of the S2S task force, representing SubX.

[6] Dirmeyer, P. A., S. Halder, and R. Bombardi, 2018: On the harvest of predictability from land states in a global forecast model. J. Geophys. Res., 123, 13,111-13,127, doi: 10.1029/2018JD029103

III. Potential Future for SubX

The pseudo-operational forecast suite and multi-model products of SubX are already positively impacting operational subseasonal forecasts at CPC, as documented above. The future of SubX was discussed from multiple angles including potential appropriate support program(s) within NOAA, changes to governance, modifications of the experiment protocols, and new potential scientific goals. Cultural assumptions of the program were acknowledged, with the intent to clarify program intent and organization.

Support Program

SubX was originally funded under the NOAA CPO CTB program, which is intended for projects destined for transition to operations or applications, and therefore CTB projects are normally small grants to accelerate the transition of high readiness-level research to operations. SubX grew out of a much more mature project, NMME, which provided a springboard to SubX; however, subseasonal prediction with dynamical models has a much shorter history and SubX modeling capabilities have been tested for a shorter time. Many pieces of the project are not ready for transition to operations, nor is there any intention to formally transition the participating non-NOAA models to NOAA operations.

The willingness of participating centers/institutes/agencies to provide data past the original end point allowed the project output to continue on a minimal-funding basis, and is notable for demonstrating a capability-multiplication through interagency and academic cooperation. While the contributing modeling centers have found benefits to running an operational-type cadence on their own, to contribute to SubX and to further their own missions, NWS has neither personnel nor computing capacity to transition the participating models into the NWS operational infrastructure. Thus, CTB is arguably not the ideal place to continue to foster a pseudo-operational interagency/academic partnership effort for subseasonal prediction.

Additionally, ongoing support to academic partners via competitive and non-competitive funding should be present in any future funding scenario. Competitive funding can increase interaction with the weather enterprise to foster and incorporate the best ideas, while non-competitive funding can target specific needs already identified by operational centers. This non-competitive funding typically represents a small percentage of research funding.

Therefore, **the panel recommends an interagency coordination organization such as the National Earth System Prediction Capability (National ESPC) or the Office of the Federal Coordinator for Meteorology (OFCM) provide governance and coordination**; the participating agencies have so far been willing to provide model output to support the common goals. The utility of an interagency coordination body is its ability to inform relevant agencies of emerging issues, and optimally leverage research resources for improvements.

Governance

This SubX review panel is a mix of NOAA and external personnel with interest and expertise in various aspects of subseasonal prediction with dynamical models, and drew from agencies and academia. It provides a good paradigm for future review panels. NMME, which serves in many ways as a template for SubX, has a standing executive committee that provides annual review. **As SubX has a similar pseudo-operational component, a similar standing review mechanism would be effective.** NOAA operational and research interests as well as academic research and modeling centers external to NOAA should have representation in such a review body. A SubX governing body could specify the protocols for inclusion in a potential follow-on experiment.

The governing body and the supporting program office should work together to strengthen and formalize the linkages among operational and research focused modeling centers. Fostering the linkages can result in faster transition from research to operations, enhanced information flow to the research community about discovered issues, and potential new uses for the data.

Protocol Suggestions

Currently forecasts are due 6AM on Thursdays – the question was raised whether this is the best day of the week, even for the CPC subseasonal outlooks. Some current users of CPC's week 3-4 outlooks have indicated that they would find a forecast issued earlier in the week more desirable and potentially more useful. So a consideration for **either daily or multiple deliveries per week should be considered**.

The 32-days minimum forecast period could be extended to fully accommodate week-4 forecasts when IC dates lag forecast delivery day. Forty-eight days would accommodate prediction products through week 6, while model runs out 60-62 days are common for WMO's International S2S Prediction Project models[7].

