Examining the Sensitivity of Monthly Temperature Forecast Models to Multiple Sources of Soil Moisture Data

GEOGRAPHY

Zack Leasor (leasor.4@osu.edu), Steven M. Quiring (Department of Geography, The Ohio State University)

Motivation

Previous research has demonstrated that soil moisture improves warm season temperature forecasts at lead times ranging from one month to one season. To account for the increasing availability of soil moisture data, this study examines the sensitivity of forecast models to varying data sources. Soil moisture data from direct in situ measurements and land surface models are considered in addition to the Standardized Precipitation Index (SPI) as a proxy estimate. These data are integrated into forecasts to examine how the data source impacts model performance.

Data

Data:

- Monthly in situ soil moisture data (2009-2018) retrieved from 8 networks
- Quality control used to identify stations continuously operating from 2009 to 2018
- All measurements standardized as VWC percentiles
- Converted to NLDAS layers by averaging any measurements that fall within the range of depths

Table 1. Number of stations providing a continuous period of soil moisture monitoring from 2009 to 2018 with at least one error-free measurement recorded at the shallowest monitoring depth.

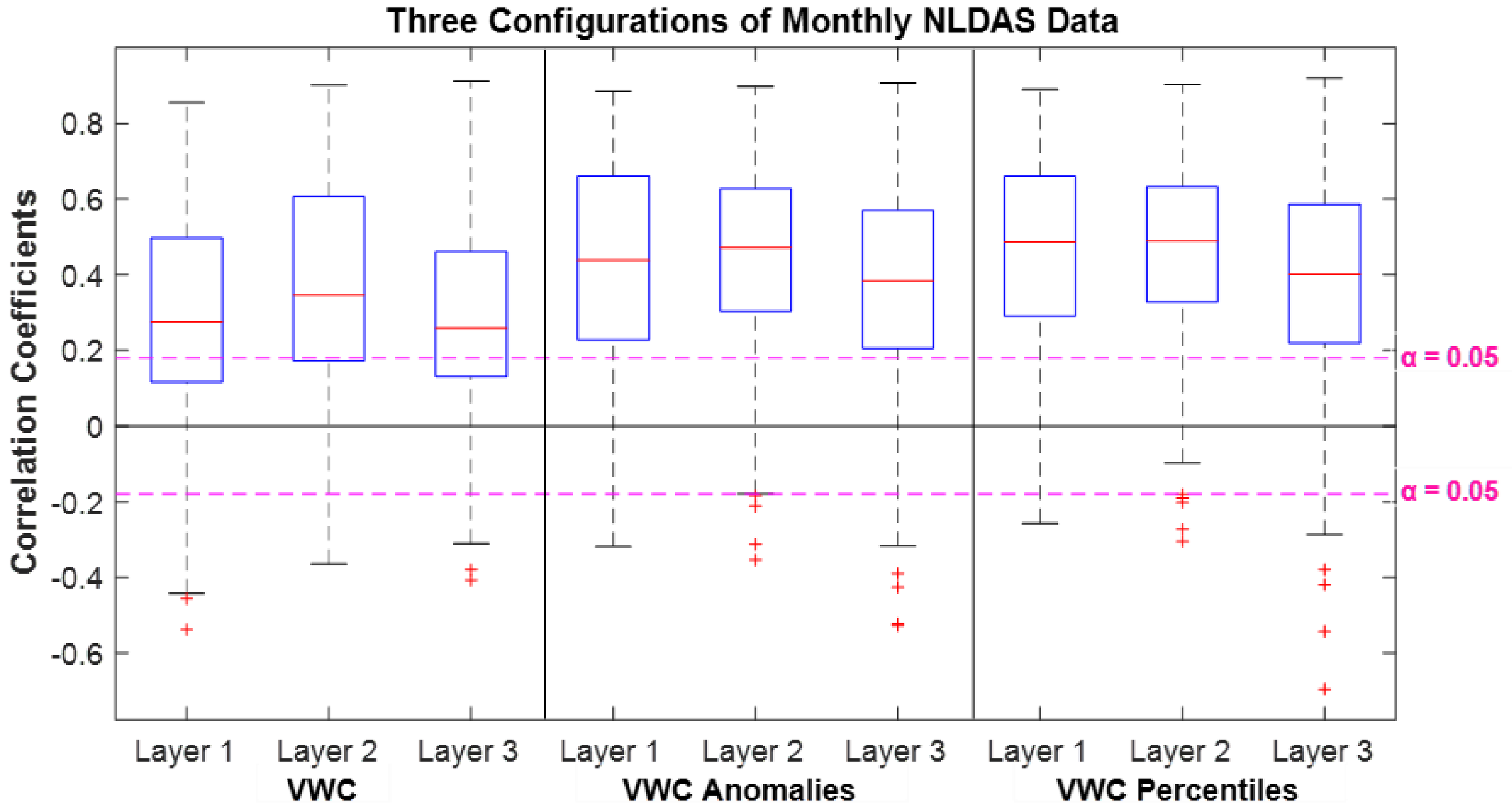
Network	Number of Stations (QC)
DEOS	9
MHAWD	3
NCEconet	34
OKM	100
SCAN	96
SNOTEL	222
UGA	71
WTM	42
All	**577**

