Spatio-Temporal Climate Data Causality Analytics – An Analysis of ENSO's Global Impacts

Hua Song[1], Jianwu Wang[1](jianwu@umbc.edu), Jing Tian[2], Jingfeng Huang[3], Zhibo Zhang[4]

1 Department of Information Systems, UMBC 2 Department of Mathematics, Towson University 3 NOAA NESDIS STAR 4 Department of Physics, UMBC

1. Motivation

- El Niño and the Southern Oscillation (ENSO) is one of the most dominating climate factors that impacts remote weather/climate through atmospheric "teleconnection".
- Most previous studies use the conventional correlation-based methods to study the impacts of ENSO on climate variables. These methods cannot identify cause-and-effect of such linkage.
- Granger causality (GC) approach is suitable to determine the causality relations with high memory data.
- We aim to find the causality relations between ENSO and climate variables by applying the GC approach to observation data.
- We also use the climate model simulation to further support the causality relations between ENSO and climate variables from the observation-based analyses.

2. Methods, Data and Model

2.1 Methods

- **Granger causality calculation via Vector Autoregressive model (VAR):**
  Using VAR(p) to denote an autoregression model of the lag order p, then VAR(p) on two time series *X* and *Y* is defined as:

$$\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} + \begin{pmatrix} d_{11}^1 & d_{12}^1 \\ d_{21}^1 & d_{22}^1 \end{pmatrix} \times \begin{pmatrix} X(t-1) \\ Y(t-1) \end{pmatrix} + \cdots + \begin{pmatrix} d_{11}^p & d_{12}^p \\ d_{21}^p & d_{22}^p \end{pmatrix} \times \begin{pmatrix} X(t-p) \\ Y(t-p) \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

  The VAR can be applied to test the Granger causality of *X* and *Y*: if at least one of the elements $d_{21}^i, i = 1 \cdots p$ are nonzero, then *Y* is Granger caused by *X*.

- **The maximum lag correlation:** $Lag_{corr[\tau]} \doteq \frac{1}{n-\tau}\sum_{k=1}^{n-\tau}\left[\frac{x(k)-\bar{x}}{\sigma_x} \times \frac{y(k+\tau)-\bar{y}}{\sigma_y}\right]$

2.2 Data

- **The Hadley Centre Sea Ice and Sea Surface Temperature data (HadiSST):**
  1870-Present, 1˚x 1˚ Lat-Lon resolution, monthly mean
- **The Global Precipitation Climate Project Precipitation (GPCP) version 2.3:**
  1979-Present, 2.5˚ x 2.5˚ Lat-Lon resolution, monthly mean
- **The NCEP/NCAR reanalysis I surface air temperature:**
  1948-Present, 2.5˚ x 2.5˚ Lat-Lon resolution, monthly mean

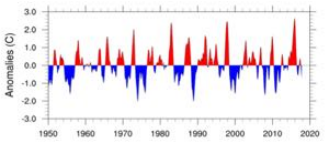

Figure 1. Time series of ENSO index: the running 3-month mean SST anomaly for the Niño 3.4 region (5⁰S-5⁰N, 170⁰W-120⁰W) from 1950 to 2017 with the 1950-2000 as the base period. Unit of SST anomaly is ⁰C.

2.3 Climate Model Simulations

- **The Community Atmospheric Model (version 5.3, CAM5.3):**
  1.9˚ Lat × 2.5˚ Lon resolution with 30 vertical levels and 30-min time step.

36-mon runs
- **CAM5.3-Control**: forced by climatological SST
- **CAM5.3-p2K:** forced by climatological SST + 2 ˚K at the Nino3.4 region
- **CAM5.3-n2K:** forced by climatological SST - 2 ˚K at the Nino3.4 region

