Project Final Report

Submitted to

National Oceanic and Atmospheric Administration
Climate Program Office
Modelling, Analysis, Predictions, and Projections (MAPP) Program
(Dr. Annarita Mariotti, Program Officer)

Research project
Exploring best practice procedures for optimal use of climate forecast for regional hydrological applications
(Grant Number: NA12OAR4310081, GC Identifier: GC12-682)
)

October 15, 2017

Principal Investigator:
Lifeng Luo
Department of Geography, Environment, and Spatial Sciences
Michigan State University
East Lansing, MI 48824
Email: lluo@msu.edu

Pang-Ning Tan
Department of Computer Science and Engineering
Michigan State University
East Lansing, MI 48824
Email: ptan@cse.msu.edu

INTRODUCTION

Decisions regarding water resource management, agricultural practice, and energy allocation often require information about future climate conditions weeks to months in advance. Skillful and reliable seasonal climate prediction can significantly facilitate and benefit the decision making process. However, it is a challenge to make skillful predictions about the climate system at these time scales because the processes that contribute to the seasonal predictability are not fully understood thus not adequately represented in climate models. Research is needed to assess the current prediction skills from the state-of-the-art climate forecast systems, and to develop "best practice" procedures for optimal use of climate forecasts in applications that are directly relevant to decision making. This project was an effort to address various issues related to seasonal hydrological prediction.

The overall goal of the proposed research is to advance the nation's capability in seasonal drought prediction through improved climate model forecasting and advanced post-processing techniques. To achieve this goal, the project has the following specific objectives:

1. Evaluate the state-of-the-art climate forecast systems, quantify their seasonal prediction quality, and assess factors that contribute to the prediction skill.
2. Assess the optimal choice of ensemble members and scales, and to develop best practice procedures for combining and post-processing multiple forecasts to achieve better forecast quality.
3. Demonstrate the usefulness of seasonal climate prediction and evaluate the new post-processing procedures with seasonal drought prediction.

The innovation of the proposed research is mainly reflected in the second activity. Our efforts focused on developing two innovative methods (multiscale Bayesian merging and structured output regression) in parallel to combine forecast information across multiple characteristic spatial and temporal scales. These methods will address outstanding issues like spatial and temporal dependence (or correlation structure) that is practically ignored when combining forecast members in an ensemble or multimodel ensemble system currently. It is our intention to compare and combine two methods that grew out of two research communities. These statistical methods have the potential to significant advance the seasonal climate forecast skills. We will demonstrate the improvement in prediction skills and usefulness of climate prediction in regional hydrological applications by performing seasonal drought forecast for selected drought events in the US using these new methods.

RESEARCH ACTIVITIES AND ACCOMPLISHMENTS

The major research activities during the project period are listed below followed by more details on the major findings and accomplishments with selected illustrations. Complete description of each study can be found in the published papers listed at the end of the report.

1. Establishing framework for drought prediction based on CFSv2

A seasonal drought prediction framework was established in this project. Figure 1 illustrates the overall data flow in the framework. The system leverages from previous research activities supported by MAPP and CTB programs, so the general concept is similar to previous studies. What is new in this framework are the following:

a) Daily fields from CFSv2 forecast are used.
Previous research always used monthly fields from global climate models for seasonal hydrological predictions. Using the Bayesian merging approach (Luo et al, 2007), posterior distributions of monthly temperature and precipitation are derived, then daily temperature and precipitation time series are created based historical time series and the posterior distributions of monthly quantities. Because of this, seasonal drought forecast can be issued only once a month from the beginning of each month. In many cases, this is not sufficient. In this project, our new approach uses daily temperature and precipitation from all CFSv2 forecast members. This allows us to temporally aggregate them into any spatial and temporal scales as needed, adding a level of flexibility.

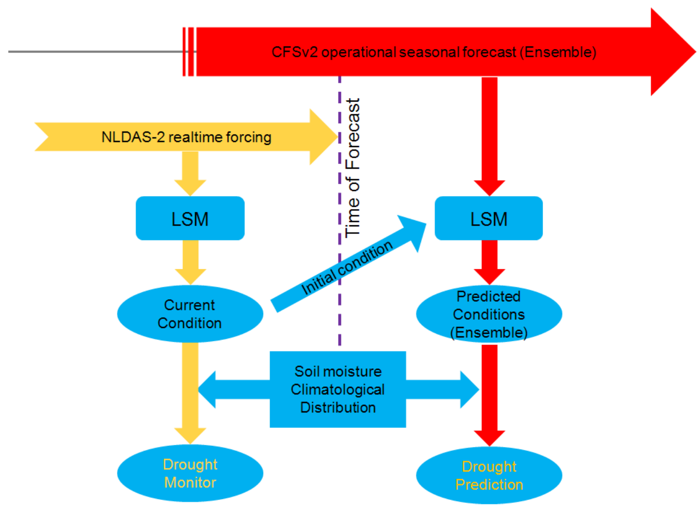


Figure 1: Schematic illustration of the seasonal drought prediction framework.

b) Posterior distributions are derived for various temporal scales.
In the operational configuration, there are 16 forecast members each day, some of which go 10 months into the future while others go only 45 days into the future. By taking the daily quantities, we can then aggregate them in various temporal scales, and then derive posterior distributions for each of these aggregates. For example, we can calculate average temperature (or precipitation) for a 90-day, 80-day, 70-day… 10-day periods starting from any day. These posterior distributions derived for these temporal aggregates are used to describe the ensemble forecast. This allows a seasonal drought prediction to be made from any day of the month.

c) Accumulation of information by continuous updating of posterior distributions
Luo et al (2007) use the climatological distribution as the prior in the Bayesian merging. In the newly developed system, the prior distribution is taken from the posterior distribution of an earlier merging. Because there are 16 CFSv2 forecast runs available each day from NCEP operation, the posterior distributions can be updated daily. This will help to improve the quality of the posterior distribution by accumulating forecast

information from the past. Figure 2 illustrates how an ensemble precipitation time series can be improved with these posterior distributions of temporal aggregates.

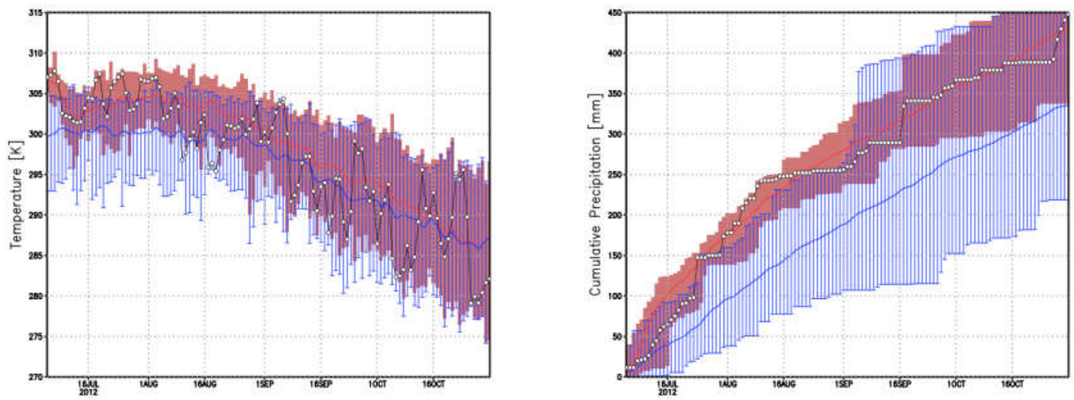


Figure 2: Temperature and precipitation time series generated from the Bayesian merging by combining information from multiple forecast runs from CFSv2 with continuous update. The black line with open circle is the observation, and the red line with shaded area is the forecast ensemble mean and spread. The blue line and associated error bars are for the climatological distributions.

d) The use of QPF for short-term rainfall

It is recognized that the initial conditions are important for the skill of seasonal hydrological predictions. It is easy to understand that short-term precipitation is just as important. To improve the precipitation forecast for the first seven days, we have been using the quantitative precipitation forecast (QPF) from the Hydrometeorological Prediction Center (HPC) at NCEP during the forecast forcing generation. The daily precipitation from QPF is blended into individual ensemble members with reduced weight as lead time increases. This not only allows for spread among members, but also significantly improves the skill of short-term drought prediction.

