Project Title: “Towards week-2 to week-4 excessive heat outlooks. Evaluation of the forecast skill of the North American Multi-Model Ensemble System”
PI: Augustin Vintzileos
Co-PI: Jon Gottschalck
Program Officer: Heather Archambault
Project Final Report
Reporting Period: 08/01/2015 - 07/31/2016

1. Summary

This project investigates the value added by the NMME in a Subseasonal Excessive Heat Outlook System (SEHOS). The research conducted here is part of a long term strategy developed by the PI for building and improving resilience to heat induced health hazards. This strategy (Figure 1) consists in three levels. In the first level we define excessive heat events based on bio-meteorological considerations and the constraints of probabilistic subseasonal forecasting of meteorological fields. These definitions allow to create monitoring and baseline forecasting systems. Monitoring systems are then used for research on predictability sources and for developing the forecast verification system. The verification system is used for assessing the quality of forecasts allowing for evaluation of new prediction systems that are based on the research on predictability sources. Finally, improvements of the SEHOS come from disseminating forecast and using the feedback from the users.

This final report discuss the development of monitoring/verification systems that allowed to test multi-model approaches. It also presents initial research on predictability sources of excessive heat events using methodology developed by the project. This research focused NMME model comparison to the forecast skill of geopotential height anomalies at 500 hPa. The central finding of this work is that multi-model approaches will allow the improvement of the operational SEHOS. A major benefit from this research is that the developed verification system facilitated the execution of experimental realtime forecasts during summer of 2016 which were available to CPC forecasters.

2. Definition of heatwaves

We define a heat event as the succession of at least two heat days. A heat day is defined as a day with maximum heat index exceeding a given percentile of the climatological distribution of maximum heat index for the geographical location and time of the year. The thresholds we use are: 90%, 95% and 98%, corresponding to heat events of increasing severity. A heat week for a given grid point is a weekly period containing at least one heat event. A heat week is characterized by the start day and the duration of the heat event.

3. Baseline forecast system

The baseline forecast system uses the GEFS and targets Week-2. For this project we are also using as baseline system the ECMWF model for Week-2 and Week-3. The forecast consists in computing whether a given week (Week-2 or Week-3) is a heat week for each of the ensemble forecast members. For heat weeks we additionally detect the starting day and compute the duration of the heat event. We then calculate statistics from the ensemble forecast: (a) probability of occurrence of a heat event, (b) mean start day and (c) mean duration of the event. The methodology we use for bias correcting the forecast heat index is an implicit quantile mapping. According to this method the percentile of the forecast heat index at a given forecast lead is computed by comparing it with the distribution of the reforecast heat index at the same forecast lead time at a given grid point. These percentiles are then mapped to the distribution of the observed values of the heat index.

4. Monitoring systems

We developed three monitoring systems. The first two are based on meteorological data from (a) the NCAR/NCEP Reanalysis and (b) the day-1 forecast from the GEFS Reanalysis. The third system is based on excessive heat warnings, advisories and watches encoded by the Valid Time Event Code (VTEC) by the WFOs. The uniqueness of the third system is that it adds the element of human perception of heat in the system (WFO forecasters). Figure 2 shows the utility of monitoring heat events based on the definition of heat we introduced; it depicts the heat wave which affected the broader Chicago area in July 1995 and which resulted in abnormal mortality in excess of 700. Figure 3a shows results from VTEC for the week ending 24 July 2016 during which a major heatwave developed in the mid part of the CONUS. Figure 3b shows the same heatwave as captured by the NCAR/NCEP monitoring system. Comparison of Figures 3a and 3b shows that the NCAR/NCEP based monitoring system quantifies heatwaves with precision; the heat events captured over the Rockies by the NCAR/NCEP system were not a threat to human health and therefore there were not emitted as such by the relevant WFOs.

