

# Causality Analysis of ENSO's Global Impacts on Climate Variables based on Data-driven Analytics and Climate Model Simulation

CyberTraining: Big Data + High-Performance Computing + Atmospheric Sciences

Hua Song<sup>1</sup>, Jing Tian<sup>2</sup>, Jingfeng Huang<sup>3</sup>,  
Faculty mentor: Jianwu Wang<sup>4</sup>, Zhibo Zhang<sup>5</sup>

<sup>1</sup>Joint Center for Earth Systems Technology, UMBC

<sup>2</sup>Department of Mathematics, Towson University

<sup>3</sup>NOAA NESDIS STAR

<sup>4</sup>Department of Information Systems, UMBC

<sup>5</sup>Department of Physics, UMBC

Technical Report HPCF-2018-14, [hpcf.umbc.edu](http://hpcf.umbc.edu) > Publications

## 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 the 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 sea surface temperature (SST) data and the NCEP/NCAR reanalysis data to demonstrate the Granger causality between ENSO and other climate attributes. Results showed that ENSO as the modulating factor can result in abnormal surface temperature, pressure, precipitation and wind circulation remotely, not vice versa. We also carry out the global climate model sensitivity simulations using the parallel computing techniques to double confirm the causality relations between ENSO and abnormal events in remote regions. Our statistical and climate model-based analyses may enrich our current understanding on the occurrences of extreme events worldwide caused by different ENSO strengths through teleconnections.

# 1 Introduction

El Niño and the Southern Oscillation (ENSO) is a local phenomenon of the variation in sea surface temperature (SST) and air pressure across the equatorial eastern Pacific Ocean. ENSO is strongly linked to remote weather and climate far away over other parts of the world through the atmospheric “teleconnection”. There are different observation attributes that reflect abnormal extreme weather/climate events, such as flooding, drought, extreme heat or cold, severe storms, etc. The relation between climate factors such as ENSO and the above-mentioned extreme events taking place somewhere else in the world is not easy to understand comprehensively.

Over the past several decades, ENSO has been found as one of the most dominating climate factors that impacts remote weather and climate through the atmospheric “teleconnection” using the conventional correlation-based methods [1, 2, 3]. These methods are useful to establish how they are linked or correlated in the spatio-temporal pattern, but cannot identify the 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 may lead to non-accurate results when one or more of the variables have high memory or autocorrelation [7, 8].

Granger causality method [9], which consists of a lagged autoregression and a lagged multiple linear regression, is suitable to determine the causality relations with high memory data [10, 11]. Recently, Granger causality has been applied to analyze the causality relationships between climate variables, such as between SST and hurricane strength [12], and between ENSO and Indian monsoon [13]. In this paper, we use the Vector Autoregressive (VAR) model estimation method to explore 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.

In this study, we aim to determine the spatio-temporal causality relationships between ENSO and abnormal events in remote regions, and provide some valuable insights for the prediction of several extreme weather/climate events under different ENSO backgrounds. We hypothesize that ENSO is one of the modulating factors of the extreme weather and

climate events, and the causality can be statistically demonstrated using observational datasets and can be consistently simulated using climate models.

The report is structured in the following sections. Section 2 lists the datasets used for the study. Section 3 introduces the Granger causality methods and global climate model simulations. Section 4 reports the main findings from our study, followed by Section 5 that discusses and concludes the study.

## 2 Data

### 2.1 Sea surface temperature data

For this study, we use the Hadley Centre Sea Ice and Sea Surface Temperature data (HadISST). The HadISST data utilize both in situ SST from ships and buoys, and bias-adjusted SST from the satellite-borne advanced very high-resolution radiometer (AVHRR). But the satellite SST only started in late 1981 after AVHRR was launched. The data include monthly mean SST and sea ice extent from 1870 to the present with  $1^\circ \times 1^\circ$  latitude-longitude resolution [14].

Same as the ENSO indices defined by National Oceanic and Atmospheric Administration (NOAA), we use the SST in the Niño 3.4 region ( $5^\circ\text{S}$ - $5^\circ\text{N}$ ,  $170^\circ\text{W}$ - $120^\circ\text{W}$ ) to derive the ENSO index (Figure 1). A full-fledged El Niño or La Niña will be classified when the anomalies exceed  $+0.5^\circ\text{C}$  or  $-0.5^\circ\text{C}$  for at least five consecutive months.