This experimental system has been running in realtime. Every night, CFSv2 realtime forecast are downloaded from NCEP ftp site. These forecast runs are preprocessed and a new set of posterior distributions are derived to describe the future temperature and precipitation time series. Every Sunday afternoon, a weekly nowcast run is performed from last week's drought condition to integrate forward by 7 days to produce a model-based drought analysis for the past Tuesday. The nowcast run takes realtime NLDAS-2 atmospheric forcing data and drives the VIC hydrological model to simulate soil moisture, thus providing us with a new initial condition for hydrological predictions. This is done to match the release of weekly drought maps by the U.S. Drought Monitor. The USDM release a weekly drought map on Thursday morning of the drought condition on Tuesday. Because of the current 4-day lag in the NLDAS-2 realtime forcing, our system can only produce the drought analysis for the same valid period on Sunday (3 days late). After the nowcast run is finished, an ensemble forecast is performed. This forecast uses both ESP method and the new Bayesian merging method to produce atmospheric forcing. Each method produce 33 ensemble members. Drought prediction based on the soil moisture forecast is produced. Predicted drought conditions are depicted on a weekly interval match the weekly drought monitoring. This makes it easier to visually compare drought predictions against drought monitor on a weekly basis for past forecasts.

Figure 3 shows the weekly drought prediction before the 2012 summer Central Plain drought issued in the middle of May 2015. The ESP-based forecast (last column) was not able to predict the intensified drying; while our forecast approach (third column) was able to produce a much consistent drying trend for the following number of weeks. More information and realtime drought prediction is available on our project web site at http://drought.geo.msu.edu/research/forecast/smi.weekly.php

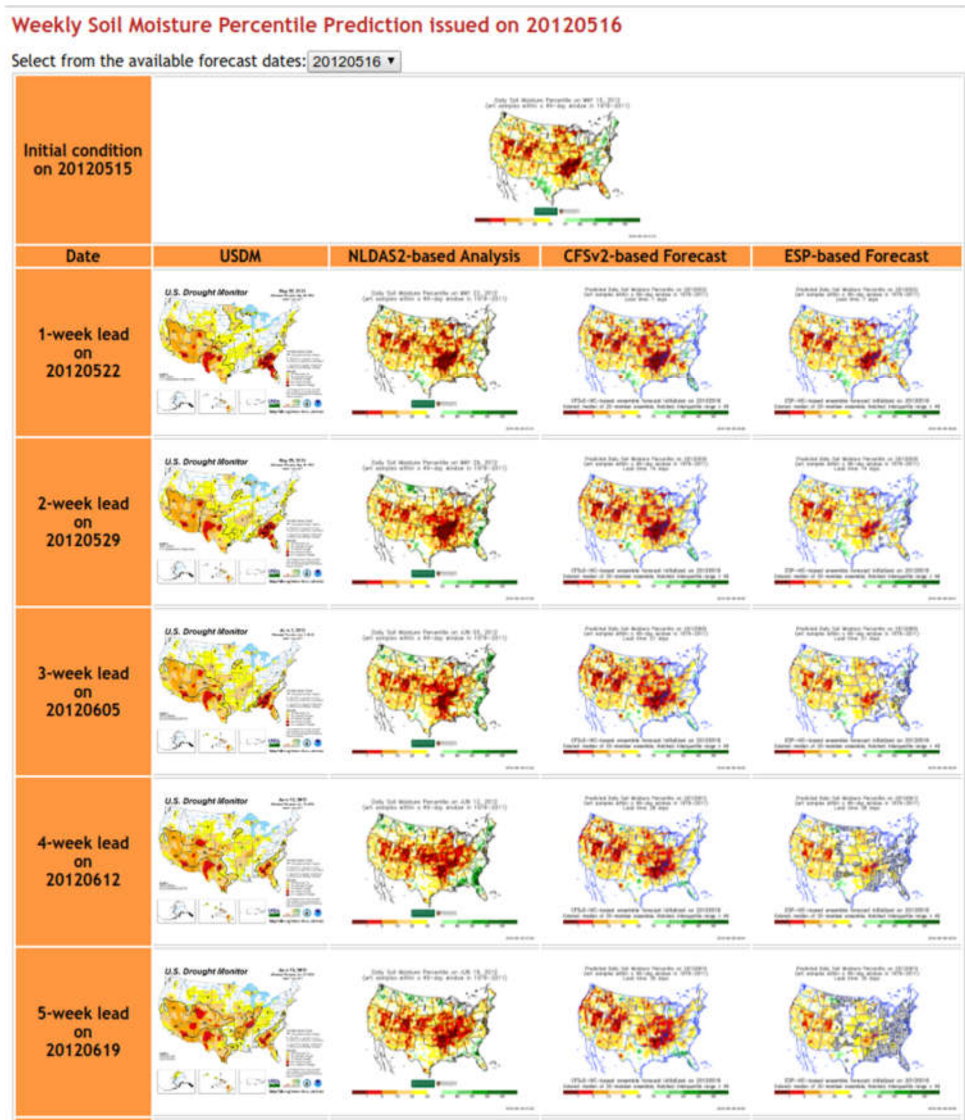


Figure 3: Weekly drought forecast issued on May 16, 2012 before the major drought over the Central US. The third column of maps are forecast from our system using the Bayesian merging approach, and the last column is the ESP approach. The first column are the maps from USDM, and the second column are from our nowcast system that can be compared with the forecasts.

2. *Improving the CFSv2 based drought prediction methodology*

The above results shown in Figure 2 and Figure 3 are the results of continuous improvement to the CFSv2 based drought prediction methodology, which was a major research activity during the project. We worked on improving the methodology, especially on the preprocessing of climate model forecast. We have developed programs to perform the multi-scale Bayesian merging procedure. The idea is to derive posterior distributions of precipitation and temperature at various spatial and temporal scales, and these posteriors distributions are then used to provide constraints on forecast ensembles based on historical records. Meanwhile, such a procedure will also allow continuous update of these posterior distributions, so useful forecast information can be accumulated to improve the skill and robustness of the seasonal forecast.

These forecasts are based on historical observation time series but guided by the probability distributions derived from CFSv2 model forecast. In particular, we have improved the transfer scheme that is used to convert precipitation probability distributions to normal distributions to enable Bayesian merging in the normal space. The earlier effort using logarithmic transformation can sometimes produce extreme rainfall conditions when the posterior distribution is transferred back to precipitation space. The current quantile mapping method seems to be more reliable to preserve the proper distributions without generating extreme precipitation values.