Having models better synchronize their initialization dates and ensemble structuring would aid construction of multi-model ensembles, and facilitate forecast comparisons. Currently, models that routinely initialize several days before the weekly deadline deliver suboptimal forecasts. It is clear this must degrade multi-model skill as it degrades individual model skill, but it remains to be determined how much the multi-model ensemble skill suffers through a delay.

An infrastructure that can easily accommodate updated model versions or new models, including timely generation of corresponding hindcasts (crucial for determination of anomalies), would facilitate evolutionary improvement of the multi-model forecasts as individual models are improved. The ability to provide a continuous record for any real-time portion of the experiment should also be a criteria for inclusion.

Currently there are two categories of users of SubX forecast and hindcast data: those who use the data in a forecast context, and those who use it as a research data set. The two are distinct but not without overlapping interests. **New metrics could be added that inform decisions by both communities**. The International S2S Prediction Project has developed metrics for model forecast comparison and validation that are directly applicable to SubX. In addition to metrics based solely on forecast skill, **process-based metrics could help illuminate addressable model shortcomings**. Some communities, such as the ocean-atmosphere and land-atmosphere research communities, have

[7] http://s2sprediction.net/

developed and tested sets of coupled metrics that can inform both model development and the question of predicting predictability, i.e., identifying promising situations for forecasts of opportunity.

LaJoie showed that model diversity was a clear contributor to increased skill, especially when the full SubX multi-model suite was included. LaJoie stated that this seemed to be more important than increasing the ensemble size of individual models that are part of a multi-model ensemble. Thus, focus on maintaining a large selection of models appears paramount. It is clear that for hindcasts, small ensembles (3 or 4 members) are sufficient for establishing a stable climatology for determining anomalies; extending the hindcast record back in time would likely be more useful than increasing hindcast ensemble size.

The SubX team has already generated and agreed to a list of **additional variables** that would be good to deliver (Appendix B). Complete delivery of all the variables requested by, and agreed to, by the current SubX projects should be a high priority in any continuation or subsequent rounds of experiments.

Potential Goals

The FY16 funded program under review had six objectives, with some specifically aimed at opening the data up to the broadest range of stakeholders. **A second round of experiments could be more specific and focused around science and implementation objectives.** The objectives could be possibly quantitative or requirements driven, as some of the models run for the first round are likely to be implemented in operations, on each operational contributing center's infrastructure, by the start of a second experiment, and most definitely informed by the volume of research conducted and ongoing with the existing datasets.

SubX naturally intersects with NMME via its common heritage and design elements. They fill adjacent/overlapping bands in the spectrum from weather to climate forecasting. The program designs and overlap times (monthly from one to twelve months for NMME, weekly from two to six weeks for SubX) were created through other program constraints, rather than optimal prediction skill, which remains a research issue. It should be discussed in any future governance panel **whether it makes sense in the long run to merge the SubX and NMME efforts into a single forecast suite, with outputs and timing tailored to address both subseasonal and seasonal forecast needs.** That question was not pursued as part of this review.

SubX also has clear similarities with the International S2S Prediction Project coordinated via the WMO (s2sprediction.net). There are three main differences between SubX and the WMO S2S Project. First, SubX forecasts are released near-real-time without embargoes, whereas most S2S forecasts are not accessible for several weeks. Real-time delivery and release offers great potential for expanding the applications of the data. Second, SubX is designed for producing multi-model forecasts (common data resolutions, forecast delivery dates, etc.), whereas the WMO S2S Project collects operational forecasts as they come and is only beginning its Real-Time Pilot Initiative for 1 Nov 2019. Stricter protocols for a second round of experiments could remove uncertainty based on differing model protocols, rather than allowing delivery schedules of operational centers to override protocols for an optimized experiment. Third, the WMO S2S Project includes only operational forecast models, and discussions to include non-operational models have consistently met with resistance. There is already good coordination between the two projects, and the IRI is a principal bridge between the projects. However, SubX fills a niche in both S2S prediction research and multi-model forecasting and supplements the International S2S Prediction Project. **Closer connection with the WMO S2S Project might benefit both**.

