



## **Dimensionality reduction and network inference for sea surface temperature data**

Fabrizio Falasca (1), Annalisa Bracco (1), Athanasios Nenes (1), Constantine Dovololis (2), and Ilias Fountalis (2)

(1) Georgia Institute of Technology, School of Science, Earth and Atmospheric Sciences, Atlanta, United States

(fabrifalasca@gmail.com), (2) Georgia Institute of Technology, College of Computing, Atlanta, United States

Earth's climate is a complex dynamical system. The underlying components of the system interact with each other (in a linear or non linear way) on several spatial and time scales. Network science provides a set of tools to study the structure and dynamics of such systems. Here we propose an application of a novel network inference method,  $\delta$ -MAPS, to investigate sea surface temperature (SST) fields in reanalyses and models.  $\delta$ -MAPS first identifies the underlying components (domains) of the system, modeling them as spatially contiguous, potentially overlapping regions of highly correlated temporal activity, and then infers the weighted and potentially lagged interactions between them. The SST network is represented as a weighted and directed graph. Edge direction captures the temporal ordering of events, while edge weights capture the magnitude of the interaction between the domains.

We focus on two reanalysis datasets (HadISST and COBE ) and on a dozen of runs of the CESM model (extracted from the so-called large ensemble). The networks are built using 45 years of data every 3 years for the total dataset temporal coverage (from 1871 to 2015 for HadISST, from 1891 to 2015 for COBE and from 1920 to 2100 for CESM members).

We then explore similarities and differences between reanalyses and models in terms of the domains identified, the networks inferred and their time evolution.

The spatial extent and shape of the identified domains is consistent between observations and models.

According to our analysis the largest SST domain always corresponds to the El Niño Southern Oscillation (ENSO) while most of the other domains correspond to known climate modes. However, the network structure shows significant differences.

For example, the unique role played by the South Tropical Atlantic in the observed network is not captured by any model run.

Regarding the time evolution of the system we focus on the strength of ENSO: while we observe a positive trend for observations and model members in the second half of the 20th century and the first decade of the 21st century, the trajectory of different ensemble members differs significantly in the projections.

Fully coupled climate models represent a primary tool for predicting the evolution of the climate system under increasing anthropogenic emissions: our findings help identifying important limitations in the models' ability to reproduce historical climate records and therefore predict future climate.