Figure 4 shows the prediction of drought area in the California/Nevada river forecast center region. Each week, we calculate the soil moisture percentile for each grid over the entire NLDAS domain, then counted the number of grid cells that has a value smaller than 20 in the RFC. This is shown in Figure 4 as the black line. Then every week, a 13-week forecast was made and the drought area within the RFC is also calculated. The solid color line is the ensemble mean and the two dashed lines are the upper and lower quartile of the ensemble spread. The forecasts made on 20131108(red) and 20131122 (green) were not able to predict the forthcoming drought. But the following forecasts, including 20131213 (blue), 20131227(magenta), 20140110(cyan), were all able to consistently predict the dramatic increase in drought area in the region with high confidence (narrow ensemble spread).

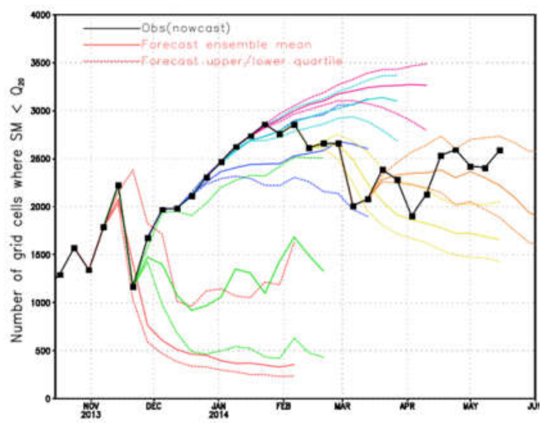


Figure 4 The number of grid cells within the California/Nevada RFC area where the soil moisture is lower than the 20th percentile of the historical distribution. A forecast is made every week with a lead-time up to 13 weeks. The system was able to predict the 2013-2014 drought in the region with a reasonable confidence.

The above result not only demonstrate the skill of the forecast system, but also support our argument that drought predictions need to be made more frequently than once-a-month as some quick changes can take place and affect the overall skill and usefulness of subseasonal and seasonal drought outlook.

3. Examining the effect of continuous update with Bayesian merging on precipitation and temperature forecast skills.

One major feature in the seasonal drought prediction system at MSU is to use Bayesian merging approach to derive the posterior distribution of precipitation and temperature at various temporal aggregates. Because the operational CFSv2 forecast are produced every 6hr cycle, these forecast runs with 9-month lead times will have slightly different lead time when considering a target forecast period. Should we treat these forecast runs equally as members of an ensemble, or should we treat them differently by considering their difference in lead time? During the project, we also tried to address this issue by examining how posterior distribution of temperature and precipitation for a target period would change when we do Bayesian update sequentially. Figure 5 illustrates how the monthly temperature distribution for a single point would change from the prior (climatology, left most box-whisker) to posterior distributions when one to six updates. The green horizontal line is the observation for this particular month (forecast target). In this case, it is clear that this month is on the warmer side of the climatological distribution. The posterior distribution from one update indeed shows much warmer, but a bit overconfident. As more CFSv2 forecast is used in the Bayesian merging update, the final posterior distribution after six updates is actually much closer to the target. This demonstrates the needs to do sequential update in the Bayesian framework to improve forecast quality. The results from this investigation will help us to decide how to best treat CFSv2 forecast members to achieve better forecast skill in seasonal hydrological predictions.

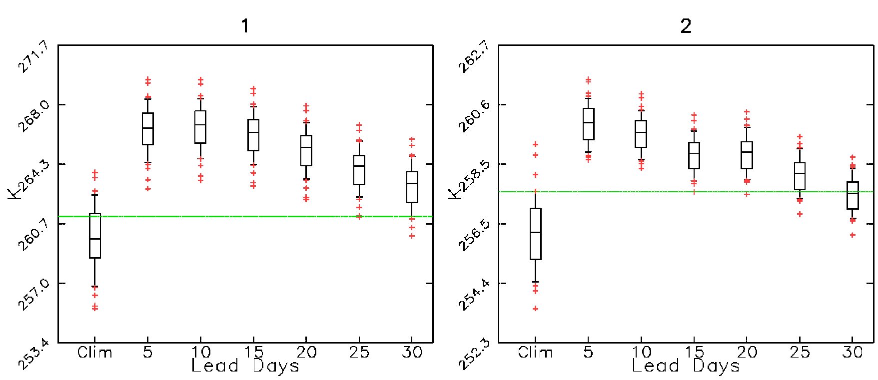


Figure 5 Climatological distribution and posterior distributions of monthly mean temperature for February at two given locations. The posterior distributions are produced with different number of updates using CFSv2 forecast from 5-days earlier to up to 30-days earlier.

4. Examining the CFSv2 hindcast skill dependence on spatial and temporal scales

A climate model's predictive skill for seasonal temperature and precipitation generally varies with multiple factors, such as location, lead-time, season, and temporal and spatial scales. To fully understand the potential and limitation of the Climate Forecast System version 2 (CFSv2) in predicting seasonal drought, we investigated how the seasonal drought predictive skill varies with such multiple factors. Six-month standardized precipitation index (SPI6) is used as the primary drought indicator to measure the medium-term meteorological drought. The predictive skill was then assessed by the correlation coefficient between observation-based SPI6 and CFSv2 forecast-based SPI6 at multiple spatial scales as well as multiple lead times during the period 1982-2008. Most of the analysis was carried out by a visiting PhD student from China through the collaboration with Chinese colleagues. This analysis helps us better characterize the capability of CFSv2 in seasonal climate forecast, thus to better utilize forecast information from such a system in drought prediction and water resource management.

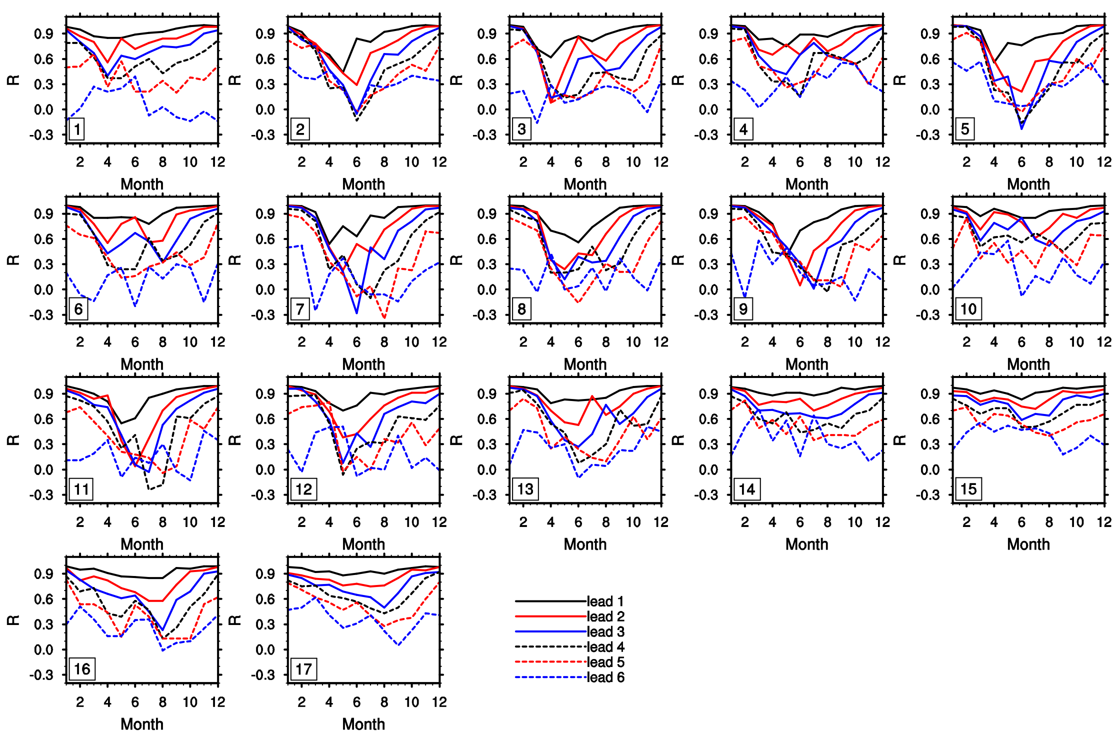


Figure 6: Forecast skill (measure by correlation) for SPI6 at 17 climate regions in China and how the forecast skill varies with lead time (different colors) and season.