5. Predictability sources

The excessive heat monitoring systems we developed were used as a tool for investigating the subseasonal predictability of heat waves based on the following methodology. We first identify a major heatwave (in this case the July 1995 heatwave). Then, we detect heatwaves from 1948 to 2015 over the CONUS based on the NCAR/NCEP Reanalysis. Then we calculate the pattern correlation between the major heatwave and all heatwaves from 1948-2015. We finally choose similar heat events by setting a correlation threshold. For example, Figure 4 shows the composite heat event when the correlation threshold is set to 0.55 for heat events at 90%. This results to 42 similar events from 1948 to 2015. Finally we composite weekly mean geopotential anomalies at 500 hPa for the week during the heat events, one week before and 3 weeks before. Results are depicted in Figure 5. The resulting modes are very robust and can be traced for at least three weeks before a heat event. These modes of variability that may embed extreme heat events set a minimal forecast target that models should be able to reach for predicting excessive heat events.

6. Utility of multi-model approaches

Before evaluating NMME models for excessive heat event forecast skill we focused on demonstrating that multi-model approaches are useful for predicting excessive heat events. For this we used as basis the ECMWF monthly forecasting system which we compared with the GEFS for forecast Week-2 and with the CFS for forecast Week-3. Although the GEFS is not a member of the NMME ensemble and is uncoupled, it shares the same atmospheric model with the CFS but with a higher resolution. The verification methodology used in this work is the Receiver Operating Characteristics (ROC) and the corresponding Area Under Curve (AUC). Briefly, ROC consists of attributing a forecast heat event when a threshold P of the forecast probability of occurrence is exceeded and comparing the probability of detection versus the probability of false detection for different values of P. The area under the resulting ROC curve (AUC) is a measure of forecast skill of the underlining probabilistic forecast system. AUC values close to 0.5 suggest a forecast system with no skill as the probability of false detection equals the probability of detection. AUC values equal to 1 indicate a perfect forecasting system. Figure 6 depicts the AUC for Week-2 and Week-3 for three models. Results from the ECMWF, GEFS and CFS models are shown respectively on rows (a), (b) and (c). Columns one and two refer respectively to Week-2 and Week-3 forecast skill. The cell corresponding to the GEFS at Week-3 is empty as currently such forecasts are not available. Row four shows AUC when we use the multi-model ensemble approach. Cell d.1 shows that when the GEFS is combined to the ECMWF model the forecast skill becomes generally higher that for the ECMWF model alone. At this time we have not combined the CFS model in the Week-2 multi-model system as we were not expecting the high forecast skill shown in cell c.1 (the CFS was not designed for Week-2 forecasts). However, a comparison of cells b.1 and c.1 suggests that including the CFS in Week-2 forecasts will likely enhance the forecast skill of the SEHOS. We must underline that in reforecast mode the GEFS is initialized by the CFS-Reanalysis and not by its own analysis system which hampers reforecast skill. On the other hand, when evaluating the CFS in reforecast mode we need to combine forecasts for the day and the day before (lagged ensemble generation technique) to generate an ensemble of 8 members. This is also not ideal for Week-2 forecasts. It follows that it is important to conduct diagnostic analyses of real time Week-2 forecasts. Finally, column 2 of Figure 3 shows that CFS forecasts are improving the skill of the ECMWF model in the mid-Atlantic and the northeast corridor.