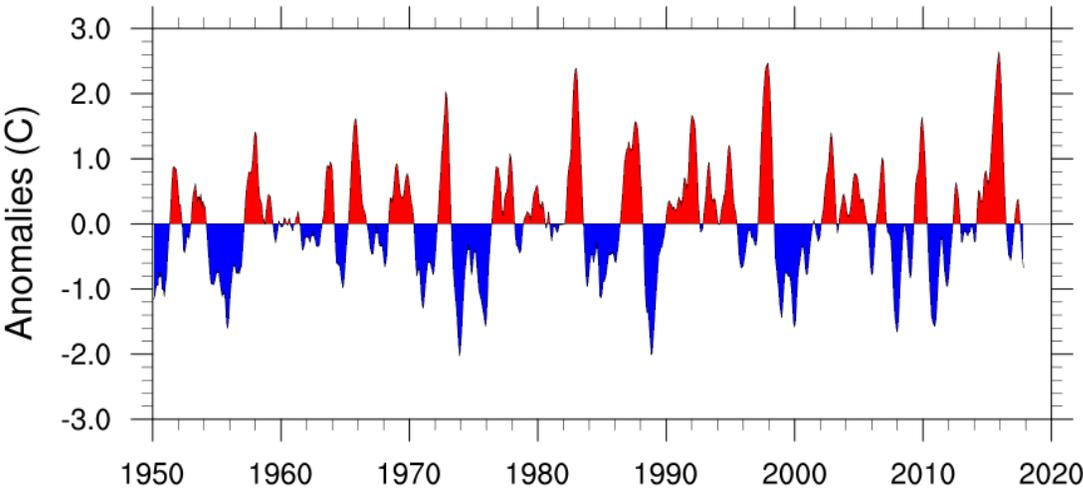


Figure 1. The running 3-month mean SST anomaly for the Niño 3.4 region from 1950 to 2017 with the 1950-2000 as the base period. Unit of SST anomaly is degree Celsius.

## 2.2 Meteorology reanalysis data

The NCEP/NCAR reanalysis I data are employed in this study. This reanalysis data is produced using a state-of-the-art analysis/forecast system that performs data assimilation using past observational data from 1948 to the present. The data span from 1948 to present at the  $2.5^\circ \times 2.5^\circ$  latitude-longitude resolution with 17 vertical levels [15]. To investigate the relation between climate variables and ENSO, we use NCEP/NCAR monthly mean surface air temperature, sea level pressure (SLP), and wind data for any identification of extreme heat or cold events, anomalies in large-scale atmospheric circulation etc.

## 2.3 Precipitation data

For flooding and drought extreme events, we use the Global Precipitation Climate Project Precipitation (GPCP) version 2.3 data from 1979 to the present at the  $2.5^\circ \times 2.5^\circ$  latitude-longitude resolution [16]. The GPCP monthly product provides a consistent analysis of global precipitation from an integration of various satellite datasets over land and ocean, and a gauge analysis over land. Observational data from rain gauge stations, satellite and sounding observations are merged to estimate monthly rainfall.

# 3 Methods and Model

In this study, we use two statistic methods (Granger causality method and Maximum lag correlation) and a global climate model (Community Atmospheric Model) to investigate the global impacts of ENSO on the climate variables.

## 3.1 Granger causality method

Granger causality method was first introduced by Granger [9] to predict economics. It defines a causal relationship from one time series  $\mathbf{x}$  to another time series  $\mathbf{y}$ . Given two time series  $\mathbf{x}$  and  $\mathbf{y}$ , we first perform the autoregression (Eq. (1)) to find the proper lagged value  $k$  of  $\mathbf{y}$ , then we perform the multiple linear regression (Eq. (2)).

$$y(t) = c_0 + c_1 \times y(t-1) + \dots + c_k \times y(t-k) + \varepsilon_t \quad (1)$$

$$y(t) = a_0 + a_1 \times y(t-1) + \dots + a_k \times y(t-k) + b_l \times x(t-l) + \dots + b_n \times x(t-n) + \varepsilon_t \quad (2)$$

The lagged values  $l$  and  $n$  of  $\mathbf{x}$  depend on the significance of the statistical test. If no lagged values of  $\mathbf{x}$  are retained in (Eq. (2)) (i.e. all the coefficients of the  $x$  terms are 0), we can conclude that  $\mathbf{y}$  is not Granger caused by  $\mathbf{x}$ . Granger causality could be calculated using different approaches such as vector autoregressive model (VAR), Graphical Lasso and SIN methods [17]. In this study, we use the VAR to determine the causality relation between ENSO and climate variables.

### 3.1.1 Vector autoregressive model

An autoregressive model is usually used to measure the dependency of a variable on its own previous values.

Using AR( $s$ ) to denote an autoregressive model of order  $s$ , then the AR( $s$ ) on a series  $x(t)$  is defined as

$$x(t) = \phi_0 + \phi_1 \times x(t-1) + \phi_2 \times x(t-2) + \dots + \phi_s \times x(t-s) + \varepsilon_t \quad (3)$$

The vector autoregressive model is a particular case of the autoregressive model: VAR is used when we have more than one variable. Therefore, we will have the autoregressive model (Eq. (3)) on a vector. Given two time series  $\mathbf{x}$  and  $\mathbf{y}$ , VAR( $p$ ) on  $\mathbf{x}$  and  $\mathbf{y}$  is defined as

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} + \begin{pmatrix} d_{11}^1 & d_{12}^1 \\ d_{21}^1 & d_{22}^1 \end{pmatrix} \times \begin{pmatrix} x(t-1) \\ y(t-1) \end{pmatrix} + \dots + \begin{pmatrix} d_{11}^p & d_{12}^p \\ d_{21}^p & d_{22}^p \end{pmatrix} \times \begin{pmatrix} x(t-p) \\ y(t-p) \end{pmatrix} + \begin{pmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{pmatrix} \quad (4)$$

The VAR can be applied to test the Granger causality of  $\mathbf{x}$  and  $\mathbf{y}$ : if at least one of the elements  $d_{21}^i$ ,  $i = 1 \dots p$  in (Eq. (4)) are nonzero, then  $\mathbf{y}$  is Granger caused by  $\mathbf{x}$ .

### 3.1.2 Implementation of the vector autoregressive model

In this study, we use the VAR package in Python to implement the vector autoregressive model.