Figure 6 shows how forecast skill (measured by correlation) for predicting SPI6 variation with lead time and season over 17 climate regions in China. SPI6 is a common index for drought monitoring and prediction, which is purely based on precipitation. The persistence characteristics of SPI are derived from its cumulative, time integrated design without seasonality of monthly precipitation. Depending on the way SPI6 is constructed (i.e., how many month of observation is included in the 6-month period), SPI6 at different lead time show different forecast skill, and the skill generally decays with increase of lead time. But in regions with strong seasonality of precipitation, the including of observed precipitation during the wet seasonal makes SPI6 more predictable with stronger persistence. Such a characteristic can be useful for using SPI6 in drought monitoring and prediction. A manuscript that describes this study in detail is currently under revision.

5. ***Reanalysis products for realtime drought monitoring***

During the project, we also investigated the feasibility of doing realtime drought monitoring using various global reanalysis products. This is especially important for regions without sufficient resources to observe rainfall and soil moisture conditions. This work was not initially proposed in the proposal, but we found it is necessary to carry out such an analysis, so the graduate students worked on this as a side project. We compared precipitation data from the ERA-Interim, CFSR, MERRA and GLDAS-2 projects. As an example, Figure 7 shows the correlation between CRU and these reanalysis in terms of SPI6. Smaller correlations are found in regions like Africa and South America where observational network tends to be less sufficient. For these regions, it is possible to develop a reanalyses-based drought monitoring system but using multiple reanalysis

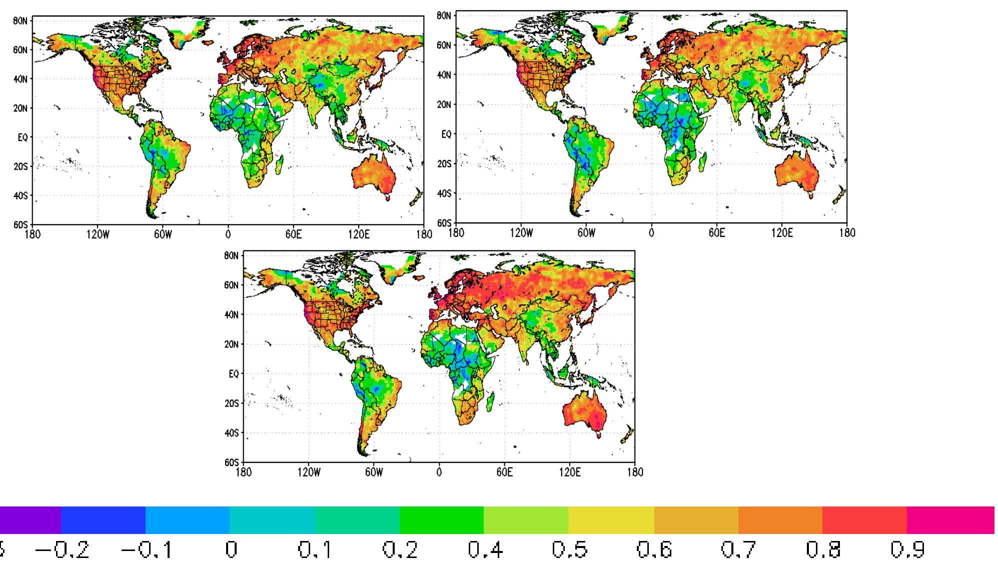


Figure 7: Correlation Coefficient (R) between CRU observations and each of the three reanalyses (CFSR, MERRA and ERA-Interim, left to right).

products is recommended. Further analysis also show that CFSR has inconsistent precipitation before and after around 1998.

6. ***Preprocessing and postprocessing methods using an online multi-task regression approach***

One of the component of the proposed study was to explore the usefulness of machine learning methods, in particular, multitask learning (MTL) methods, in dealing with prediction problems. Our effort on this has led to the development of three algorithms that can be used for preprocessing climate forecast information (e.g., downscaling of temperature and precipitation from climate models) and postprocessing of ensemble predictions (e.g., weighting schemes for ensemble members to achieve best prediction). This section summarizes the development of these algorithms and their evaluation.

ORION algorithm

We have developed a novel method for postprocessing of ensemble forecast using an online multi-task regression approach (see Figure 8). Current postprocessing methods, which are often based on using the ensemble mean or median, assume each ensemble member is equally likely in the probabilistic forecast. Weighing the ensemble members equally is not desirable as some members might be more skillful than others, a specific problem in our ensemble hydrological prediction system that uses historical records as the basis for individual ensemble members.