SubX would be an ideal proving ground for the nascent Unified Forecast System (UFS). There was hearty agreement that **beginning testing of the new coupled UFS in such a pseudo-operational setting would help ensure a better model at time of delivery, informing model development at an early stage.** Inclusion of the beta-version of UFS's Seasonal Forecast System (SFS) in the multi-model ensemble would be entirely consistent with the presence of other non-operational models in SubX. UFS will be the common global model for NOAA for years to come, so its inclusion in SubX is every bit as important for UFS as for SubX – having the eyes on the community on SubX performance and behavior before and after it becomes operational will contribute greatly to the quality of the model. Additionally, leveraging the outcomes of the work accomplished as part of the Earth Prediction Innovation Center (EPIC) should be a major priority of any future subseasonal prediction experiment.

Another potential goal could encompass the **use of machine learning to glean more information from model datasets**. Suitable learning algorithms using appropriate reforecast datasets might glean additional prediction skill from existing models for various regions. A potential application for these methods could be the real-time identification of forecasts of high opportunity where enhanced skill can be expected based on physical principles, is a promising avenue to increased usefulness and value for users (akin to the distinction between "equal chances" and "near normal" in current outlooks).

A future governance panel could look at **options to streamline data flow and coordinate with stakeholders.** The rigorous and quantitative use of forecast information in the subseasonal timeframe is in its relative infancy, so engagement with stakeholders could drive future research priorities and changes to operational product suites. This could be helped by collecting, via voluntary form submission prompted by data download, information about the people and organizations that use SubX data and the applications built on the data.

The panel recommends the **establishment of an official list of peer-reviewed publications citing the SubX dataset DOI and/or voluntarily reporting to use the SubX dataset in a significant manner**. The current listing established by the funded investigators could be transitioned to a NOAA library. Any future campaign should include support for the already established channels of communication that detail changes to model configurations.

Appendix A: SubX Citations

Paper Submitted, In Press or Appeared

Diro, G. T., and H. Lin, 2019: Subseasonal forecast skill of snow water equivalent and its link to surface air temperature in three SubX models. Weather and Forecasting, (submitted).

Guan, H., Zhu, Y., Sinsky, E., et al. (2018b) Systematic Error Analysis and Calibration of 2-m Temperature for the NCEP GEFS Reforecast of the Subseasonal Experiment (SubX) Project. Wea Forecasting 34:361–376. doi: 10.1175/WAF-D-18-0100.1

Janiga, M.A., Schreck C.J., Ridout, J.A., et al. (2018) Subseasonal Forecasts of Convectively Coupled Equatorial Waves and the MJO: Activity and Predictive Skill. Mon Wea Rev 146:2337–2360. doi: 10.1175/MWR-D-17-0261.1

Keyel, A.C., Elison Timm, O., Backenson, P.B., Prussing, C., Quinones, S., McDonough, K.A., et al. (2019) Seasonal temperatures and hydrological conditions improve the prediction of West Nile virus infection rates in Culex mosquitoes and human case counts in New York and Connecticut. PLoS ONE 14(6): e0217854. https://doi.org/10.1371/journal.pone.0217854

Kim, H., Janiga, M.A., Pegion, K. MJO propagation processes and mean biases in the SubX and S2S reforecasts (In Revision). Journal of Geophysical Research Atmospheres

Kim, H., Vitart, F., Waliser, D.E. (2018b) Prediction of the Madden–Julian Oscillation: A Review. J Climate 31:9425–9443. doi: 10.1175/JCLI-D-18-0210.1

Li, W., Y. Zhu, X. Zhou, D. Hou, E. Sinsky, C. Melhauser, M. Peña, H. Guan, R. Wobus, 2018: Evaluating the MJO Prediction skill from Different Configurations of NCEP GEFS Extended Forecast. *Climate Dynamics*, 52, 4923–4936