3. Impacts of ENSO on Surface Air Temperature and Surface Precipitation

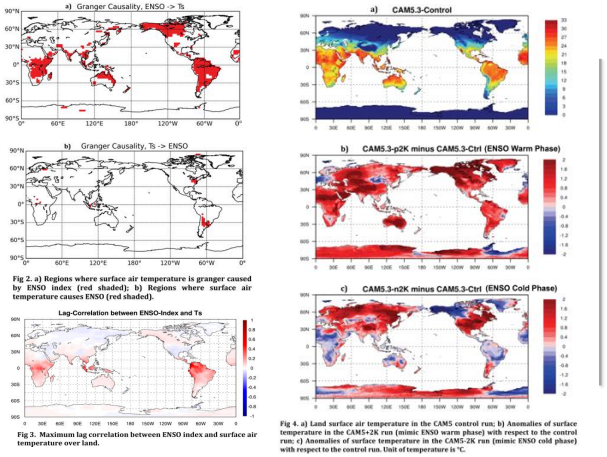

Fig 2. a) Regions where surface air temperature is granger caused by ENSO index (red shaded); b) Regions where surface air temperature causes ENSO (red shaded).

Fig 3. Maximum lag correlation between ENSO index and surface air temperature over land.

Fig 4. a) Land surface air temperature in the CAM5 control run; b) Anomalies of surface temperature in the CAM5+2K run (mimic ENSO warm phase) with respect to the control run; c) Anomalies of surface temperature in the CAM5-2K run (mimic ENSO cold phase) with respect to the control run. Unit of temperature is °C.

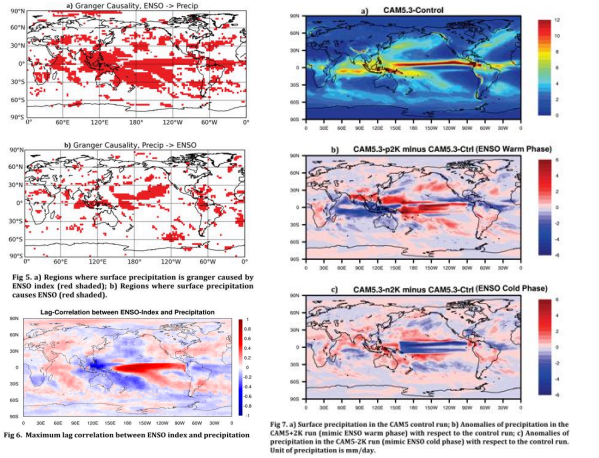

Fig 5. a) Regions where surface precipitation is granger caused by ENSO index (red shaded); b) Regions where surface precipitation causes ENSO (red shaded).

Fig 6. Maximum lag correlation between ENSO index and precipitation

Fig 7. a) Surface precipitation in the CAM5 control run; b) Anomalies of precipitation in the CAM5+2K run (mimic ENSO warm phase) with respect to the control run; c) Anomalies of precipitation in the CAM5-2K run (mimic ENSO cold phase) with respect to the control run. Unit of precipitation is mm/day.

- Figure 2 indicates that ENSO is a driver of surface temperature anomalies in remote regions such as South America, northwest North America, equatorial South Africa, and northern Australia; while ENSO variation is not caused by surface temperature over land.
- ENSO has strong positive relationship with surface temperature in South America and equatorial South Africa, which indicates that El Niño events are most likely accompanied with higher surface temperature over these lands (Figure 3).
- Results of the climate model sensitivity simulations (Figure 4) are consistent with the observational-based analyses. Surface temperature over South America, northwest North America and south Africa show opposite anomalies for the ENSO warm and cold phases.

- ENSO is leading the changes in surface precipitation anomalies in many regions such as tropical Ocean and tropical land, with significant granger causality correlation over broad area in Figure 5(a), but not vice versa in Figure 5(b).
- ENSO has strong negative relationship with surface precipitation in Tropical Western Pacific and tropical South American (Figure 6).
- Model simulations (Figure 7) indicate that in the ENSO warm-phase events, there are positive anomalies (floods) in surface precipitation over Tropical Central and Eastern Pacific, and negative anomalies (droughts) in surface precipitation over Tropical Western Pacific.

4. Conclusions

- We analyzed observational data, reanalysis data and model data to comprehensively investigate the global impacts of ENSO using statistical methods (VAR GC method and maximum lag correlation) and global climate model simulations.
- Results show that the VAR method is able to clearly show ENSO as a cause instead of an effect to influence the remote climate variables and thus cause extreme weather events such as flooding, drought, extreme heat and cold, etc.
- Our model simulations using the CAM5.3 also successfully simulated ENSO's remote impacts on surface temperature and precipitation, consistent to the findings from our observation-based statistical analysis.