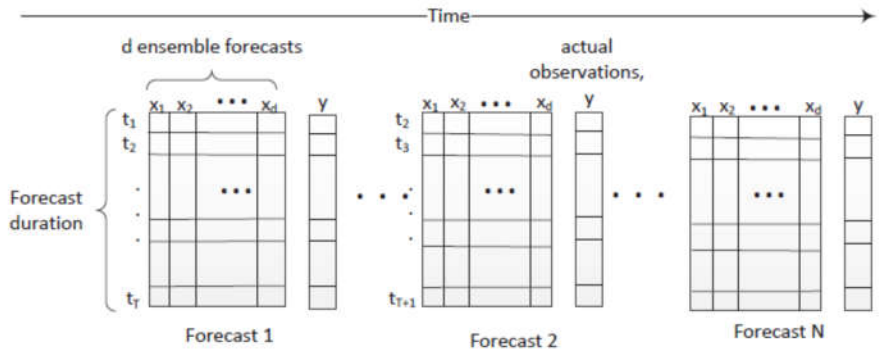
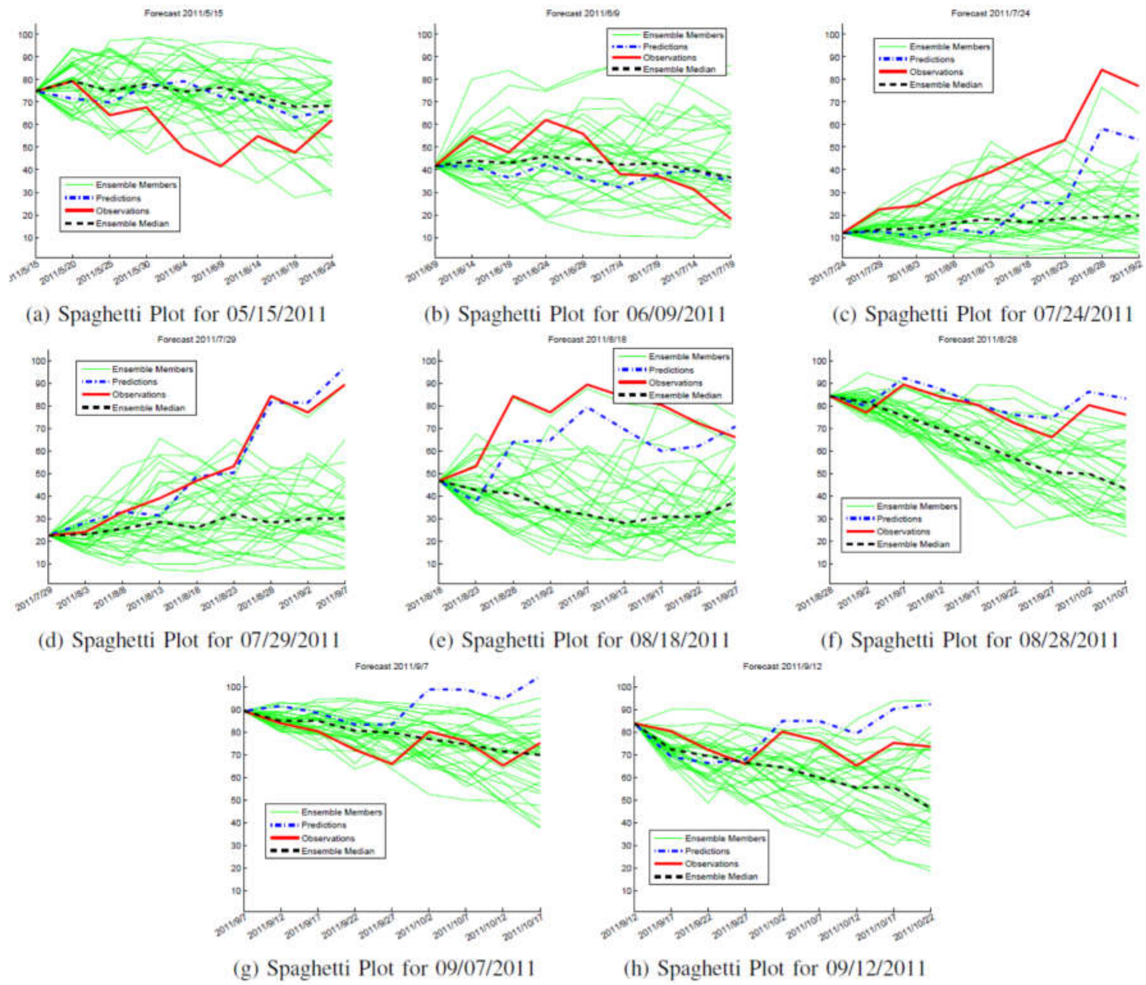


Figure 8: Schematic illustration of ensemble prediction as a multitask learning problem.

During the project, we have developed a framework called ORION, which stands for **O**nline **R**egularized mult**I**-task regression, to estimate the optimal weighted combination of ensemble members. The framework employs graph regularization constraints to ensure smoothness in the model parameters while taking into account the inherent relationships between forecasts in each run. The framework is applicable to different types of loss functions, including quantile loss, which is particularly appealing for forecasting extreme events in a time series.

We evaluated the performance of the framework on soil moisture data produced by the seasonal hydrological prediction system using the VIC model, for 12 major river basins in North America. The ensemble members of the data sets are sequences of basin-averaged soil moisture forecasts generated every 5 days from April 2011 to September 2011. We trained our model on the first 23 forecasts and evaluated it on the last 10 forecasts. We compared the performance of our method against 4 competing baseline methods and showed that our proposed method (ORION) outperforms the baseline methods (including ensemble median, EM) for the majority of the data sets in terms of their mean absolute error.

To illustrate the effectiveness of the method, Figure 9 compares the forecasts generated by ORION against the ensemble median for one of the selected data sets. Note that this is an online learning algorithm, so initially, the forecasts generated by ORION are similar to the ensemble median. However, as more data becomes available, the predictions by ORION get closer to the observed time series as the past information is used to learn the weights gradually.



(a) Spaghetti Plot for 05/15/2011 (b) Spaghetti Plot for 06/09/2011 (c) Spaghetti Plot for 07/24/2011

(d) Spaghetti Plot for 07/29/2011 (e) Spaghetti Plot for 08/18/2011 (f) Spaghetti Plot for 08/28/2011

(g) Spaghetti Plot for 09/07/2011 (h) Spaghetti Plot for 09/12/2011

Figure 9: Illustration of ORION algorithm on developing weights for different ensemble members of basin-average soil moisture values.

Later during the project, co-PI Tan and his student continued their investigation into the development of novel computational approaches for combining and post-processing multiple forecasts to achieve better forecast quality. Although the results obtained using ORION are promising, there are two limitations to the framework. First, it was not suitable for extreme value predictions (e.g., forecasting of extreme drought or wet conditions) as it was designed to optimize a quadratic loss function. Predicting the frequency and timing of extreme events are important for applications such as hydrological forecasting due to their impacts on both human and natural systems. Second, the framework trains a forecast model for each basin location independently. As a consequence, it may not be able to utilize information from other nearby basins to improve the quality of its seasonal forecasts.

To overcome the first limitation, we extended the ORION framework to deal with extreme value predictions by incorporating a quantile loss function. The proposed extension, known as ORION-QR was applied to the soil moisture time series data for 5 major river basins in North America. Extreme events were initially identified as values exceeding 1.64 standard deviations away from the mean (i.e., beyond 90% confidence interval of the mean). We divided each time series into a separate training and validation periods. Our goal was to accurately predict the timing of occurrence for such extreme events during the validation period. Using an evaluation measure known as F1 (which combines sensitivity and specificity of the detection into a single metric), the experimental results shown in Table I suggest that ORION-QR outperforms four other baseline methods (including the original ORION framework) in 3 out of the 5 basins evaluated. A manuscript reporting the results of this analysis is currently in preparation.

As the ORION framework applies the post-processing approach to the ensemble forecasts at each basin independently, it will not be able to exploit any information from nearby basins to improve its aggregated forecast even though there may be spatial autocorrelations present in the data. Towards this end, we begun some preliminary work to develop multi-task learning techniques for modeling time series data at multiple locations simultaneously. Specifically, our spatial MTL method considers the forecast at each location (basin) as a separate task and learns the forecast model for all locations simultaneously, taking into account the inherent spatial autocorrelations in the data.

Our results have shown that the performance of ORION on ensemble soil moisture forecasting is superior over the existing state-of-the-art baselines, as well as the ensemble mean or median, which is commonly used for ensemble forecasting. We have extended the conference version of the paper and submitted it to a journal, IEEE Transactions on Knowledge and Data Engineering (TKDE), which is one of the premier journals in Data Mining. The extension includes theoretical analysis regarding convergence of the proposed algorithm as well as performing sensitivity analysis on the results of the experiments. The initial prototype of the ORION algorithm was implemented in Matlab. Our team has also implemented the algorithm in Python so that it can be easily integrated into the seasonal hydrology forecasting system.

We also examined whether the framework will generalize well to other ensemble forecasting problems. To evaluate this, the ORION framework was applied to the ensemble prediction of Niño3.4 index (sea surface temperature over the Niño 3.4 region in the equatorial Pacific) using data from NCDC daily SST analysis (from Dec. 1981 to Nov. 2014). The duration of each forecast is 9 months, using 20 ensemble members. The first 95 months of SST data were reserved for training while the remaining 300 months were used for validation. The mean absolute error (MAE) for the overall forecast using ORION is 0.408 whereas the MAE using Ensemble Median (EM) is 0.562. The MAE results for each start month and lead time are shown in Tables 1 and 2, for ORION and EM, respectively. The results showed that ORION significantly outperformed EM for most of the start months, especially for the longer lead time, except for November and December.