7. Adding more NMME models

We established the importance of multi-model ensemble approaches for excessive heat outlook systems and detected persistent large scale modes of variability with the potential of embedding excessive heat events. Next, we investigate pattern correlation of geopotential at 500 hPa as a function of lead time. The NMME models that we are considering for this investigation are the CanCM3, CanCM4, CCSM4 and CESM1. The phase-2 NMME database is not containing daily geopotential reforecast data for the CFS. Such data could be found from the 45-day reforecasts with the CFS but only cover the period 1999-2010. The CanCMs are operational models that are initialized by their own analysis. The CCSM4 model is initialized by interpolation of the CFS-Reanalysis and the atmospheric component of the CESM1 model is not initialized. We first calculate the mean drift of these models during the first 4 weeks from initialization by averaging all reforecasts initialized from June, July and August. We then performed an EOF analysis on the resulting fields. The first principal component for all models represents a quasi-linear drift (not shown) with percentage of variance explained equal to 95% for the initialized models and 85% for the CESM1. Figure 7 shows the first EOF for each of the models. As expected the drift of the initialized models is more pronounced. The CanCMs and CCSM4 show very similar drift patterns despite their different formulations. It is interesting to note that over the eastern and central CONUS the geopotential height is linearly increasing and that the northern hemisphere height drift pattern projects on the excessive heat related height pattern. This projection could affect the forecast of heat events through non-linearities in the evolution of the atmosphere. Figure 8 shows the pattern correlation of geopotential anomalies as a function of lead week time. The MME approach offers a very small improvement of skill for forecast Week-3. However this could be the result of the initialization strategy of the CCSM4. Further, as discussed in this paragraph, the mean drift pattern projects to heat modes and therefore linear bias correction is not enough.

8. Figures

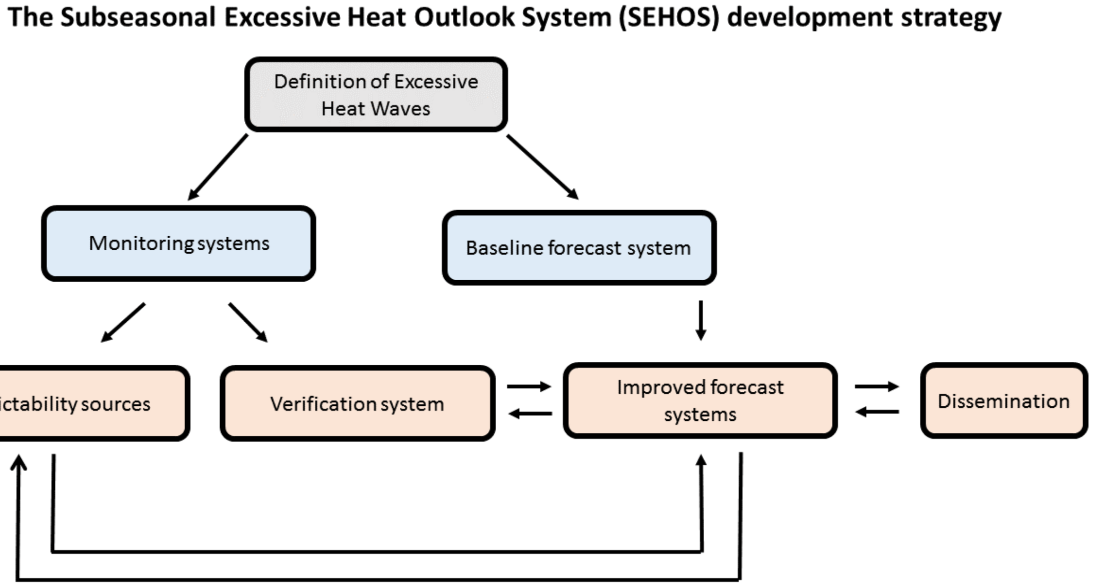


Figure 1: The long term strategy for the development and improvement of a subseasonal excessive heat outlook system.