Reference and Acknowledgement

- H. Song, J. Wang, J. Tian, J. Huang, Z. Zhang, "Spatio-Temporal Climate Data Causality Analytics – An Analysis of ENSO's Global Impacts", in Proceedings of the *8th International Workshop on Climate Informatics (CI 2018)*, 2018.

- This work was supported by a grant from the National Science Foundation (grant no. OAC–1730250).

8th International Workshop on Climate Informatics
September 19-21, 2018
Hosted by the National Center for Atmospheric Research in Boulder, CO

SPATIO-TEMPORAL CLIMATE DATA CAUSALITY ANALYTICS – AN ANALYSIS OF ENSO'S GLOBAL IMPACTS

Hua Song[1], Jianwu Wang[1], Jing Tian[2], Jingfeng Huang[3], Zhibo Zhang[4,5]

Abstract—**Numerous studies have indicated that El Niño and the Southern Oscillation (ENSO) could have determinant impacts on remote weather and climate using the conventional correlation-based methods, which however cannot identify cause-and-effect of such linkage and ultimately determine a direction of causality. This study employs the Vector Auto-Regressive (VAR) model estimation method with the long-term observational data and reanalysis data to demonstrate that ENSO is the modulating factor that can result in abnormal surface temperature, pressure, precipitation and wind circulation remotely. We also carry out the sensitivity simulations using the Community Atmospheric Model (CAM) to further support the causality relations between ENSO and abnormal climate events in remote regions.**

I. MOTIVATION

El Niño and the Southern Oscillation (ENSO) is a local phenomenon of the variation in sea surface temperature (SST) and surface air pressure across the equatorial eastern Pacific Ocean. Numerous studies have indicated that ENSO could have determinant impacts on remote weather and climate through atmospheric "teleconnection" using the conventional correlation-based methods [1, 2, 3], which however cannot identify cause-and-effect of such linkage and ultimately determine a direction of causality. Lagged linear regression is frequently used to infer causality between climate variables [4, 5, 6]. This method has weaknesses when one or more of the variables have high memory or autocorrelation [7, 8].

Granger causality [9] method, which consists of a lagged autoregression and a lagged multiple linear regression, is suitable to determine the causality relations with high memory data [10]. Recently, the Granger causality (GC) approach has been applied to analyze the causality relationships between climate variables, such as between SST and hurricane strength [11], and between ENSO and Indian monsoon [12]. In this paper, we use the Vector Auto-Regressive (VAR) model estimation method to find the Granger causality relations between ENSO and some climate variables (surface air temperature, sea level pressure (SLP), precipitation and wind). We also use the climate model simulation to double confirm the causality relations between ENSO and climate variables from the observation-based analyses.

This work is carried out in order to determine the spatiotemporal causality relationships between ENSO and abnormal extreme climate events (such as drought/flood) in remote regions, and provide some valuable insights for the prediction of some extreme weather/climate events under different ENSO backgrounds.

Corresponding author: Jianwu Wang, jianwu@umbc.edu
[1]Department of Information Systems, University of Maryland, Baltimore County, [2]Department of Mathematics, Towson University, [3]NOAA NESDIS STAR, [4]Department of Physics, University of Maryland, Baltimore County, [5]Joint Center for Earth Systems Technology, University of Maryland, Baltimore County.

II. METHOD, DATA AND MODEL

Granger causality defines a causal relationship from one time series X to another time series Y if and only if the regression for Y based on past values of both X and Y is statistically significant than the regression only based on past values of Y. Granger causality could be calculated using different approaches such as VAR, LASSO and SIN [13]. We use the VAR method to determine the causality relation between ENSO and climate variables.

Using VAR(p) to denote an autoregression model of the lag order p, then VAR(p) on two time series X and Y is defined as:

$$\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} + \begin{pmatrix} d_{11}^1 & d_{12}^1 \\ d_{21}^1 & d_{22}^1 \end{pmatrix} \times \begin{pmatrix} X(t-1) \\ Y(t-1) \end{pmatrix} + \cdots + \begin{pmatrix} d_{11}^p & d_{12}^p \\ d_{21}^p & d_{22}^p \end{pmatrix} \times \begin{pmatrix} X(t-p) \\ Y(t-p) \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

An F-test is applied to obtain a p-value and significance level to check the statistical significance for whether or not Y is Granger caused by X.