GSpartan algorithm

Building upon the ORION framework, we also developed a multi-task learning framework for spatio-temporal data known as GSpartan, which stands for GeoSPAtio-tempoRal mulTi-tAsk learNing. GSpartan learns a set of base models for all the locations and expresses the local model at each location as a linear combination of the base models. GSpartan differs from ORION in that it considers each prediction site as a separate learning task, which is related to the learning tasks at other prediction sites. Multi-task learning (MTL) is a well-known machine learning paradigm for solving multiple, related prediction tasks simultaneously by considering their shared information. The rationale for using MTL is that the information propagated between related tasks may enhance the overall accuracy if the models are trained jointly instead of independently. The algorithm is useful for spatial downscaling variables across multiple locations (regions) simultaneously considering their spatial correlations in seasonal climate prediction. GSpartan learns a set of base models for all the locations and expresses the local model at each location as a linear combination of the base models.

GSpartan was applied to downscale monthly precipitation data for 37 randomly selected weather stations from January, 1961 to December, 2000. We used 26 coarse-scale variables from NCEP reanalysis data as predictor variables for our statistical downscaling algorithm. Several state-of-the-art baselines were compared against GSpartan, including LASSO regression, MRMTL (mean regularized MTL), and SLMTL (sparse low-rank MTL). Figure 10 compares the prediction error of GSpartan against the baselines for all 37 stations. The results show that GSpartan outperforms other baselines for most of the stations. In particular, Table 1 summarizes the number of stations in which the prediction of each method is better than one of its competing method. The results suggest GSpartan outperformed other baseline methods in at least 32 (86.5%) out of 37 stations.

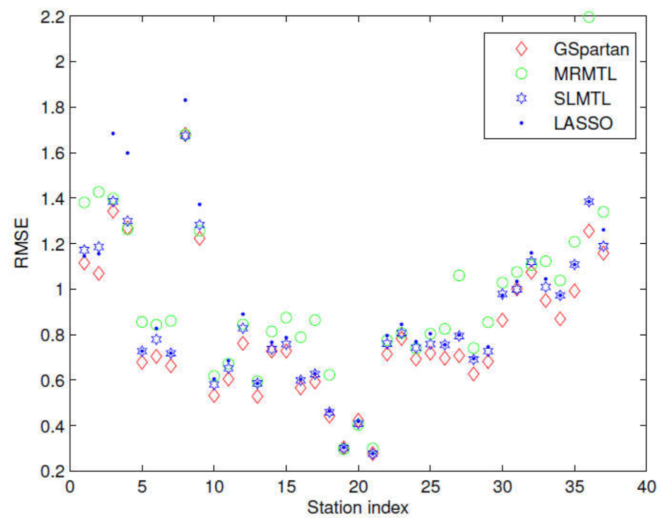


Figure 10: Comparison of GSpartan against three baseline methods

Table 1: Win-loss table comparing performance of various methods when applied to the data set with limited training examples (1 year of training data and 39 years of test data)

	GSpartan	GSpartan-norm	GSpartan-NTR	MRMTL	SLMTL	LASSO
GSpartan	-	32	37	33	32	36
GSpartan-norm	5	-	20	31	27	34
GSpartan-NTR	0	17	-	31	27	34
MRMTL	4	6	6	-	6	13
SLMTI	5	10	10	31	-	34
LASSO	1	3	3	24	7	-

WISDOM algorithm
Previous methods are useful but they haven't fully taken advantage of the spatial and temporal autocorrelations within the observations. A multi-task learning framework called WISDOM, which stands for Weighted Incremental Spatiotemporal multi-task learning via tensor decomposition was developed. The algorithm represents the spatiotemporal data as a third-order tensor, where the dimensions (modes) of the tensor represent the temporal, spatial, and predictor variables of the data. By performing tensor decomposition, the latent factors that characterize the variability of the data along each of the three dimensions can be identified. For climate data, known temporal patterns such as El Niño can be directly integrated as a constraint on one of the temporal latent factors of the spatiotemporal tensor. Sparsity-inducing norms can also be added as additional constraints to avoid model overfitting and enhance model interpretability. Figure 11 illustrates the concept of the framework.

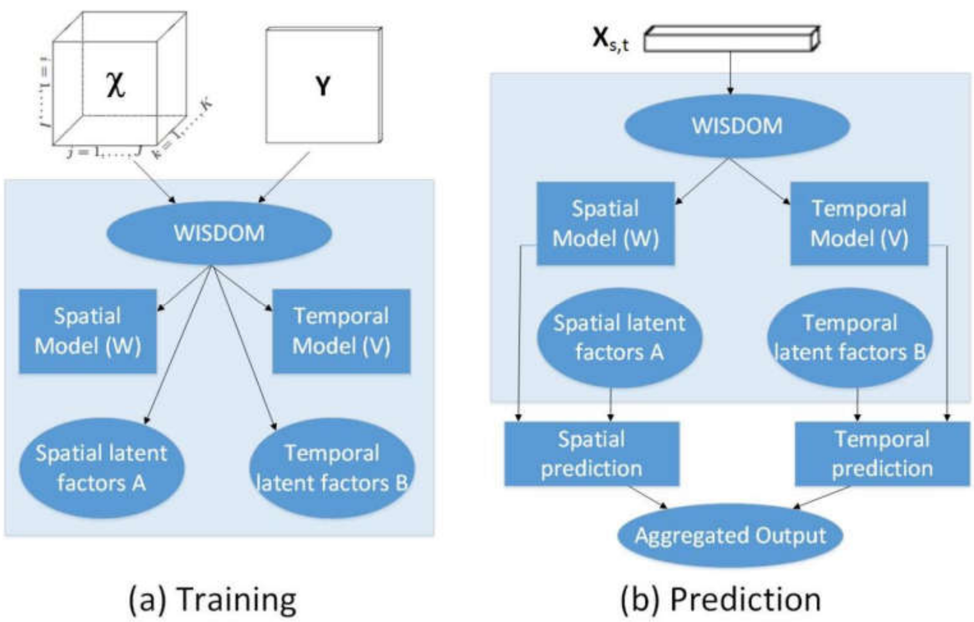


Figure 11: Overview and illustration of the WISDOM framework for statistically downscaling climate variables over multiple locations simultaneously by considering their spatial and temporal correlation in the multi-task learning algorithm

We applied WISDOM to a global-scale climate data set and compared its performance against several baseline algorithms. The climate data was obtained from two sources. First, we downloaded the monthly climate observation data from the Global Surface Summary of Day (GSOD) website. GSOD provides climate data from more than 30,000 monitoring sites worldwide, spanning a time period from 1942 to present. These monthly values of total precipitation (prcp), maximum (tmax), minimum (tmin), and average (tmean) temperature are used to define the target/response variable for our prediction tasks. We created 4 data sets, one for each response variable, to evaluate the performance of WISDOM. The second source corresponds to a coarse-scale gridded climate data from NCEP reanalysis2. We use the data to define the predictor variables for our climate prediction task. Although there are hundreds of variables available in the NCEP reanalysis data, we selected 13 of them as our predictor variables that are more related to GSOD variables. We use the monthly data from January 1985 to November 2015 (for a total of 371 months) in our experiment.