The lag order  $p$  is selected by an information criteria-based order selection. We choose the maximum number of lags to be 12 since there are 12 months per year. Then, the Schwarz's criterion (Bayesian information criterion) [18] is used to select the "optimal" lag based on the data. Therefore, for different data set, the  $p$  value varies.

The F-test is used to check the statistical significance. When we test the Granger causality of  $\mathbf{x}$  and  $\mathbf{y}$ , the null hypothesis is:  $\mathbf{y}$  is not Granger caused by  $\mathbf{x}$  (this is equivalent to all of the coefficients of  $\mathbf{x}$  in the regression relation are 0). When we apply the F-test in the VAR, we perform two regression models: an unrestricted model (Eq. (2)) and a

restricted model (Eq. (1)). The F-test will give us two values: test statistic and critical value. The test statistic is also called F-statistic, giving by:

$$f = \frac{(R_u^2 - R_r^2)/\gamma}{R_u^2/(\theta - \delta)} \quad (5)$$

In (Eq. (5)),  $R^2$  is the coefficient of determination which measures the strength of the linear relationship in the regression.  $R_u^2$  is the  $R^2$  value from the unrestricted regression model (Eq. (2));  $R_r^2$  is the  $R^2$  value from the restricted regression model (Eq. (1)).  $\gamma$  is the number of restrictions that depends on the lag order.  $\theta$  is the number of observations and  $\delta$  is the number of explanatory variables in the unrestricted model. In the F-test, this test statistic  $f$  will follow the F-distribution under the null hypothesis. The critical value can be obtained from the F distribution table. We compare the two values: if the test statistic is greater than the critical value, we reject the null hypothesis. When the test statistic is greater than the critical value, the unrestricted regression model (Eq. (2)) is more statistically significant than the restricted regression model (Eq. (1)). Therefore, we can use this procedure to determine the significance of Granger causality.

Another way to use the F-test is to compare two values: the significance level  $\alpha$  and the p-value  $p$ . The significance level is the probability of the study rejecting the null hypothesis; the p-value is the probability of obtaining a result at least as extreme, giving the null hypothesis were true. When we have  $p < \alpha$ , we can conclude the result is statistically significant. In most cases, we choose the significance level to be 0.05. Therefore, if the p-value we get from the F-test is smaller than 0.05, we reject the null hypothesis which leads to the conclusion that  $\mathbf{y}$  is Granger caused by  $\mathbf{x}$ . In fact, when we have the test statistic is greater than the critical value, we will also have the p-value less than  $\alpha$ . Therefore, the two approaches match with each other.

### 3.2 Maximum lag correlation

To compare with and to complement the Granger causality model, we also calculate the maximum lag correlation (i.e., cross correlation) between ENSO index and climate variables. It provides the maximum correlation coefficients between ENSO and climate variable and the corresponding lag time. The lag correlation coefficient between two series  $\mathbf{x}(\mathbf{n})$  and  $\mathbf{y}(\mathbf{n})$  is defined by Eq. (6), in which  $\tau$  is the lag time,  $\bar{x}$  and  $\sigma_x$  are the mean and standard

deviation of the series  $\mathbf{x}$  respectively, and  $\bar{y}$  and  $\sigma_y$  are the mean and standard deviation of the series  $\mathbf{y}$  respectively.

$$Lag_{corr}[\tau] \doteq \frac{1}{n-\tau} \sum_{k=1}^{n-\tau} \left[ \frac{x(k)-\bar{x}}{\sigma_x} \times \frac{y(k+\tau)-\bar{y}}{\sigma_y} \right] \quad (6)$$

### 3.3 Global climate model

Based on physical hypothesis and sophisticated schemes, climate model is frequently used to find out the impact of a causing factor as effects on other parameters. For this study, series of sensitivity simulations are carried out with the global climate model forced by different simulated ENSO-like SST patterns to see the corresponding responses of atmospheric fields. The climate model we use in this study is the Community Atmospheric Model (version 5.3, CAM5.3) with the CAM5 standard parameterization schemes [19]. The CAM5.3 uses the finite volume dynamical core at  $1.9^\circ$  latitude  $\times$   $2.5^\circ$  longitude resolution with 30 vertical levels and 1800-s time step. Simulations for three ENSO scenarios are carried out: 1) the control run forced with climatological SST; 2) the p2K run forced with climatological SST +  $2^\circ\text{C}$  at the Nino3.4 region and climatological SST +  $1^\circ\text{C}$  at the central Pacific ( $10^\circ\text{S}$ - $10^\circ\text{N}$ ,  $160^\circ\text{E}$ - $90^\circ\text{W}$ ); and 3) the n2K run forced with climatological SST -  $2^\circ\text{C}$  at the Nino3.4 region and climatological SST +  $1^\circ\text{C}$  at the central Pacific ( $10^\circ\text{S}$ - $10^\circ\text{N}$ ,  $160^\circ\text{E}$ - $90^\circ\text{W}$ ). The global distribution of SST forcing used in three simulations are shown in Figure 2. The three simulations are run using MPI with 32 processors at UMBC Maya cluster [20]. Each simulation is integrated for 36 months, and the last 24-month simulation outputs are used for analysis. We compare the changes in wind, SLP, cloud, precipitation and temperature fields from three simulations to our observation-based results using the statistical methods (i.e., Granger causality method and maximum lag correlation method) for consistency and discrepancy identifications.