Description of the July 1995 Heat Event

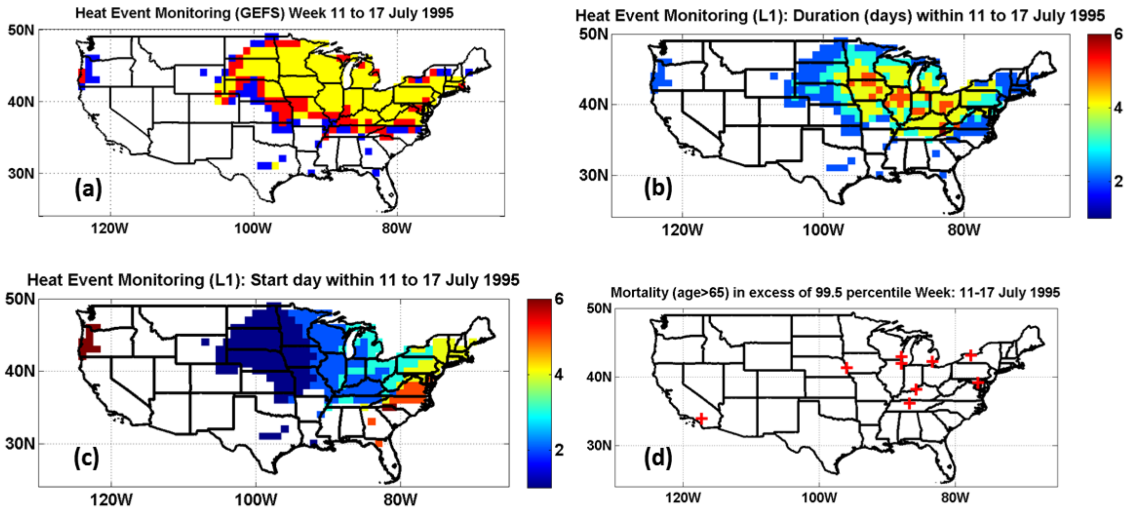
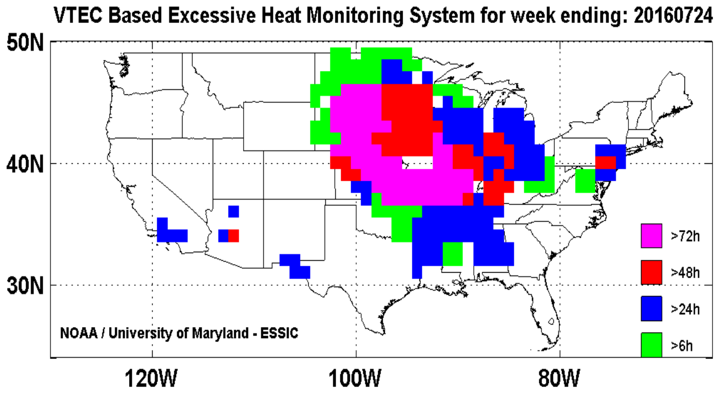
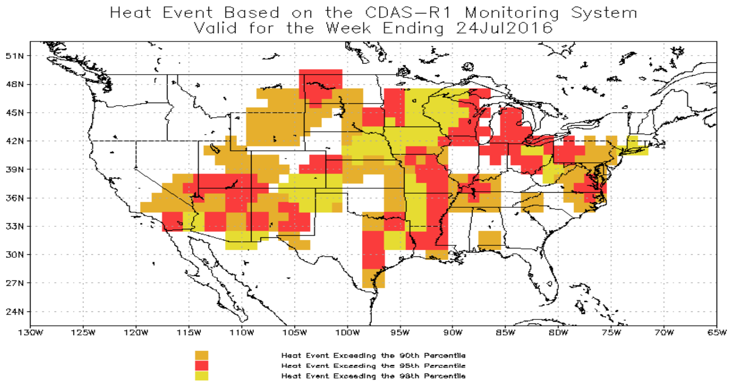
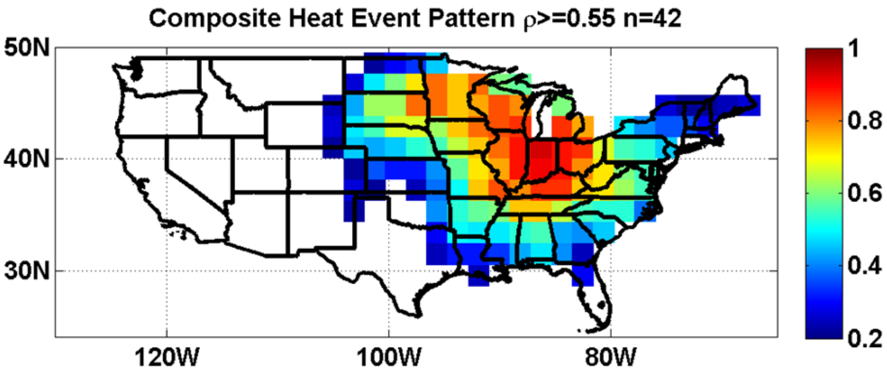