We also calculate the maximum lag correlation between ENSO index and climate variables, which provide the maximum correlation coefficients between ENSO and climate variables and the corresponding lag time.

The monthly-mean observational data we use are the Hadley Centre Sea Ice and Sea Surface Temperature data (HadiSST) from 1870 to the present with 1° x 1° latitude-longitude resolution [14], and the Global Precipitation Climate Project Precipitation (GPCP) version 2.3 data from 1979 to the present at the 2.5° x 2.5° latitude-longitude resolution [15]. We also use the NCEP/NCAR reanalysis I data from 1948 to present at the 2.5° x 2.5° latitude-longitude resolution with 17 vertical levels [16]. The NCEP/NCAR monthly mean horizontal wind, vertical wind, SLP, and air temperature data are used.

Sensitivity simulations with global climate model forced by different SST patterns are carried out to investigate the responses of atmospheric fields to different ENSO phases and further support the causality relations between ENSO and abnormal climate events in remote regions. We use the Community Atmospheric Model (version 5.3, CAM5.3) with the CAM5 standard parameterization schemes [17]. The CAM5.3 uses the finite volume dynamical core at 1.9° latitude × 2.5° longitude resolution with 30 vertical levels and 1800-s time step. The simulation is run using MPI with 32 processors at the UMBC Maya cluster [18]. Three sensitivity simulations include: the CAM5.3-Control run forced by climatological SST; the CAM5.3-p2K run forced by climatological SST + 2K at the Nino3.4 region to mimic ENSO warm phase; the CAM5.3-n2K run is forced by climatological SST – 2K at the Nino3.4 region to mimic ENSO cold phase. We compare the simulated wind, SLP, precipitation and temperature fields from three simulations.

III. EVALUATION

Following [10], we first determine the causality relation between ENSO and surface air temperature on the global land using the VAR method. As the significant differences between Figure 3(a) and Figure 3(b), the changes in ENSO clearly leads the changes in surface air temperature in Figure 3(a), but not vice versa in Figure 3(b). This indicates that ENSO is a driver of surface temperature anomalies in remote regions such as South America, northwest North America, equatorial South Africa, and northern Australia; while ENSO variation is not caused by surface temperature over land. This result is consistent with the study of McGraw and Barnes [10]. The global distribution of the maximum lag correlation between ENSO index and surface temperature (Figure 2) shows that ENSO has strong positive relationship with surface temperature in South America and equatorial South Africa, which indicates that El Niño events (i.e., ENSO warm phase) are most likely accompanied with higher surface temperature over these lands. Results of the climate model sensitivity simulations (Figure 3) are consistent with the observational-based analyses. In the ENSO warm phase (Figure 3(b)), there are positive anomalies in surface temperature over South America, northwest North America while in the ENSO cold phase (Figure 3(c)) there are negative anomalies in surface temperature over these regions.

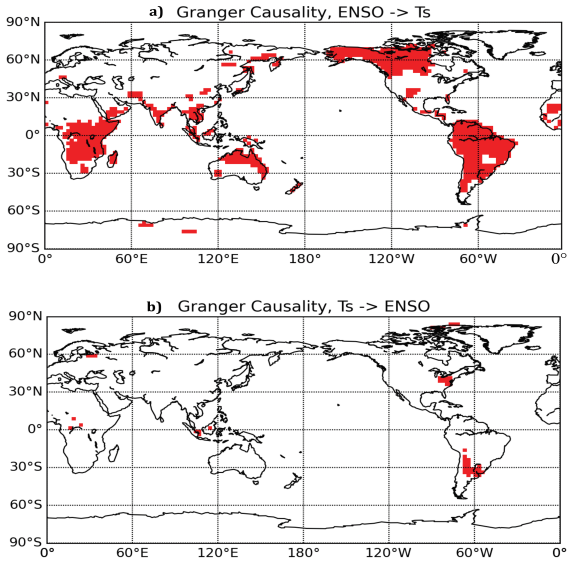

Fig 1. a) Regions where surface air temperature is granger caused by ENSO index (red shaded); b) Regions where surface air temperature causes ENSO (red shaded).