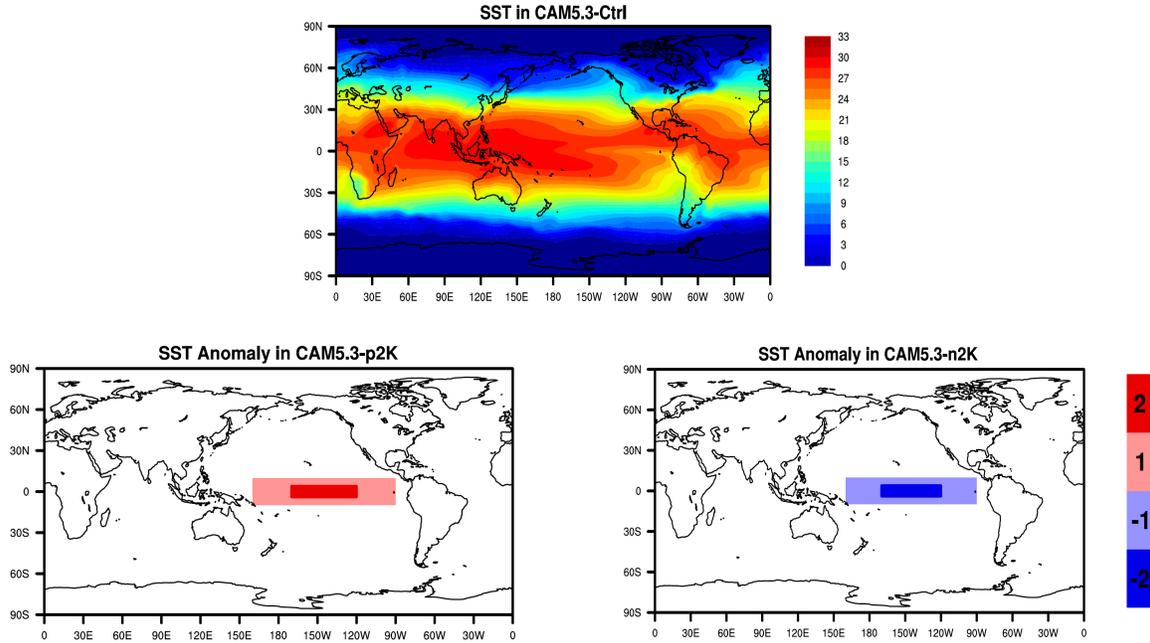


Figure 2. Global distribution of annual mean SST used in the control run (upper panel). The SST anomaly used in the +2K run with respect to the control run (left lower panel). The SST anomaly used in the -2K run with respect to the control run (right lower panel). Unit of SST is degree Celsius.

## 4 Results

### 4.1 ENSO vs. Surface Air Temperature (SAT)

First, we determine the cause-and-effect relation between ENSO and SAT on the global scale using the VAR method for Granger causality model. As the significant differences between Figure 3(a) and 3(b) are found, the changes in ENSO index clearly leads the changes in SAT in Figure 3(a), but not vice versa in Figure 3(b). This indicates SAT changes are granger caused by ENSO, and ENSO is attributable for SAT anomalies, such as extreme heat or cold events, in remote regions such as South America, northwest North America, equatorial South Africa, and northern Australia. But ENSO variation is not caused by surface temperature over land. This result is consistent with the study of McGraw and Barnes [11]. The global distribution of the maximum lag correlation between ENSO index and surface temperature (Figure 4) 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 5) are consistent with the observational-based analyses. In the ENSO warm-phase events,

there are positive anomalies in surface temperature over South America, northwest North America while in the ENSO cold-phase events there are negative anomalies in surface temperature over these regions.

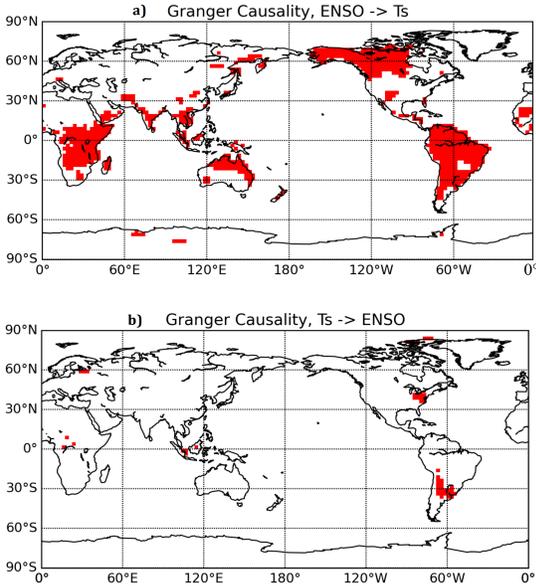


Fig 3. a) Regions where surface air temperature is granger caused by ENSO index (red shaded); b) Regions where surface air temperature causes ENSO (red shaded).

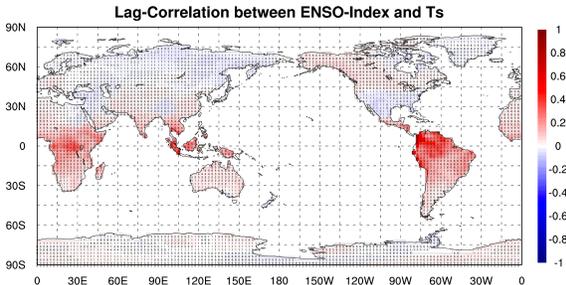


Fig 4. Maximum lag correlation between ENSO index and surface air temperature over land.