Figure 2: A description of the Chicago July, 1995 heat wave as captured by our definition of heat events. Meteorological data are from the NCEP/NCAR reanalysis. Panels represent: (a) whether the week between 11-17 of July was a heat week in respect to intensities of L1 (90%) in blue, L2 (95%) in red and L3 (98%) in yellow, (b) the duration in days of an L1 heat event during the Chicago event, (c) the start day of the L1 intensity heat wave and (d) locations with abnormal mortality exceeding the 99% percentile of daily mortality for at least one day during the week of 11-17 July, 1995 (red crosses). This description reveals a very intense heat event (L3) covering the upper left quarter of the CONUS which lasted for up to 5 days in the greater Chicago area and propagated from west to east. Abnormal daily mortality exceeding the 99% percentile is collocated with the heat event captured by our method.





<u>Figure 3:</u> The heatwave that occurred in late July 2016 as captured by the monitoring system based on VTEC (upper panel) and the monitoring system based on NCAR/NCEP reanalysis surface data (lower panel). Both systems capture extreme heat in the middle of the country and in the Southwest. The reanalysis based system also captures a heat event over the Rockies but this was not dangerous to human health and therefore it is not appearing in the VTEC based system.



<u>Figure 4:</u> Composite of heat events (0 or 1) resembling the Chicago 1995 heat event. There are 42 cases within the 1948-2015 period satisfying a pattern correlation exceeding 55%.

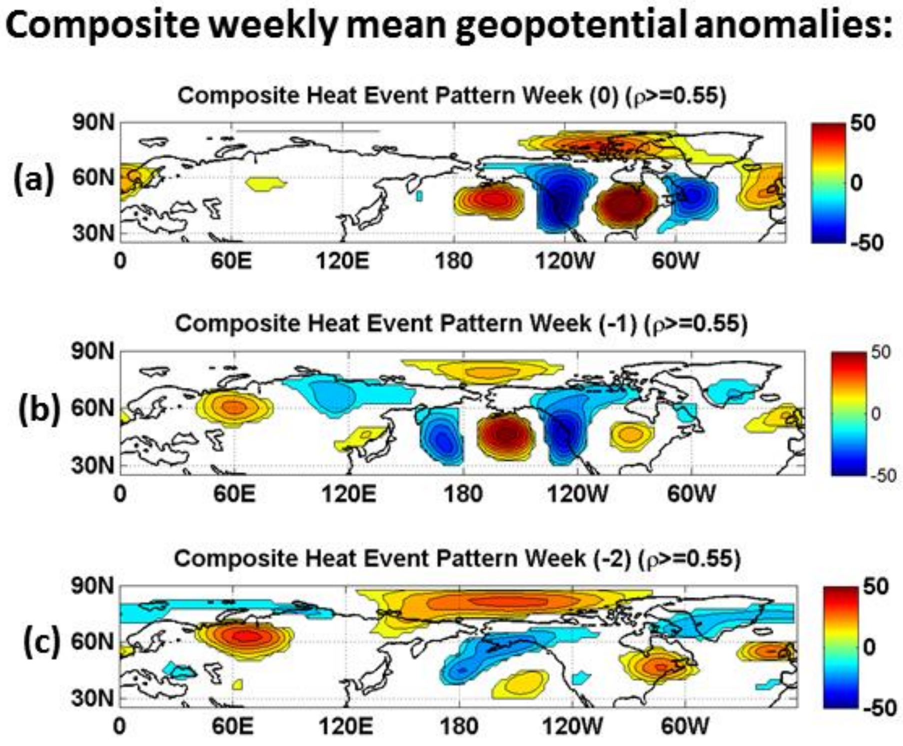


Figure 5: Composites of weekly mean anomalies of 500mb geopotential (in meters) for L1 – Heat Events resembling the Chicago 1995 event. The diagnostic shows a midlatitude high wavenumber structure similar to ones recently reported (Teng et al., 2013; McKinnon et al. 2016). The three panels refer to composites (a) during the week resembling the Chicago July, 1995 event, (b) during the week prior the event and (c) two weeks prior to the event.