Lukens, K.E., Berbery, E.H. (2019) Winter Storm Tracks and Related Weather in the NCEP Climate Forecast System Weeks 3–4 Reforecasts for North America. Wea Forecasting 34:751–772. doi: 10.1175/WAF-D-18-0113.1

Mariotti, A., Ruti, P.M., Rixen, M. (2018) Progress in subseasonal to seasonal prediction through a joint weather and climate community effort. npj Climate and Atmospheric Science 1:4. doi: 10.1038/s41612-018-0014-z

Mariotti, A., E. A. Barnes, E. Chang, A. Lang, K. Pegion, D. Barrie, 2018: Bridging the weather-to-climate prediction gap: progress by the NOAA S2S Prediction Task Force, EOS, 100, https://doi.org/10.1029/2019EO115819.

Merryfield, B. and Co-authors, Current and Emerging developments in subseasonal to decadal predictions, *submitted to BAMS*

Pegion K, et al. The Subseasonal Experiment (SubX):A multi-model subseasonal prediction experiment (Accepted). Bulletin of the American Meteorological Society.

Poan, E., and H. Lin, 2019: Subseasonal Forecast Skill of the PNA and NAO in boreal winter by the Sub-seasonal Experiment (SubX) Models (submitted).

Ross, A., C. Stock, K. W. Dixon, M. A. M. Friedrichs, R. R. Hood, M. Li, K. Pegion, V. Saba, P. St-Laurent, and G. Vecchi, Estuarine forecasts at weather to subseasonal timescales, submitted to JGR-Oceans

Sun, S., Bleck, R., Benjamin, S.G., et al. (2018a) Subseasonal Forecasting with an Icosahedral, Vertically Quasi-Lagrangian Coupled Model. Part I: Model Overview and Evaluation of Systematic Errors. Mon Wea Rev 146:1601–1617. doi: 10.1175/MWR-D-18-0006.1

Sun, S., Green, B.W., Bleck, R., Benjamin, S.G. (2018b) Subseasonal Forecasting with an Icosahedral, Vertically Quasi-Lagrangian Coupled Model. Part II: Probabilistic and Deterministic Forecast Skill. Mon Wea Rev 146:1619–1639. doi: 10.1175/MWR-D-18-0007.1

Tippett, M.K., Koshak, W.J. (2018) A Baseline for the Predictability of U.S. Cloud-to-Ground Lightning. Geophysical Research Letters 45:10,719-10,728. doi: 10.1029/2018GL079750

Wayand, N.E., Bitz, C.M., Blanchard-Wrigglesworth, E. (2019) A Year-Round Subseasonal-to-Seasonal Sea Ice Prediction Portal. Geophysical Research Letters 46:3298–3307. doi: 10.1029/2018GL081565

Xiang, B., Lin, S.-J., Zhao, M., et al. (2019) Subseasonal Week 3–5 Surface Air Temperature Prediction During Boreal Wintertime in a GFDL Model. Geophysical Research Letters 46:416–425. doi: 10.1029/2018GL081314

Zhu, Y., Zhou, X., Li, W., et al. (2018) Toward the improvement of subseasonal prediction in the National Centers for Environmental Prediction global ensemble forecast system. Journal of Geophysical Research: Atmospheres 123:6732–6745.

Zhu, Y., X. Zhou, M. Peña, W. Li, C. Melhauser, D. Hou, 2017: Impact of sea surface temperature forcing on weeks 3 & 4 prediction skill in the NCEP global ensemble forecasting system. *Weather Forecast*, 32:2159–2174

Zhu, Y., W. Li, X. Zhou and D. Hou, 2018: Stochastic Representation of NCEP GEFS to Improve Sub-seasonal Forecast. Book chapter.