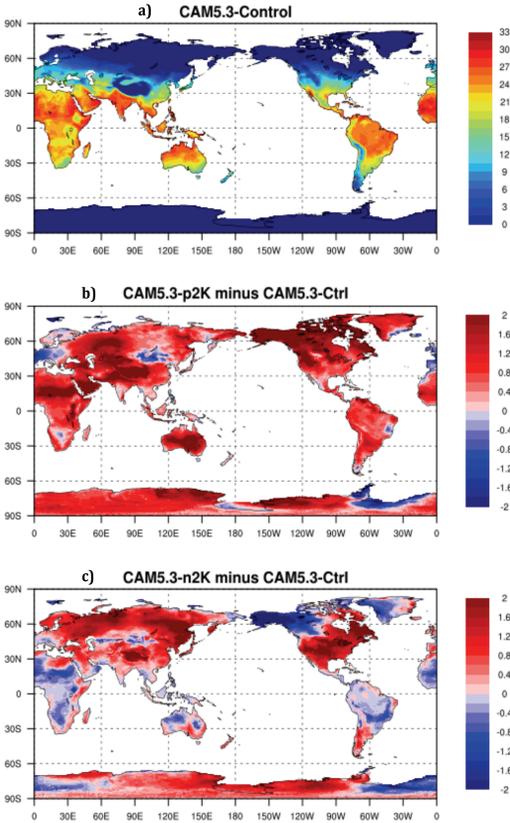


Fig 5. a) Land surface air temperature in the CAM5 control run; b) Anomalies of surface temperature in the CAM5+2K run with respect to the control run; c) Anomalies of surface temperature in the CAM5-2K run with respect to the control run. Unit of temperature is °C.

### 4.2 ENSO vs. Sea Level Pressure (SLP)

In this study, we also investigate the causality relation between ENSO and SLP on the global scale for their spatio-temporal patterns. Comparing Figure 6(a) to 6(b), we clearly observe that ENSO changes is leading the SLP changes and causing the SLP anomalies in the local and remote regions such as Pacific Ocean, Indian Ocean, and central Atlantic Ocean in Figure 6(a); while ENSO changes are only caused by SLP over eastern Indian Ocean, tropical northern Pacific Ocean and southeastern Pacific Ocean, but at a much less

significant level in Figure 6(b). This is another indication that ENSO is the modulating factor in the cause-and-effect analysis with SLP, not vice versa.

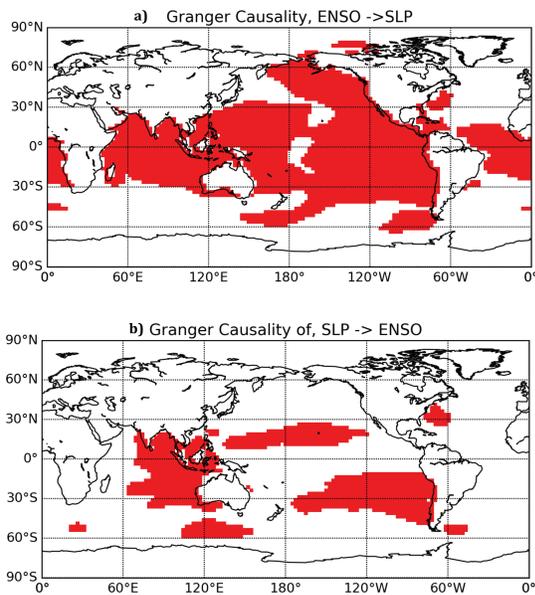


Fig 6. a) Regions where sea level pressure is granger caused by ENSO index (red shaded); b) Regions where sea level pressure causes ENSO (red shaded).