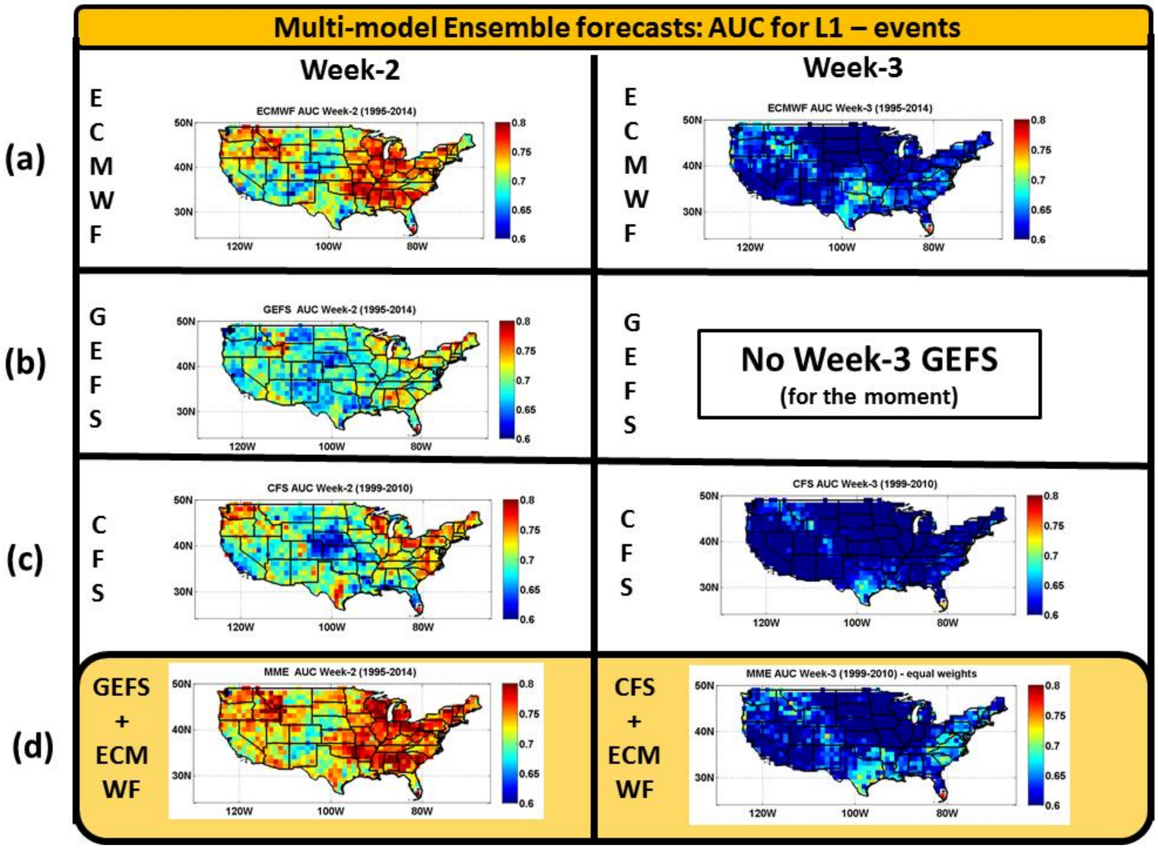


Figure 6: Area Under Curve (AUC) for three different models ECMWF, GEFS and CFS and multi-model combinations respectively in rows (a), (b), (c) and (d). Column 1 shows forecast skill for Week-2 and column 2 for Week-3 forecasts. The multi-model forecast skill for Week-2 includes only the GEFS and ECMWF models. For Week-3 the multi-model uses the ECMWF and CFS models.