Presentations

Achuthavarier, D, R. Koster, S. Schubert, J. Marshak and A. Molod, 2019: MJO Teleconnection Signals and Prediction Skill in the NASA GEOS Subseasonal Reforecasts, Topics in Atmospheric and Oceanic Sciences, School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY. (Oral)

Achuthavarier, D., R. Koster, S. Schubert, J. Marshak and A. Molod, 2018: MJO Teleconnection Signals in the NASA GEOS-5 Subseasonal Reforecasts, Subseasonal to Seasonal Prediction of Weather and Climate, Washington D.C., Amer. Geophy. Union Fall Meeting. (Oral)

Achuthavarier, D., R. Koster, J. Marshak, S. Schubert and A. Molod, 2018: prediction skill of the MJO teleconnection signals in the NASA GEOS-5 subseasonal reforecasts, Second International Conference on Subseasonal to Seasonal Prediction, Boulder, CO. World Climate Research Program. (Poster)

Achuthavarier, D., R. Koster, J. Marshak, S. Schubert and A. Molod, 2018: Prediction and predictability of the Madden Julian Oscillation in the NASA GEOS-5 seasonal-to-subseasonal system, Sixth Symposium on Prediction of the Madden-Julian Oscillation: Processes, Prediction and Impact, Austin, TX, Amer. Meteor. Soc. (Poster)

Achuthavarier, D., R. Koster, J. Marshak, S. Schubert and A. Molod, 2017: Prediction and predictability of the Madden Julian Oscillation in the NASA GEOS-5 seasonal-to-subseasonal system, National Multi-model Ensemble Experiment (NMME) and Subseasonal Experiment (SubX) Science Meeting, NOAA Center for Weather and Climate Prediction, College Park, MD. (Poster)

Bell, M.A., Robertson, A.W., Tippett, M.K., et al. (2018) The IRI Climate Data Library as a Platform to Archive, Analyze, and Distribute SubX Project Data

Benjamin, S., Sun, S., Grell, G., et al. (2017) Improved subseasonal prediction with advanced coupled models including the 30km FIM-HYCOM coupled model. p 11097

Collins, D.C., Jia, L., Lajoie, E., et al. (2018) Evaluating the Potential for Forecasts of Opportunity and of Extremes on Subseasonal Timescales using SubX

Deangelis, A.M., Wang, H., Koster, R.D., et al. (2018) Prediction Skill of US Flash Droughts in Subseasonal Experiment (SubX) Model Hindcasts

DelSole, T., “Recent Developments in Forecast Quality Assessment", 2018 Second International Conference on Seasonal to Decadal Prediction (S2D), Boulder CO, September 19, 2018.

Green, B., Sun, S., Benjamin, S., et al. (2017a) The FIM-iHYCOM Model in SubX: Evaluation of Subseasonal Errors and Variability

Green, B., Sun, S., Grell, G., Benjamin, S. (2018) Subseasonal Errors and Skill in FIM-iHYCOM Model Precipitation: Sensitivity to Convective Parameterization and Model Resolution

Guan, H., Zhu, Y., Sinsky, E., et al. (2018a) Systematic Error Analysis and Calibration for the NCEP GEFS Reforecast of SubX Project. p 3414

Janiga, M. A., et al. (2018): Using the S2S database to evaluate the performance of the Navy Earth System Prediction (ESPC) ensemble.

Kim, H., Janiga, M., Achuthavarier, D., Pegion, K. (2018a) Process-based MJO evaluation in SubX and S2S hindcasts

Kim, H., Matthew A. Janiga, Deepthi Achuthavarier, Kathy Pegion, "Process-based MJO hindcast evaluation in SubX", International Conference on Subseasonal to Decadal Prediction (S2D)", NCAR, Boulder, CO, USA, September 14-22, 2018

Kim, H., "Process-based MJO hindcast evaluation in S2S and SubX models", Intraseasonal oscillations and extended-range prediction workshop, Taiwan, 10/18/2018

Kim, H., Matthew A. Janiga, Deepthi Achuthavarier, Kathy Pegion, "Process-based MJO evaluation in SubX and S2S hindcasts", AGU Fall 2018.