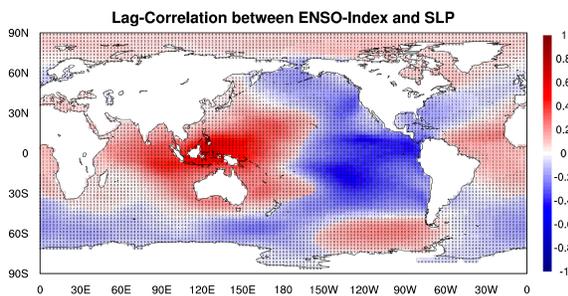


Fig 7. Maximum lag correlation between ENSO index and sea level pressure

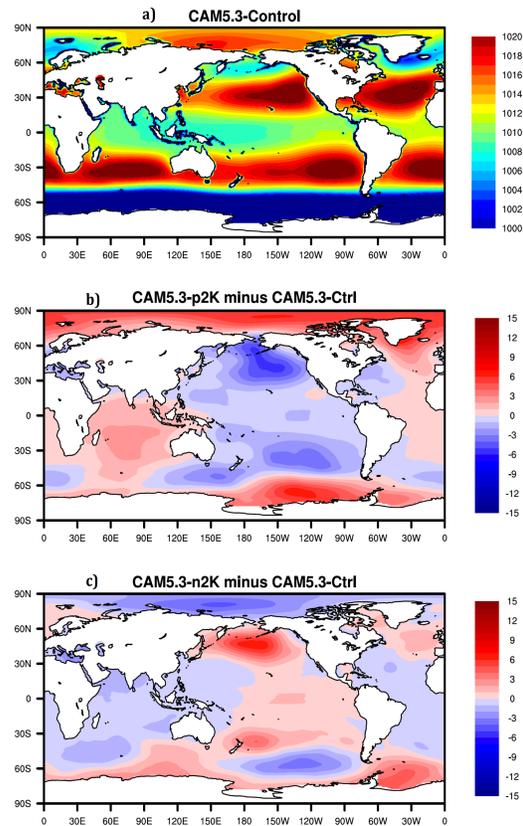


Fig 8. a) Sea level pressure in the CAM5 control run; b) Anomalies of SLP in the CAM5+2K run with respect to the control run; c) Anomalies of SLP in the CAM5-2K run with respect to the control run. Unit of SLP is hPa.

The global distribution of the maximum lag correlation between ENSO index and SLP (Figure 7) shows that ENSO has strong positive relationship with SLP in the Tropical Western Pacific Ocean and strong negative relationship with SLP in the Eastern Pacific Ocean, which implies that El Niño events (i.e., ENSO warm phase) are most likely accompanied with higher SLP over the Tropical Western Pacific Ocean and lower SLP over the Eastern Pacific Ocean. Results of the climate model sensitivity simulations (Figure 8) are consistent with the observational-based analyses. In the ENSO warm-phase events, there are negative anomalies in SLP over Eastern Pacific Ocean while in the ENSO cold-phase events there are positive anomalies in SLP over Eastern Pacific Ocean and negative

anomalies in SLP over Tropical Western Pacific Ocean and Indian Ocean. Such consistency between observational evidence and model simulations give us more confidence that ENSO is modulating SLP remotely as the causing factor, thus the spatio-temporal patterns with ENSO leading SLP can be more readily explained as ENSO causing anomalies in SLP over other parts of the world.

### **4.3 ENSO vs. Precipitation**

To explore the relationship between extreme flooding and drought with ENSO, we analyze the causality relation between ENSO and surface precipitation on the global scale. As the comparison between Figure 9(a) and 9(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 9(b), but not vice versa in Figure 9(b). The global distribution of the maximum lag correlation between ENSO index and surface precipitation (Figure 10) 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 10 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 11) 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, consistent to what we found from the observations. The patterns of precipitation anomalies in the ENSO cold-phase events are clearly different from those in the ENSO warm-phase events.

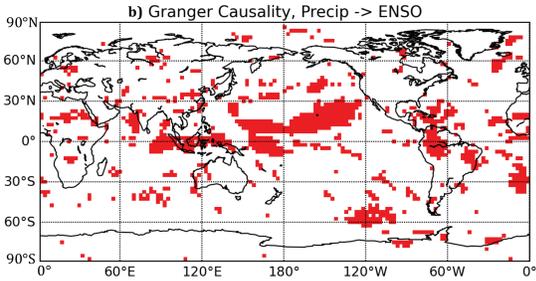
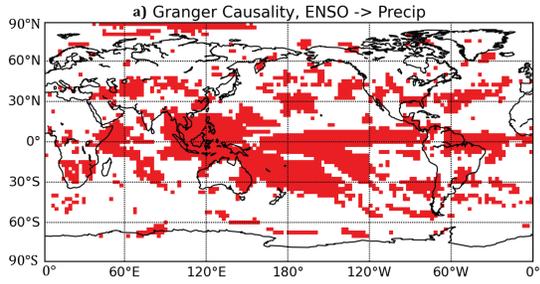


Fig 9. a) Regions where surface precipitation is granger caused by ENSO index (red shaded); b) Regions where surface precipitation causes ENSO (red shaded).

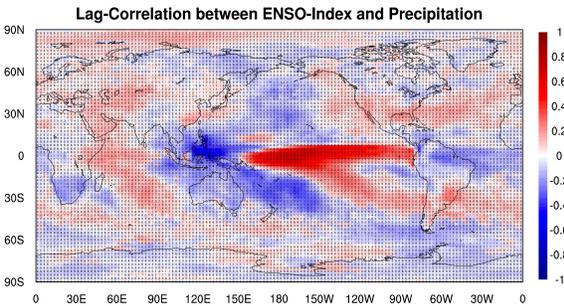


Fig 10. Maximum lag correlation between ENSO index and precipitation

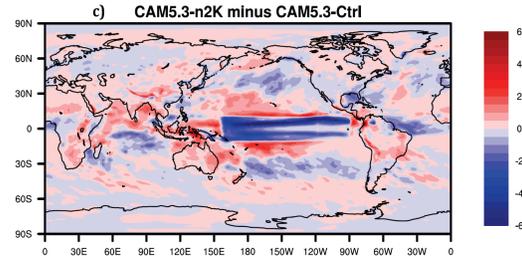
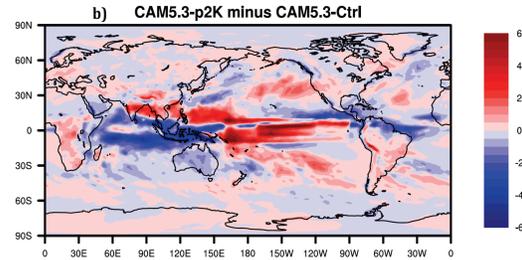
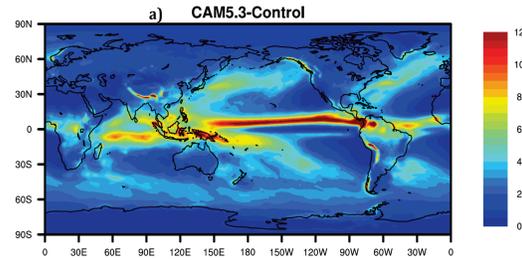