* Data provided by http://nationalsoilmoisture.com and the UGA weather network http://www.georgiaweather.net/

- NLDASv2 monthly soil moisture data simulated from the Noah model
  - Soil layer 1 (0-10 cm), Soil layer 2 (10-40cm), Soil layer 3 (40-100cm)
- Monthly temperature and precipitation data from PRISM (http://prism.oregonstate.edu)
  - Precipitation data used to calculate the Standardized Precipitation Index (SPI)

Results: Model Output vs. In Situ Soil Moisture

Correlations Between Monthly In Situ Soil Moisture Percentiles and Three Configurations of Monthly NLDAS Data



Correlations between NLDAS and In Situ Data which are not significant or <0

0 – 10 cm	10 – 40 cm	40 – 100 cm
NCEconet (10) SCAN (2) SNOTEL (78)	NCEconet (8) SCAN (9) SNOTEL (38) UGA (5)	SCAN (38) SNOTEL (59) WTM (2)

Figure 1 (top). Boxplots displaying the distribution of correlation coefficients between monthly in situ soil moisture percentiles and monthly NLDAS VWC, VWC anomalies, and VWC percentiles. . Each boxplot represents a distribution of correlation coefficients between each station and its corresponding NLDAS grid cell.
Table 2 (bottom). Networks and the associated number of stations which display negative or insignificant correlations between in situ ad NLDAS monthly soil moisture percentiles.

Results: Temperature vs. Soil Moisture

Forecast Models	3-Months Antecedent (*n-3*)	2-Months Antecedent (*n-2*)	1-Month Antecedent (*n-1*)	Current Month (*n*)
	July	*August*	*September*	*October*
0-month lead				Prediction
1-month lead				Prediction
2-month lead				Prediction