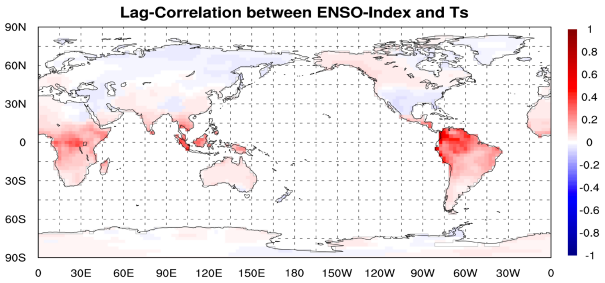

Fig 2. Maximum lag correlation between ENSO index and surface air temperature over land.

To explore the relationship between extreme drought/flood events with ENSO, we analyze the causality relation between ENSO and global surface precipitation. As the comparison between Figure 4(a) and 4(b) shows, ENSO changes is leading the changes in surface precipitation anomalies in many regions such as Tropical Ocean and Tropical Land, with significant Granger causality correlation over broad area in Figure 4(a), but not vice versa in Figure 4(b).

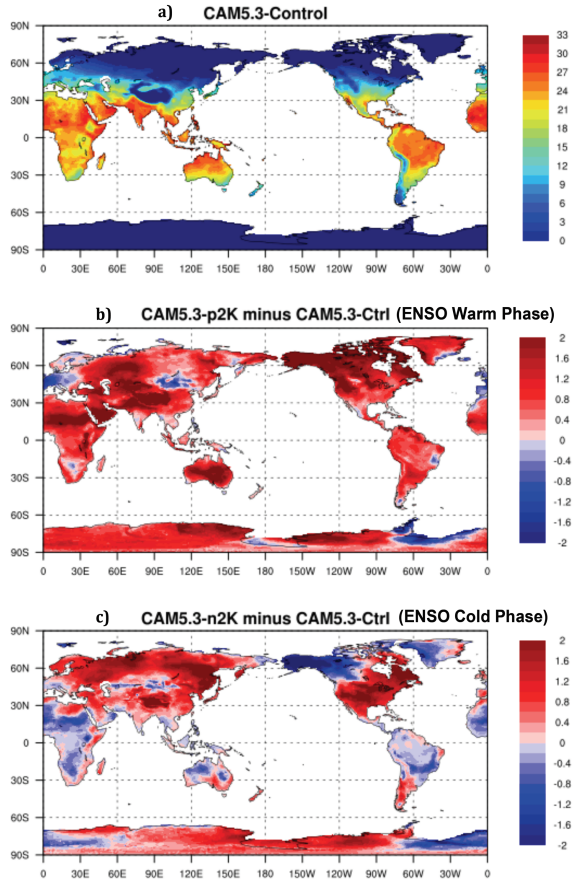

Fig 3. a) Land surface air temperature in the CAM5 control run; b) Anomalies of surface temperature in the CAM5+2K run (mimic ENSO warm phase) with respect to the control run; c) Anomalies of surface temperature in the CAM5-2K run (mimic ENSO cold phase) with respect to the control run. Unit of temperature is °C.

The global distribution of the maximum lag correlation between ENSO index and surface precipitation (Figure 5) shows that ENSO has strong negative relationship with surface precipitation in Tropical Western Pacific and Tropical South American, indicating ENSO's remote impact on extreme drought events. Figure 5 also shows ENSO has strong positive relationship with surface precipitation in Tropical Central and Eastern Pacific, which means ENSO may potentially result in extreme flooding events over these regions. Similarly, the climate model sensitivity simulations (Figure 6) indicate that in the ENSO warm-phase events (Figure 6(b)), there are positive anomalies (floods) in surface precipitation over Tropical Central and Eastern Pacific, and negative anomalies (droughts) in surface precipitation over Tropical Western Pacific, consistent to what we found from the observations. The patterns of precipitation anomalies in the ENSO cold-phase events (Figure 6(c)) are clearly different from those in the ENSO warm-phase events.