Fig 11. a) Surface precipitation in the CAM5 control run; b) Anomalies of precipitation in the CAM5+2K run with respect to the control run; c) Anomalies of precipitation in the CAM5-2K run with respect to the control run. Unit of precipitation is mm/day.

#### 4.4 ENSO vs. Circulation

Occurrence of different climate events strongly depends on the large-scale atmospheric circulation. Mid-tropospheric (500hPa) vertical pressure velocity is widely used as a proxy for the large-scale tropical circulation [21]. Lastly for this study, we explored the causality relation between ENSO and 500hPa vertical pressure velocity using the same VAR method for granger causality. As shown in Figure 12 with 12(a) versus 12(b), ENSO seems to cause 500hPa vertical velocity anomalies significantly over very broad areas in most Tropical Pacific Ocean in Figure 12(a), but not vice versa in Figure 12(b). The global distribution of the maximum lag correlation between ENSO index and 500hPa vertical velocity (Figure 13) is almost same as that for surface precipitation (Figure 10) but with opposite signs, i.e., ENSO has strong positive relationship with 500hPa vertical velocity in Tropical Western Pacific and tropical South American, an indication of ENSO warm phase suppresses deep

convection and cold phase invigorates deep convection instead. At the same time, ENSO has strong negative relationship with vertical wind over Tropical Central and Eastern Pacific, which means ENSO warm phase potentially invigorates deep convection over these areas and cold phase suppresses deep convection instead. Results of the climate model sensitivity simulations (Figure 14) are considerably similar to those of the observational-based analyses in Figures 12 and 13. In the ENSO warm-phase events, there are positive anomalies in 500hPa vertical velocity over the Tropical Western Pacific and negative anomalies in 500hPa vertical velocity over the Tropical Central and Eastern Pacific. Putting the results in the context of Pacific atmospheric circulation, ENSO warm phase weakens Pacific walker circulation but ENSO cold phase enhances Pacific walker circulation, consistent to what is reported in Yeh et al. [22].

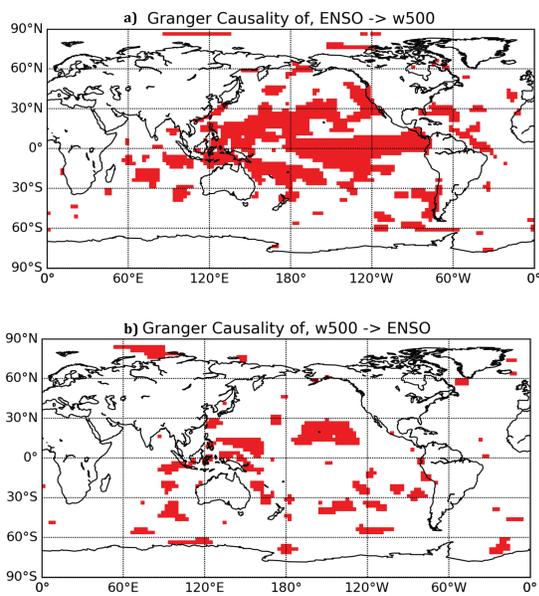


Fig 12. a) Regions where 500hPa vertical velocity is granger caused by ENSO index (red shaded); b) Regions where 500hPa vertical velocity causes ENSO (red shaded).