Kirtman, B. P., 2019: The NMME and SubX projects: Status, Challenges, and Opportunities. AMS Annual Meeting, Phoenix Az, Jan. 2019 (intived).

Kirtman, B. P., 2019: The Sub-seasonal prediction experiment. AMS Annual Meeting, Phoenix Az, Jan. 2019 (intived).

Kirtman, B. P., 2018: The Sub-seasonal prediction experiment. AGU Fall Meeting, Washington, D.C., Dec 10-14, 2018.

Lajoie, E., Collins, D. (2018) Evaluation of the SubX Reforecast Skill under Real-time Considerations

L'Heureux, M. and M. K. Tippett, Skill of California Precipitation on S2S timescales, AGU Fall Meeting, Washington, D.C., Dec 10-14, 2018.

Mariotti, A., Barrie, D., Archambault, H. (2017) The Subseasonal Experiment (SubX) to Advance National Weather Service Predictions for Weeks 3–4

O'Connor, A., Bell, R., Kirtman, B., Gorman, J. (2017) Long-range forecasting using COMPASS machine learning. p 4

Pegion, K., Earth System Prediction with the Subseasonal Experiment (SubX), NCAR CESM Workshop, CESM Climate Variability and Change Working Group, Boulder, CO, June 2019.

Pegion, K., Bridging the Gap between Weather and Climate Prediction using Multi-model Ensembles, University of Colorado, Boulder, CO, Feb 2019

Pegion, K., The Subseasonal Experiment (SubX), Dynamics Group Meeting/CSU, Fort Collins, CO Nov 2018

Pegion, K., The Subseasonal Experiment (SubX), NCAR/CGD Seminar, Boulder, CO, Nov 2018

Pegion, K., The Subseasonal Experiment (SubX), NOAA/ESRL/GSD Seminar, Boulder, CO, Oct 2018

Pegion, K., The Subseasonal Experiment (SubX), NOAA/Climate Diagnostic and Prediction Workshop, Santa Barbara, CA, Oct 2018

Pegion, K., The Subseasonal Experiment (SubX), Second International Conference on Subseasonal to Seasonal Prediction (S2S), Boulder, CO, Sep 2018

Pegion, K., Metrics for S2S: Examples from SubX, National Earth System Prediction Capability Workshop on Metrics, Postprocessing, and Products for Subseasonal to Seasonal Timeframes, College Park, MD, Feb 2018

Pegion, K., The Subseasonal Experiment, NOAA/ESRL, Boulder, CO, Jan 2018

Robertson, A.W., N. Vigaud, J. Yuan, M. Tippett, D. Collins, 2018: Calibrated Multi-model Probabilistic Sub-seasonal Forecasts Based on SubX and S2S Models. AGU Fall Meeting, Dec 12, 2018.

Robertson, A.W., Vigaud, N., Yuan, J., et al. (2018) Calibrated Multi-model Probabilistic Sub-seasonal Forecasts Based on SubX Models

Robertson, A.W, 2018: “Connecting Climate, Weather, and Water: Status, Challenges, and Needs for Subseasonal-to-Seasonal Forecasting for Water Use and Management”, National Academy Of Sciences, Washington DC, November 30, 2018. Keynote Speaker.

Strazzo, S., Collins, D.C., Schepen, A., et al. (2018) Harnessing forecast skill from SubX and the MJO for sub-seasonal prediction of North American climate

Sun. S. and K. Pegion, Real-time Subseasonal Forecast with SubX, American Meteorological Society, Annual Meeting, Phoenix, AZ, Jan 2019.

Wang, H., Schubert, S., Koster, R., et al. (2018) Causes and Predictability of Drought in the US High Plains

Zhou, J., DeWitt, D. (2019) NOAA’s MAPP-CTB Projects Update: Community R2O Contributions to the Improvement of Operational S2S Climate Prediction. Climate Prediction S&T Digest 180

Zhu, Yuejian, Wei Li, Eric Sinsky, Hong Guan, Xiaqiong Zhou and Bing Fu, 2019: An Assessment of Predictability and Prediction of NCEP GEFS for Subseasonal Forecast. S2S workshop, Reading, UK.