Warm Season Correlations with Temperature

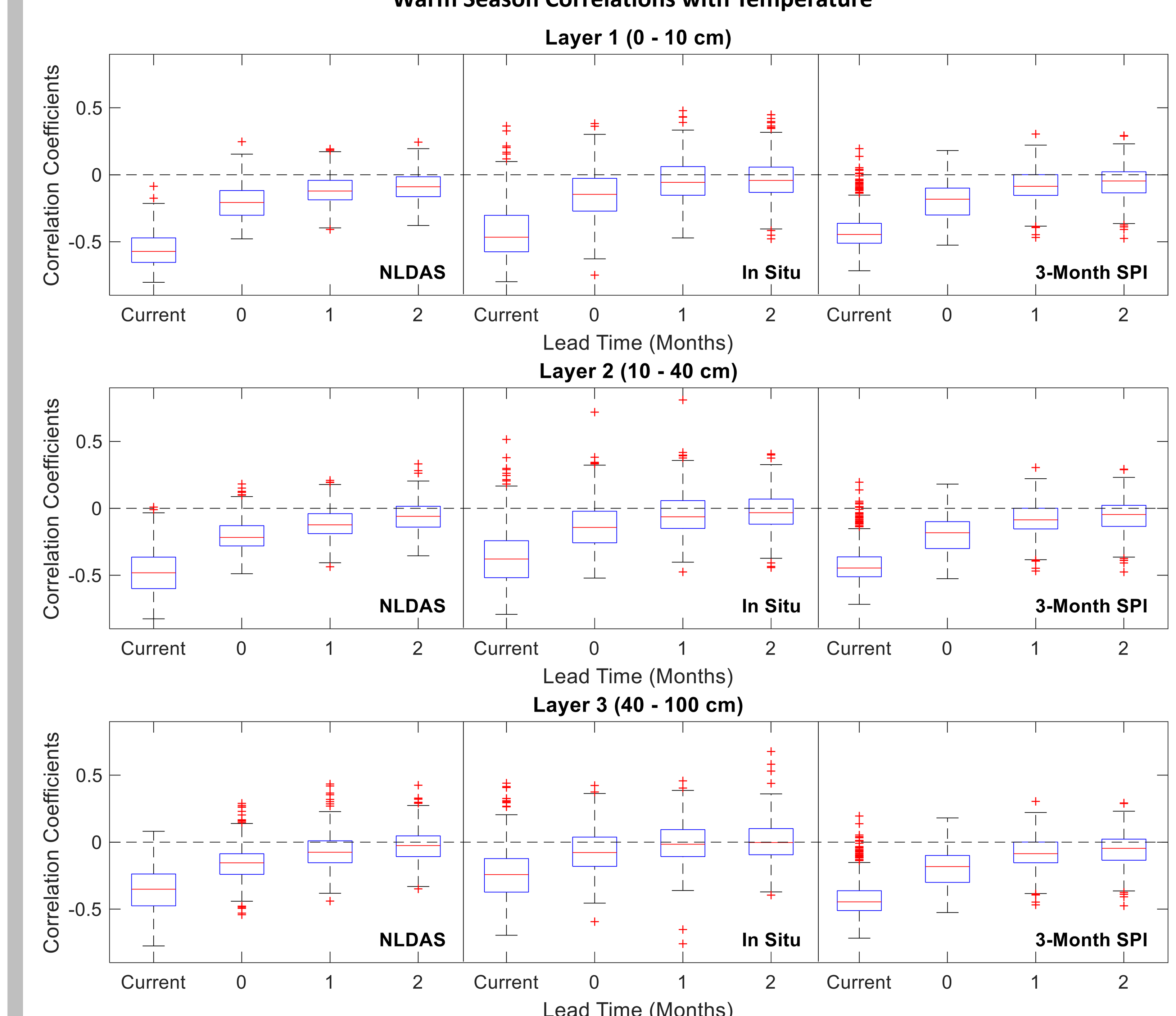


Table 3 (top). Schematic describing forecast lead times and an example of a prediction made using the designated forecast construction.
Figure 2 (bottom). Boxplots displaying distributions of warm season (May – October) correlations between monthly maximum temperatures and various configurations of antecedent soil moisture data.

Methods: Quantile Regression Forecast Models

Quantile Regression

- Used to define unique model parameters for each location based on five quantiles of the temperature data
- Linear slope and intercept terms vary based on the relationship between temperature and antecedent soil moisture

In-Sample Model Validation

- Predictions binned into categorical terciles labeled as normal, above normal, and below normal

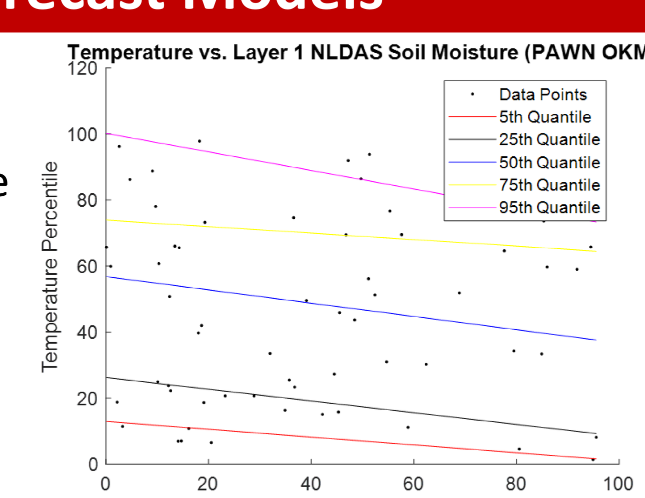


Figure 3. Example displaying the quantile regression model fits between temperatures and NLDAS soil moisture at the PAWN OKM site.

