**Motivation**

The CLDERA Grand Challenge Project is identify causal pathways between various climate features and spatial components.

**Goal:** Generate a weighted directed graph where the edges indicate influence from one feature of interest to another. Chaining these edges together defines a pathway which we call a Source to Impact pathway.

Source to Impact pathways can help analysts identify which features of interest have the largest influence on their results

- Which sources are the largest contributors to Arctic Sea Ice extent?
- What are the potential downstream effects of a drought in Western Europe?
- Find the path of least resistance from a source to impact?
- What are the bottleneck features/nodes between source and impact?

**Results Summary**

**Example:** Synthetic Dataset

- Set of coupled equations
- Dependencies are 1 timestep, the algorithm will look at 1 and 2 timesteps

\[ W_t = 0.9W_{t-1} + \varepsilon_t \]

\[ X_t = 0.8X_{t-1} + 0.5W_{t-1} + \varepsilon_t \]

\[ Y_t = -0.9W_{t-1} + \varepsilon_t \]

\[ Z_t = 0.3X_{t-1} + 0.5Y_{t-1} + \varepsilon_t \]

**Feature Importances**

- Importance Weights calculated using SHAP feature importance
- Weights are pruned using a random variable as a cutoff point

**Technical Approach**

**Random Forest Regressors (RFRs):** Machine Learning (ML) predictive models

- Require training data comprised of pairs of inputs and outputs
- Once trained, can predict outputs given a set of input variables

**Feature Importances:** a scalar value that indicates the predictive power that a particular input variable has on a given output variable

RFR is commonly used for and well-developed for regression, classification and prediction tasks.

- Our approach extends RFR for purpose of discovering pathways

**Approach:**

- Train a multi-variate RFR (python package: sklearn)
- Determine feature importances weights between inputs and outputs
- Convert weights into a weighted directed graph

**Impacts & Successes to date**

**Contributions to CLDERA**

- Developing a suite of verification metrics for feature pathway graphs
- Used as tool for explaining model outputs from other thrusts
- Identify key pathway differences between simulation runs

**Presentations / Papers**

- Poster at American Geophysical Union (AGU) 2022
- In Progress of writing a journal paper

**Signature-based Clusters**

We will be using these types of clusters as inputs to identify these moving clustering influence one another

**0.5 FTE for Subthrust in CLDERA Grand Challenge**