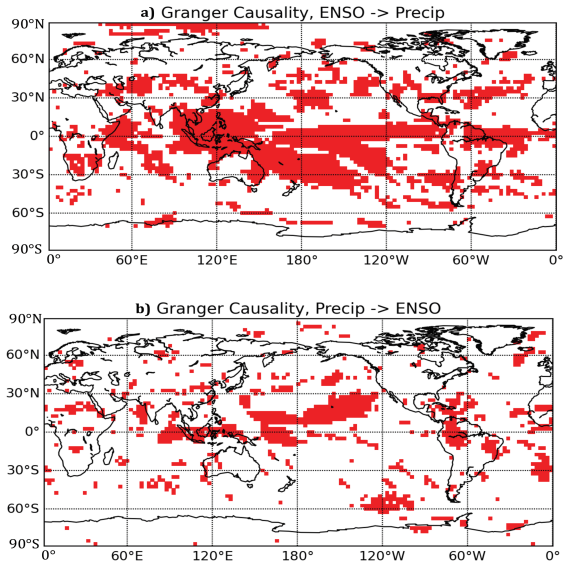

Fig 4. a) Regions where surface precipitation is granger caused by ENSO index (red shaded); b) Regions where surface precipitation causes ENSO (red shaded).

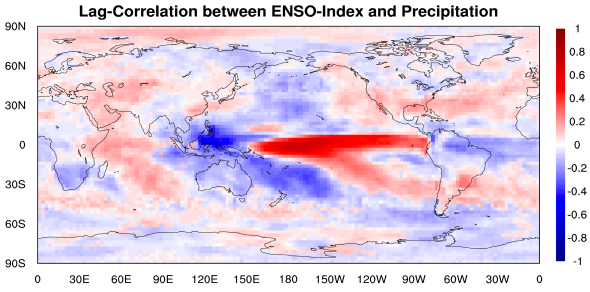

Fig 5. Maximum lag correlation between ENSO index and precipitation

IV. CONCLUSIONS

In this study, we analyzed different observational data, reanalysis data and model data to comprehensively investigate the global impacts of ENSO using statistical methods, namely the VAR GC method, maximum lag correlation, and global climate model simulations. Results show that the VAR method is able to clearly show ENSO as a cause instead of an effect to influence the remote climate variables and thus cause extreme weather events such as flooding,

drought, extreme heat and cold, etc. Our model simulations using the CAM5.3 also successfully simulated ENSO's remote impacts on other weather variables, consistent to the findings from observational evidence. More details of our causality analysis of the ENSO's global impacts can be found in [19].

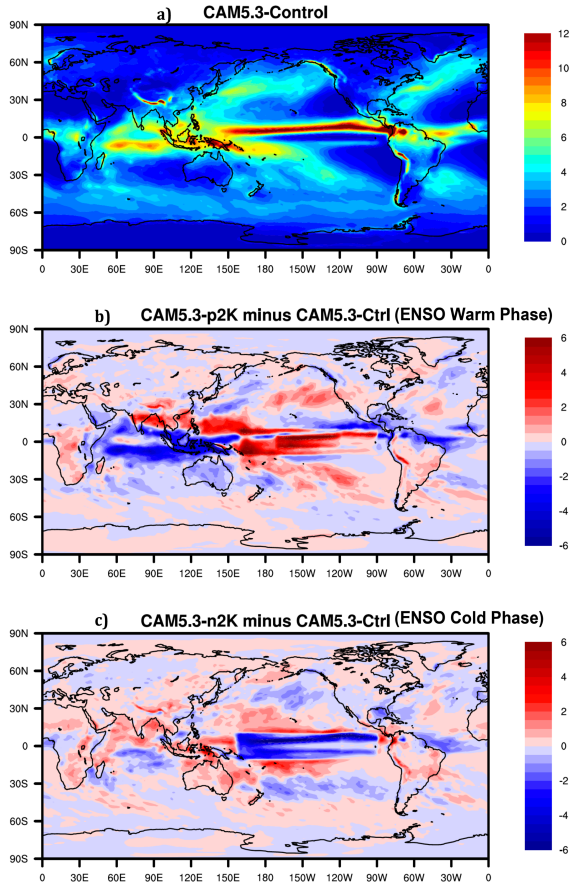

Fig 6. a) Surface precipitation in the CAM5 control run; b) Anomalies of precipitation in the CAM5+2K run (mimic ENSO warm phase) with respect to the control run; c) Anomalies of precipitation in the CAM5-2K run (mimic ENSO cold phase) with respect to the control run. Unit of precipitation is mm/day.