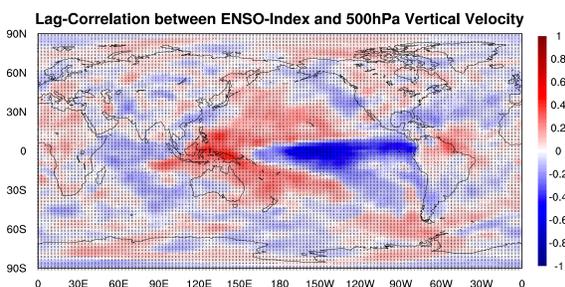


Fig 13. Maximum lag correlation between ENSO index and 500hPa vertical velocity

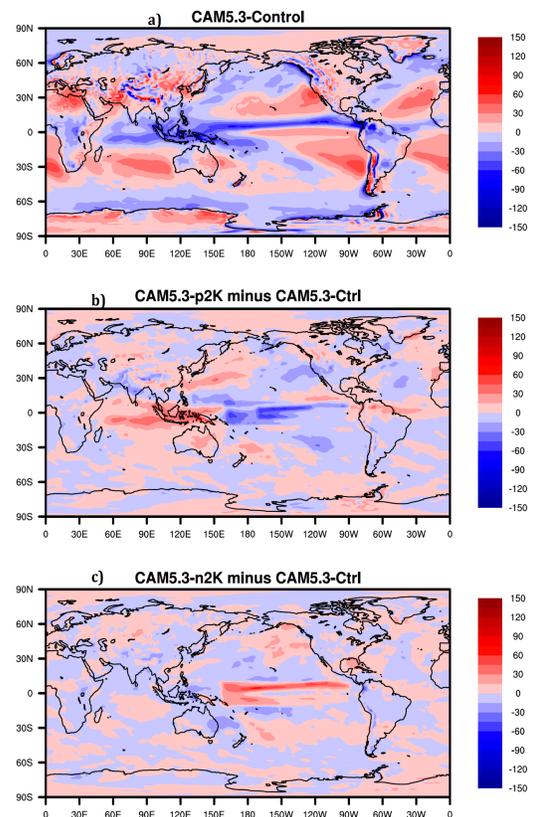


Fig 14. a) 500hPa vertical velocity in the CAM5 control run; b) Anomalies of 500hPa vertical velocity in the CAM5+2K run with respect to the control run; c) Anomalies of 500hPa vertical velocity in the CAM5-2K run with respect to the control run. Unit of 500hPa vertical velocity is hPa/day.

## 5 Conclusions

In this study, we use statistical methods, namely the VAR method for Granger Causality model, and global climate model simulations to investigate ENSO causality as one of the modulating climate factors that cause the anomalies in surface air temperature, precipitation, surface pressure and vertical wind over remote regions through teleconnection with lagged temporal variability. We analyzed different observational data, reanalysis data and model data to comprehensively investigate the global impacts of ENSO. The Granger causality analysis was 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. Besides, all the source codes used in this study can be found on Github [23].

For future work, we will compare the granger causality on multiple variables for intercomparison, and we will add analysis on the impact of ENSO on clouds and aerosol using NASA satellite remote sensing data. We plan to use the Granger causality methods to predict climate variations and other interesting economic factors, such as crop yield and wheat stock price, under different ENSO backgrounds. At the same time, we plan to explore more efficient use of high performance parallel computing in the studies that use much broader big data of satellite observations with higher spatial and temporal resolutions.

## Acknowledgments

This work is 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). The hardware in the UMBC High Performance Computing Facility (HPCF) is supported by the U.S. National Science Foundation through the MRI program (grant nos. CNS-0821258, CNS-1228778, and OAC-1726023) and the SCREMS program (grant no. DMS-0821311), with additional substantial support from the University of Maryland, Baltimore County (UMBC). See [hpcf.umbc.edu](http://hpcf.umbc.edu) for more information on HPCF and the projects using its resources.

## References

- [1] I. Mohkov, D. Smirnow, P. Nakonechny, S. Kozlenko, E. Selez, and J. Kurths, “Alternating mutual influence of El-Niño/Southern Oscillation and Indian monsoon”, *Geophys. Res. Lett.*, 38, 2011.
- [2] 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.
- [3] A. Kumar, L. Zhang and W. Wang, “Sea Surface Temperature - Precipitation Relationship in Different Reanalyses”, *Monthly Weather Review*, 1118-1123, 2012.
- [4] Y. Chen, and coauthors, “Forecasting fire season severity in south America using sea surface temperature anomalies”, *Science*, Vol 334., 787-791, 2011.
- [5] 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.
- [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] A. C. Lozano, A. Naoki, Y. Liu, and S. Rosset, “Grouped graphical Granger modeling methods for temporal causal modeling.” In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 577–86. KDD '09. New York, NY, USA: ACM. 2009.
- [11] M. McGraw, and E. A. Barnes, “Memory matters: A case for Granger causality in climate variability studies”, *Journal of Climate*, 3289-3300, 31, 2018.
- [12] J. Elsner, “Granger causality and Atlantic hurricanes”, *Tellus*, 59A, 476–485, 2007.
- [13] I. Mohkov, D. Smirnow, P. Nakonechny, S. Kozlenko, E. Selez, and J. Kurths, “Alternating mutual influence of El-Niño/Southern Oscillation and Indian monsoon”, *Geophys. Res. Lett.*, 38, 2011.
- [14] 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.
- [15] E. Kalnay, and Co-authors, “The NCEP/NCAR 40-year reanalysis project”, *Bull. Amer. Meteor. Soc.*, 77, 437-470, 1996.
- [16] R. Adler, and Co-authors, “New global precipitation climatology project monthly analysis product corrects satellite data drifts”, *GEWEX News*, 26 (4), 7-9, 2016.

- [17] 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.
- [18] R. E. Kass and L. Wasserman, “A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion”, *Journal of American Statistical Association*, 92 (437), 1995.
- [19] 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.
- [20] High Performance Computing Facility, UMBC: <http://hpcf.umbc.edu/>
- [21] S. Bony, J.-L. Dufresne, H. LeTreut, J.-J. Morcrette, and C. Senior, “On dynamic and thermodynamic components of cloud changes”, *Climate Dyn.*, 22, 71–86, 2004.
- [22] S.-W. Yeh, W. Cai, S.-K. Min, M.J. McPhaden, D. Dommenges, B. Dewitte, and J.-S. Kug, “ENSO atmospheric teleconnections and their response to greenhouse gas forcing”. *Reviews of Geophysics*, 56. <https://doi.org/10.1002/2017RG000568>.
- [23] H. Song, J. Tian, J. Huang, J. Wang, and Z. Zhang, <https://github.com/big-data-lab-umbc/cybertraining/tree/master/year-1-projects/team-4>, Source code, 2018.