Zhu, Y., W Li, E Sinsky, H Guan, X Zhou and B Fu, 2018: An Investigation of Prediction and Predictability of NCEP Global Ensemble Forecast System (GEFS). Science and Technology Infusion Climate Bulletin, NOAA's National Weather Service, 43rd NOAA Annual Climate Diagnostics and Prediction Workshop

Appendix B: Supplemental Output Variables

PRIORITY III, IV, and Optional Variables

VARIABLE	LEVEL	FREQUENCY
Geopotential Height	850 hPa	Average at 0,6,12,18Z
Temperature	100 hPa	Daily average
Zonal velocity	10, 30, 50 hPa	Average at 0,6,12,18Z
Meridional velocity	10, 30, 50 hPa	Average at 0,6,12,18Z
Geopotential Height	10, 30, 50 hPa	Average at 0,6,12,18Z
Temperature	10, 30, 50 hPa	Average at 0,6,12,18Z
Root zone soil moisture expressed in terms of volumetric soil moisture content as volume of water per volume of soil	subsurface	Accumulated every 24 hrs
Vertically integrated (sfc to 300 hPa) zonal moisture flux in the atmosphere (for ARs)	multiple	Accumulated every 24 hrs
Vertically integrated (sfc to 300 hPa) meridional moisture flux in the atmosphere (for ARs)	multiple	Accumulated every 24hrs
Vertically integrated (sfc to 100 hPa) zonal moisture flux in the atmosphere (for MJO moisture budget)		Accumulated every 24hrs
Vertically integrated (sfc to 100 hPa) meridional moisture flux in the atmosphere (for MJO moisture budget)		Accumulated every 24hrs
Vertically integrated (sfc to 100 hPa) precipitable water(for MJO moisture budget)		Accumulated every 24hrs
TOA net shortwave radiation		
Vertically integrated (100 to 100 hPa) dry static energy		
Vertically integrated (100 to 100 hPa) dry static energy advection by u and v winds		
LW component of net surface radiation (MJO community)		
SW component of net surface radiation		

(MJO community)		
Common indices of some particular phenomena		
Wave heights (optional)		Daily Average 00Z-00Z

Appendix C: Review Agenda and Speakers

Review panel members:
Duane Waliser (NASA/JPL, chair), David DeWitt (CPC), Vijay Tallapragada (EMC), Matthew Rosencrans (CTB), Jessie Carman (OWAQ), Paul Dirmeyer (GMU/COLA), Joshua Cossuth (ONR)

SubX Review Agenda		
Time	***Topic***	***Presenter(s)/Parties (Potential)***
08:00 - 08:30	Registration and Continental Breakfast	
Introduction		
08:30 - 09:00	SubX programmatic overview and guidance for the review	Jessie Carman
Morning Session		
09:00 - 10:15	Report of the SubX Project	Ben Kirtman/Kathy Pegion
10:15 - 10:30	Discussion	All
10:30 - 10:50	Break	
10:50 - 11:30	CPC Evaluation and Metrics	Dan Collins, Emerson LaJoie
11:30 - 12:00	Reports from selected participating modeling centers	Neil Barton (Navy ESPC)
		Randy Koster (NASA)
		Yuejian Zhu (EMC)
12:00 - 13:00	Lunch	
Afternoon Session		
13:00 - 13:30	Remarks by NOAA-OWAQ Management	Bill Lapenta, Jessie Carman
13:30 - 14:00	SubX Partner Future Ideas	Ben Kirtman, Kathy Pegion
14:00 - 15:00	General Discussion/Recommendations	All
15:00 - 15:20	Break/Adjourn	

Closed Session		
15:20 - 17:00	Review Panel Discussion	Review Panel Members Only