ACKNOWLEDGMENTS

This work was supported by the grant CyberTraining: DSE: Cross-Training of Researchers in Computing, Applied Mathematics and Atmospheric Sciences using Advanced Cyberinfrastructure Resources from the National Science Foundation (grant no. OAC–1730250).

REFERENCES

[1] G. Gu, and R. F. Adler, "Precipitation and temperature variations on the interannual time scale: Assessing the impact of ENSO and volcanic eruptions". *Journal of Climate*, 24, 2258-2270, 2011.

[2] A. Kumar, L. Zhang and W. Wang, "Sea surface temperature - precipitation relationship in different reanalyses", *Monthly Weather Review*, 1118-1123, 2012.

[3] Y. Chen, and coauthors, "Forecasting fire season severity in South America using sea surface temperature anomalies", *Science*, Vol 334., 787-791, 2011.

[4] J. García-Serrano, C. Frankignoul, G. Gastineau, and A. de la Cámara, "On the predictability of the winter Euro-Atlantic climate: Lagged influence of autumn Arctic sea ice", *J. Climate*, 28, 5195-5216, 2015.

[5] H. Tabari, and P. Willems, "Lagged influence of Atlantic and Pacific climate patterns on European extreme precipitation", Scientific Reports, 8, 2018.

[6] H. Tabari, and P. Willems, "Lagged influence of Atlantic and Pacific climate patterns on European extreme precipitation", *Scientific Reports*, 8, 2018.

[7] J. Runge, V. Petoukhov, and J. Kurths, "Quantifying the strength and delay of climatic interactions: The ambiguities of cross correlation and a novel measure based on graphical models", *Journal of Climate*, 27, 720-739, 2014.

[8] M. Kretschmer, D. Coumou, J.F. Donges, and J. Runge, "Using causal effect networks to analyze different Arctic drivers of midlatitude winter circulation", *Journal of Climate*, 29, 4069-4081, 2016.

[9] C. Granger, "Investigating causal relations by econometric models and cross-spectral methods", *Econometrica*, 37, 424-438, 1969.

[10] M. McGraw, and E. A. Barnes, "Memory matters: A case for Granger causality in climate variability studies", *Journal of Climate*, 3289-3300, 31, 2018.

[11] J. Elsner, "Granger causality and Atlantic hurricanes", *Tellus*, 59A, 476–485, 2007.

[12] I. Mohkov, D. Smirnow, P. Nakonechny, S. Kozlenko, E. Selezn, and J. Kurths, "Alternating mutual influence of El-Niño/Southern Oscillation and Indian monsoon", *Geophys. Res. Lett.*, 38, 2011.

[13] N. A. Rayner, D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, "Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century", *J. Geophys. Res.*, 108, 2003.

[14] A. Arnold, Y. Liu, and N. Abe, "Temporal causal modeling with graphical Granger methods," in Proceedings of the *13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Jose, California, USA, pp. 66-75, 2007.

[15] Adler, R., and Co-authors, "New global precipitation climatology project monthly analysis product corrects satellite data drifts", *GEWEX News*, 26 (4), 7-9, 2016.

[16] E. Kalnay, and Co-authors, "The NCEP/NCAR 40-year reanalysis project", *Bull. Amer. Meteor. Soc.*, 77, 437-470, 1996.

[17] R. B. Neale, and Co-authors, "Description of the NCAR community atmosphere model (CAM 5.0)", Tech. Rep. TN–486+STR, 268 pp., Natl. Cent. for Atmos. Res., Boulder, Colorado, 2010.

[18] High Performance Computing Facility, UMBC: http://hpcf.umbc.edu/

[19] H. Song, J. Tian, J. Huang, J. Wang, and Z. Zhang, "Causality analysis of ENSO's global impacts on climate variables based on data-driven analytics and climate model simulation", Technical Report HPCF-2018-14, UMBC High Performance Computing Facility, University of Maryland, Baltimore County, 2018.