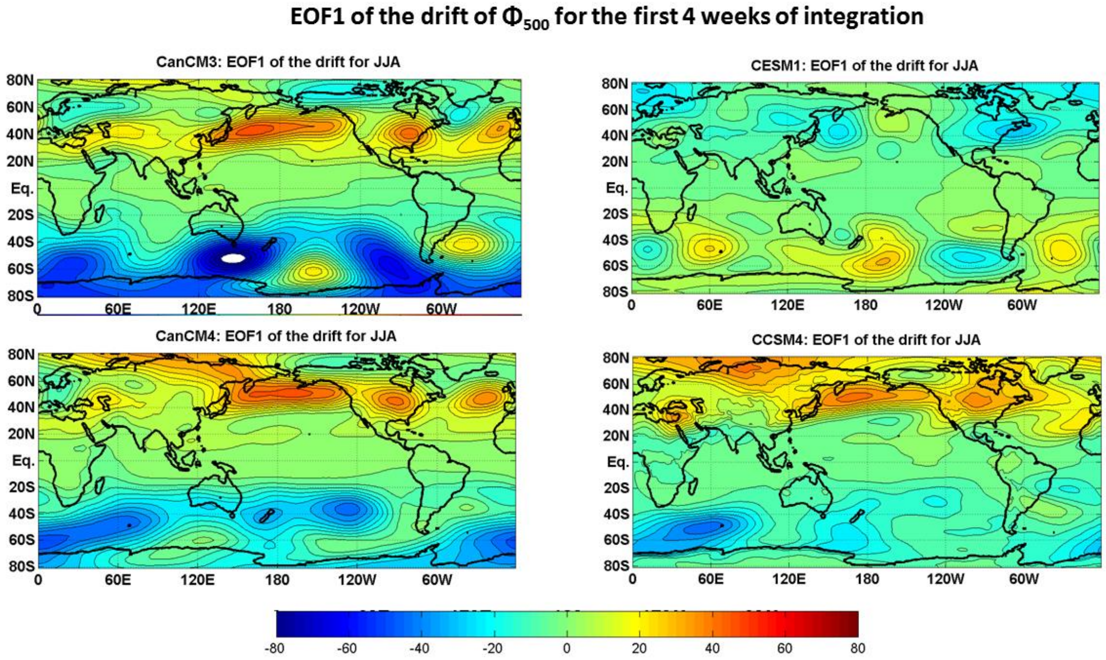


Figure 7: The first EOF for the four additional NMME models. The percentage of variance explained is 95% for the initialized models and 85% for the CESM1. All initialized models show an increase of the geopotential height at 500 hPa by approximately 70 meters over the eastern and central CONUS.