- A consensus (mode) prediction is developed based on the predictions made for each quantile of the temperature distribution
  - 'Normal' temperature forecast when a consensus is not reached
- Heidke Skill Score (HSS): used for forecast evaluation
  - A correct forecast (*C*) occurs when the predicted and observed terciles are equal. *T* is the total number of forecasts and *E* represents the number of correct random forecasts *(T/3)*

$$HSS = \frac{(C-E)}{(T-E)}$$

Results: Model Validation

0-Month Lead Warm Season Temperature Forecasts (2009-2018)

NLDAS vs. In Situ	Significant Difference?				
Model Parameters	5th Quantile	25th Quantile	50th Quantile	75th Quantile	95th Quantile
Layer 1 Slope	✓				
Layer 1 Intercept	✓				✓
Layer 2 Slope	✓	✓			
Layer 2 Intercept	✓	✓	✓		✓
Layer 3 Slope	✓	✓			
Layer 3 Intercept	✓	✓	✓	✓	✓

Table 4. Table displaying parameters for models used to predict warm season temperature at a 0-month lead from 2009-2018. National distributions of model parameters were tested using a two-sample t-test. Parameters shaded green indicate that the null hypothesis of the parameters from independent distributions with equal means and variance at a 95% confidence level.

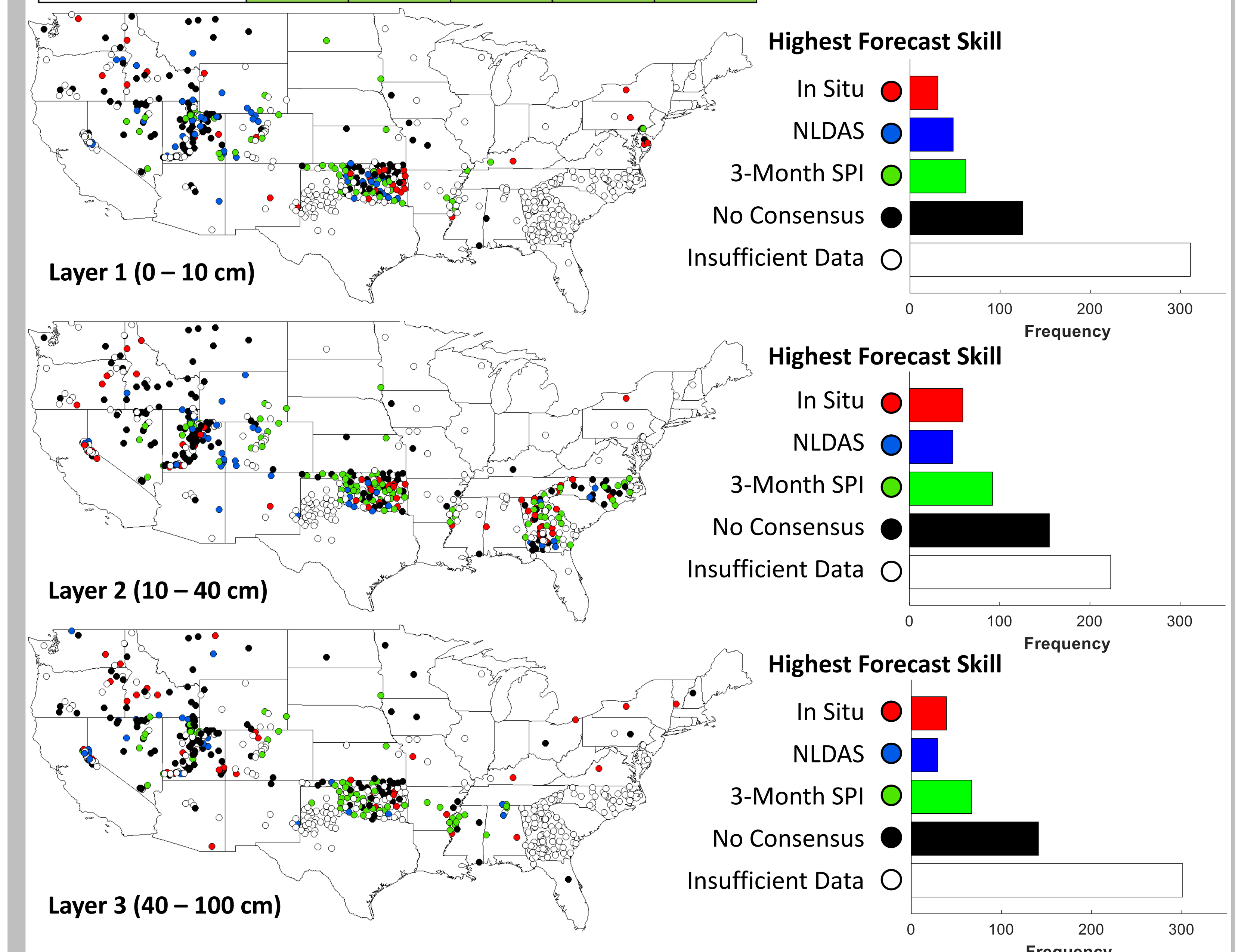