We compared WISDOM against g two baseline algorithms: single task learning (STL) and accelerated online low rank tensor learning (ALTO). The results suggest that WISDOM outperform STL and ALTO in more than 75% of the locations for all 4 data sets evaluated. The percentage is even higher (> 90%) when compared against ALTO on the three temperature data sets. By outperforming STL, this suggests the importance of incorporating spatial autocorrelation into the learning framework.

This work was presented at the 2016 IEEE international conference on Big Data, and it received the best paper award.

7. Contribution of temperature and precipitation anomalies to the multiyear California drought

This was not the original task for this project, but the idea was developed during the project as the result of the involvement in the Drought Task Force (DTF) led by Martin Hoerling. One of the major topics during the second DTF was to understand the drought mechanisms while California was experiencing a major multiyear drought that is characterized by major precipitation reduction and higher than normal temperatures. So the DTF was set off to investigate the role of temperature anomalies to this drought event.

The PI and his students took the effort to investigate contribution of temperature and precipitation anomalies to the multiyear drought over California, which will help us better understand the drought mechanism and predictability. This work was inspired by the fact that this drought event was accompanied by abnormally higher temperatures as compared to earlier multi-year drought in the region. We use the VIC model to simulate the hydrology over the region and examine the sensitivity of agricultural and hydrological drought to temperature and precipitation anomalies in a series of modeling experiments. Here agricultural drought refers to the lack of soil moisture in the total soil column simulated by the VIC model that has three soil layers with the top layer the thinnest (10 cm) and varying depths of the other layers over different regions. For hydrological drought we focus on the reduction in snow water equivalent (SWE) and runoff.

The study comprises of a number of offline simulations. The CTRL simulation is driven by the observed temperature and precipitation and is our best estimate of reality. The CTRL simulation actually starts from 1979 and continues to the present. A model state is saved at the end of September 2011 to provide the initial condition for all other experiments that start from the beginning of the 2012 water year. In the CLIMP simulation, the monthly precipitation anomaly is removed for all months in the 4-year period of October 2011 to September 2015. To remove the anomaly, we multiply the observed daily precipitation within a month by a common scaling factor so that the monthly total precipitation after adjustment matches the climatological mean precipitation for that month. The CLIMT simulation is done in a similar way except that daily temperature is adjusted by subtracting the observed monthly anomaly. The CLIMPT simulation has both temperature and precipitation anomalies removed. In the RANDP (random precipitation) ensemble experiment, we preserve the observed temperature and use the precipitation time series from four random water years between 1980 and 2015 to replace the observed precipitation. Similarly, the RANDT (random temperature) experiment preserves the observed precipitation and replaces the temperature with ones from random years. These simulations provide a useful set of data to quantify the role of temperature and precipitation anomalies to the drought development in two different ways.

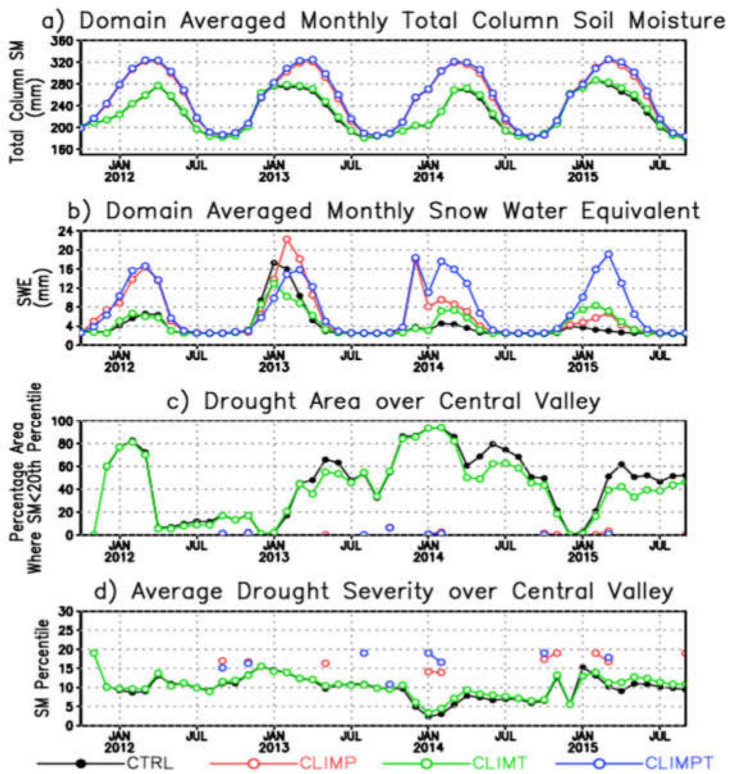


Figure 12: Land surface modeling experiments show how temperature and precipitation anomalies have contributed to the 2012-2015 California drought measured by both soil moisture and snow pack. Time series of (a) basin-averaged total column soil moisture, (b) snow water equivalent (SWE), (c) percentage area, and (d) average severity of agricultural drought over Central Valley from CTRL, CLIMP, CLIMT, and CLIMPT simulations. The drought area is represented by the percentage area of the Central Valley where the monthly total column soil moisture is below the 20th percentile of its climatological distribution, and the average severity is represented by the averaged soil moisture percentile of the grids in drought. A lower soil moisture percentile indicates a more severe drought. Missing values in the CLIMP and CLIMPT time series correspond to months when no grid cell is in drought.

The research was first presented at AGU fall meeting, and the talk generated healthy discussion about our original approach. The experiments that were performed earlier were subsequently revised to include both single and ensemble sensitivity runs as outlined above. This new experiment design is useful for quantifying the temperature contribution deterministically and probabilistically. Figure 12 is illustrates the impact of higher temperature during the period on both soil moisture and snow pack. Our investigation shows that low precipitation was the main driver of the drought conditions, as measured by soil moisture deficits. Temperature helped to exacerbate the drought - especially by reducing the snow water equivalent accumulated in the Sierra snowpack, effectively changing the timing of water availability.

The final work was presented in one of the Drought Task Forcing telecom, and was published in Geophysical Research Letters. This work also contributes to the discussion and the synthesis of DTF research on temperature and drought, and is included in the final information sheet from DTF2 report.

8. Multiyear drought hindcast over the US

We requested a one-year no-cost extension to complete the comprehensive evaluation of the drought forecast approach by conducting a 28-year (1982-2010) monthly hindcast over the NLDAS domain. This was the major task for the last year.

The hydrological hindcast dataset includes the following. Temperature and precipitation forecast from the CFSv2 model were processed by the multi-scale Bayesian merging approach in our Drought Monitoring and Prediction System (DMAPS) to produce a 120-day outlook for each day during the 28 years. Then the temperature and precipitation outlooks at the beginning of each month are used to drive the VIC hydrological model for the next four months to produce a seasonal ensemble drought forecast for the following four months. The ensemble forecast comprises of 33 members, and the predicted model states in soil moisture and other state variables are then verified against the long-term historical offline simulation by VIC. For comparison, a 33-member ensemble using the ESP approach is also produced each month. Each forecast is about 80 GB in size, so the total data produced in the 28-year hindcast is about 27TB. The post-processed results from each month's forecast are archived for comprehensive evaluation.