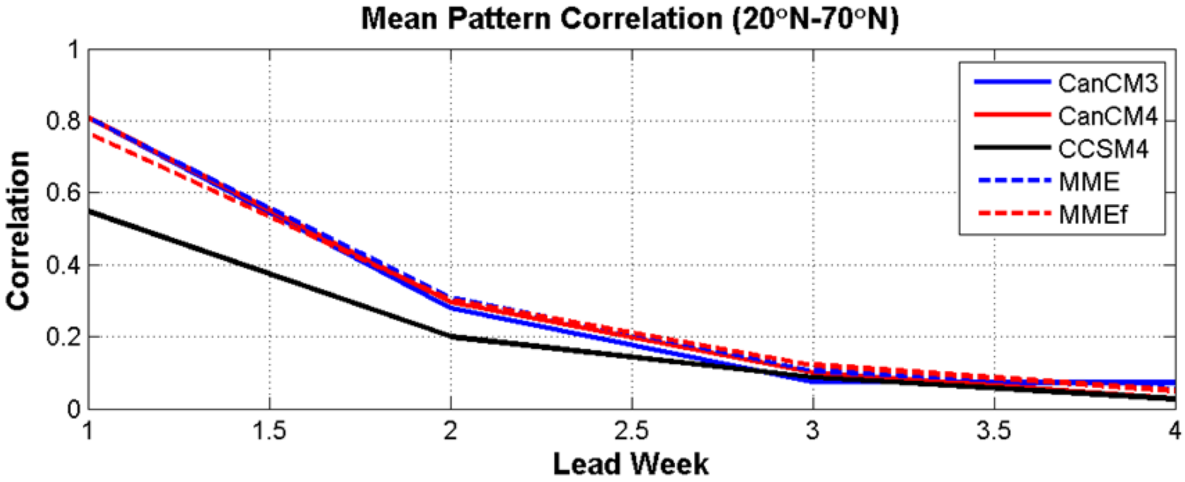


Figure 8: Pattern correlation for northern hemisphere geopotential anomalies at 500 hPa as a function of forecast lead week. Continuous lines show results from individual models. Dashed lines show pattern correlations for the combination of the CanCMs only (blue) and the combination of all models (red). The CCSM4 presents a lower forecast skill for week 1 and 2 as the result of an initial atmospheric state that is not computed by its own analysis but by interpolation from the CFS-Reanalysis and the resulting initialization shocks. The full MME combination shows an improvement albeit very small for forecast Week-3.

9. Highlights of Accomplishments

- This project contributed to the definition of monitoring systems for the SEHOS (VTEC based monitoring system).
- This project contributed to the definition of a verification methodology based on the monitoring system and the Receiver Operating Characteristic (ROC) technique.
- The project demonstrated the utility of using multi-model combinations of the ECMWF, CFS and GEFS models for forecast of heat events at Week-2 and Week-3.
- The project developed a technique for delineating large scale persistent variability modes relevant to excessive heat events.
- The project quantified the subseasonal mean drift of the CanCMs, CCSM4 and CESM1 models and identified similarities between the drift pattern and heat modes.
- The project quantified forecast skill for pattern correlations of geopotential anomalies at 500 hPa and underlined the importance of initialization for the subseasonal forecast lead times.

10. Publications from the Project

In preparation:

Vintzileos et al., 2017 : Large scale modes of variability associated with excessive heat events.
Vintzileos et al., 2017: Multi-model ensemble forecasting of excessive heat events at subseasonal lead times.

Communications:

Vintzileos, A. et al., 2016 Enhancing Resilience to Heat Extremes: Forecasting Excessive Heat Events at Subseasonal Lead Times (Week-2 to 4). 41st Climate Diagnostics and Prediction Workshop, Orono, Maine.
Vintzileos, A. et al., 2016: Extreme Heat and Health, 2016: Towards Multi-Model Ensemble Forecasting of Excessive Heat Events at Subseasonal Lead Times (Week-2 to Week-4). CICS Executive Board Meeting (INVITED), 16 May, College Park, MD.
Vintzileos, A. et al., 2016: Towards Multi-Model Ensemble Forecasting of Excessive Heat Events at Subseasonal Lead Times (Week-2 to Week-4). NOAA/CPO/MAPP Webinar Series (INVITED), 28 April, Silver Spring, MD.
Vintzileos, A. and J. Gottschalck and M. Halpert, 2016: A Baseline System for Forecasting Excessive Heat Events at Subseasonal Lead Times. Climate Prediction Applications Science Workshop. 24-26 March, Burlington, VT
Vintzileos, A. and J. Gottschalck, 2016: Towards a Multi-Model Subseasonal Excessive Heat Outlook System. AMS Annual Meeting, 10-15 January 2016, New Orleans, Louisiana.
Vintzileos, A. and J. Gottschalck, 2015: Towards a Multi-Model Subseasonal Excessive Heat Outlook System. AGU Annual fall meeting, 14-18 December 2015, San Francisco, California.

11. **PI Contact Information**.
e-mail: Augustin.Vintzileos@noaa.gov

cell phone : 240-593-4855