Figure 5. Maps and corresponding histograms displaying the data source with the highest forecast skill. The rankings are based on the maximum HSS value for 0-month lead warm season temperature forecasts from 2009-2018. No consensus indicates a tie among data source and locations with missing in situ measurements are labeled as insufficient.

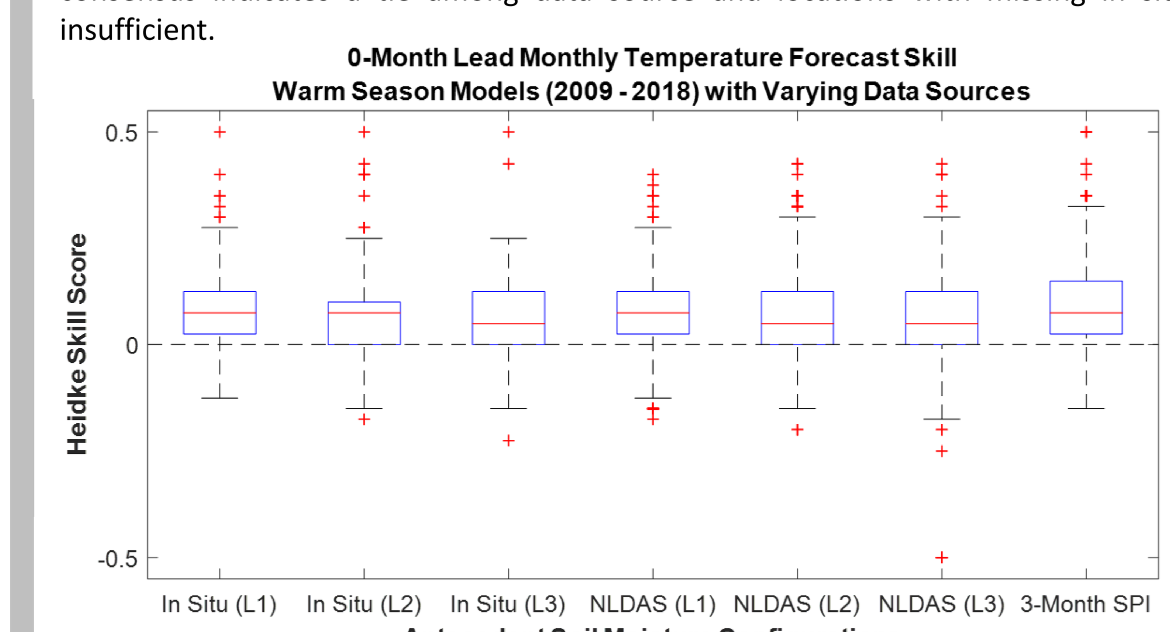


Figure 6. Boxplots displaying the distribution of HSS values for predictions based on 0-month lead warm season monthly temperature forecasts. Each model uses a unique set of predictors based on the source of soil moisture data and the depths used. HSS values are used to quantify hindcast skill for all monthly predictions ranging from 2009-2014.

Conclusions

- Soil moisture shows the strongest relationship with temperature at a 0-month forecast lead time.
- There are significant differences between in situ and NLDAS data and models are consequently sensitive to the type of soil moisture used in forecasts.
- The 3-month SPI produces more skillful temperature forecasts than in situ or NLDAS soil moisture.
- Differences in model performance between in situ and NLDAS soil moisture vary by depth, but in most cases there are no significant differences.