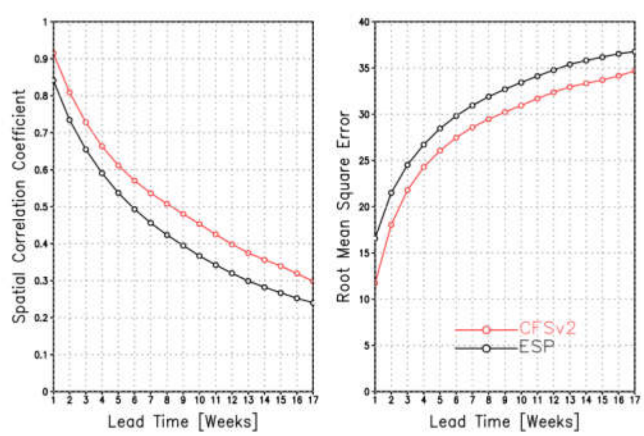


Figure 13: Spatial correlation and RMSE of SMI (soil moisture index, expressed as the percentile of total column soil moisture) forecast verified against analysis.

Meanwhile, we have also produced a real-time weekly forecast dataset in similar setup from April 2011 to present. The real-time results are available online on our project website. The verification of the hindcast and the real-time forecast are completed, and we are working on two manuscripts that are soon to be submitted. Figure 13 shows one of the results on soil moisture forecast verification from the real-time forecast. This evaluation is based on seasonal drought forecasts that were issued between April 6, 2011 and January 21, 2016 on every Wednesday. For a given forecast, we can first quantify how similar the predicted drought condition over CONUS is to the analysis from the offline simulation. This is achieved by calculating the spatial correlation and the root-mean-square error (RMSE) between two spatial fields at a given time. Figure 13 shows the spatial correlation between the predicted drought condition (ensemble mean on a given day) and the drought analysis (offline simulation driven by observed NLDAS2 forcing). It is clear that CFSv2-based forecast shows higher spatial correlation and smaller RMSE than the ESP-based forecast for all lead times.

JOURNAL PUBLICATIONS AND CONFERENCE CONTRIBUTIONS

This project has led to or contributed to the following peer-reviewed publications and conference presentations.

Peer-reviewd publications

Xu, J., P. N. Tan, J. Zhou, and L. Luo (2017), Online Multi-task Learning Framework for Ensemble Forecasting. IEEE Transactions on Knowledge and Data Engineering, 29(6) 1-13. doi:10.1109/TKDE.2017.2662006.

Luo, L., D. Apps, S. Arcand, H. Xu, M. Pan, and M. Hoerling (2017), Contribution of temperature and precipitation anomalies to the California drought during 2012–2015, Geophys. Res. Lett., 44, 1–9, doi:10.1002/2016GL072027.

Xu, J., J. Zhou, P. N. Tan, X. Liu, and L. Luo (2016), WISDOM: Weighted Incremental Spatio-Temporal Multi-Task Learning via Tensor Decomposition, 2016 IEEE International Conference on Big Data.

Xu, J., P. N. Tan, L. Luo, and J. Zhou (2016), GSpartan: a Geospatio-Temporal Multi-task Learning Framework for Multi-location Prediction, Proceedings of the 2016 SAIM International Conference on Data Mining, 657-665.

Xia, Y., B. A. Cosgrove, K. E. Mitchell, C. D. Peters-Lidard, M. B. Ek, M. Brewer, D. Mocko, S. V. Kumar, H. Wei, J. Meng, and L. Luo (2016), Basin-scale assessment of the land surface water budget in the National Centers for Environmental Prediction operational and research NLDAS-2 systems, J. Geophys. Res. Atmos., 121, 2750–2779, doi:10.1002/2015JD023733.

J. Xu, P. N. Tan and L. Luo, "ORION: Online Regularized Multi-task Regression and Its Application to Ensemble Forecasting" 2014 IEEE International Conference on Data Mining, Shenzhen, 2014, pp. 1061-1066. doi: 10.1109/ICDM.2014.90

Tang, W., Z. Lin, C, Yang, and L. Luo (2014), Evaluation of a hydrological simulation over the Huaihe River Basin using a coupled land surface and hydrologic model system and its uncertainty analysis. Climatic and Environmental Research (in Chinese), 19 (4): 463−476. doi:10.3878/j.issn.1006-9585.2013.13066

Lang, Y., A. Ye, W. Gong, C. Miao, Z. Di, J. Xu, Y. Liu, L. Luo, and Q. Duan (2014), Evaluating skill of seasonal precipitation and temperature predictions of NCEP

CFSv2 forecasts over 17 hydroclimatic regions in China, J. of Hydrometeo., 15(4), 1546-1559. doi: 10.1175/JHM-D-13-0208.1
Tang, W., Z. Lin, and L. Luo (2013), Assessing the Seasonal Predictability of Summer Precipitation over the Huaihe River Basin with Multiple APCC Models. Atmos. Oceanic Sci. Lett., 6(4), 185-190, doi: 10.3878/j.issn.1674-2834.13.0025.
Luo, L., W. Tang, Z. Lin, and E. F. Wood (2013), Evaluation of summer temperature and precipitation predictions from NCEP CFSV2 retrospective forecast over China. Climate Dynamics, 41, 2213-2230, DOI 10.1007/s00382-013-1927-1
Yang Lang, Lifeng Luo, Qingyun Duan, Zhaohui Lin, and Aizhong Ye, Does CFSv2 help improving forecast of meteorological drought over mainland China? Atmospheric Science Letters, in revision.

Conference Presentations

Lifeng Luo, Wei Tang, Pang-Ning Tan, and Zhaohui Lin (2012), CFSv2 driven seasonal drought prediction over the United States (American Geophysical Union, San Francisco, CA, December 3-7, 2012)
Lifeng Luo and Pang-Ning Tan (2012), CFSv2 driven seasonal drought prediction over the United States (37th Climate Diagnostic and Prediction Workshop and US Drought Task Force meeting, Fort Collins, CO, October 22-26, 2012)
Lifeng Luo and Andy Wood: Seasonal streamflow prediction with VIC and Climate Forecast System version 2. AGU Chapman Conference on Seasonal to Interannual Hydroclimate Forecasts and Water Management, 28 July-31 July 2013, Portland, OR.
Lifeng Luo, Jianpeng Xu, and Pang-Ning Tan: Postprocessing of hydrological ensemble prediction with an online learning method. AGU Fall meeting, 8 December – 13 December, 2013, San Francisco, CA.
Yang Lang, Maura Casey, Lifeng Luo, Qingyun Duan (2014), Evaluation of Drought Monitoring using Three Global Reanalysis Products (Association of American Geographers annual meeting, Tampa, FL, April 8-11, 2014)
Yang Lang, Lifeng Luo, Maura Casey, Qingyun Duan (2014), Characterizing the skill of CFSv2-based seasonal drought prediction at multiple spatiotemporal scales over China (American Geophysical Union Fall Meeting, San Francisco, CA, December 15-19, 2014)
Maura Casey, Lifeng Luo, Yang Lang, Application of Multi-Model CMIP5 Analysis in Future Drought Adaptation Strategies (American Geophysical Union Fall Meeting, San Francisco, CA, December 15-19, 2014)
Jianpeng Xu, Pang-Ning Tan, Lifeng Luo, and Jiayu Zhou. GSpartan: a GeoSpatio-Temporal Multi-task Learning Framework for Multi-location Prediction. In Proceedings of the SIAM International Conference on Data Mining, 2016.
Lifeng Luo et al., Contribution of Temperature and Precipitation Anomalies to the 2012-2015 California Drought, in preparation. (Presented at AGU Fall meeting, San Francisco, CA)