SAND2022-7561C

Sandia National Laboratories

CLDERA CLimate impact: Determining Etiology thRough pAthways

A Novel Foundational Approach for Attributing Climate Impacts

Diana Bull, Kara Peterson, Irina Tezaur, Lyndsay Shand, Laura Swiler
Sandia National Laboratories

8th European Seminar on Computing (ESCO)
June 14, 2022

U.S. DEPARTMENT OF ENERGY
NNSA National Nuclear Security Administration
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Tools to identify source-impact relationships are needed

The ability to distinguish the impacts of intervention, anthropogenic climate change, and natural variability will become increasingly important as the effects of climate change compound.

Earth System Models (ESMs) are the primary testbed to predict climate impacts, but complex coupling of processes in ESMs can obscure the relationships between sources and impacts.

Observations have been historically underutilized to improve our knowledge of the earth system.

Attributing a predominant source to an impact is an ill-conditioned problem due to many possible sources leading to similar climate impacts.

Modified from: http://www.thesourgrapevine.com/2019/11/the-ball-of-string-theory-for-learning.html

These limitations create a **significant barrier** to developing **quantitative relationships** between **sources and impacts.**

Establishing Connective Relationships in Earth's Climate: ***Current Approaches***

Confounding characteristics of the climate attribution problem:

- High internal variability in the Earth system;
- Limited ESM ensembles;
- Historically limited observational data;
- Multiple sources contributing to an impact.

Techniques currently applied to climate attribution:

- Fingerprinting (e.g. Hasselmann, 1997; Hegerl & North, 1997; Marvel et al., 2020)
- Causal Inference (e.g. Runge et al., 2019; Nowack et al., 2020)
- Emergent constraints (e.g. Hall et al., 2019; Williamson et al., 2021)

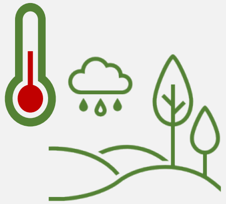
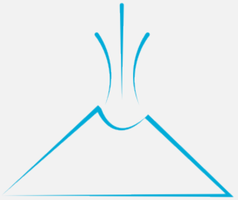
"Pecking Order", Edgar Hunt

Connective relationships are often unbounded from etiological relationships allowing physically meaningless source-impact correlations.

(Caldwell et al., 2014)

Moving beyond a correlative approach – establishing connective relationships.

Attribution in Climate

				**CLDERA** 
What is the question?	How do rising concentrations of GHGs in the atmosphere affect climate state variables?	What are the relative contributions of different sectors, activities, and entities to concentrations of GHGs in the atmosphere?	How do changes in the global climate system affect the relative frequency and severity of extreme events ?	What are the downstream impacts from large, spatio-temporally localized sources within the climate?
What is the source?	Long-term, slowly accumulating GHG emissions	Sector-specific emissions	Changes to global climate system	Geographically and temporally localized source forcings in the climate
Examples	WG1 and WG2 IPCC reports	Cement industry, transportation	Likelihood of Hurricane Katrina	Volcanic eruptions, wildfires
Is the attribution direct / quantitative?	WG1- yes (temp/precip/…) WG2 – normally no, notional linkage	yes	no, notional connection	yes, linked through multiple process-nodes

The goal of CLDERA is to develop new tools to enable *downstream impact attribution* from geographically and temporally localized source forcings in the climate.

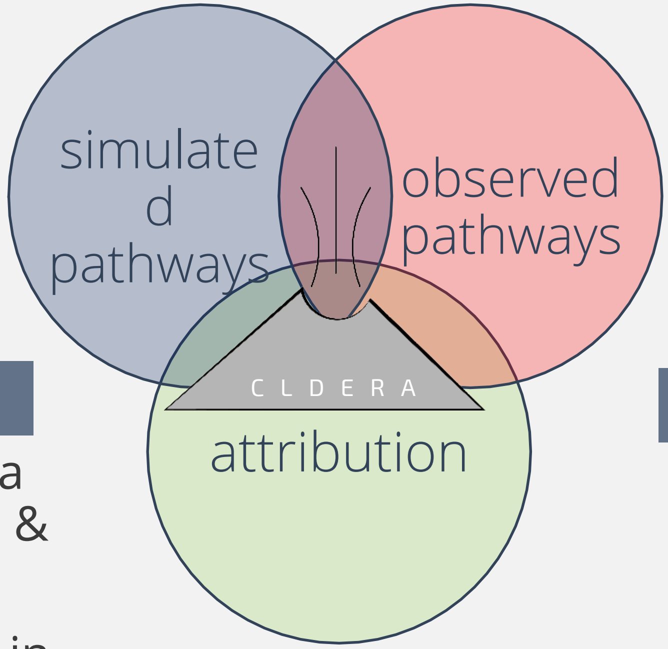
CLDERA
a novel foundational approach

Need

Determine how a geographically and temporally localized source drives the climate system to respond with particular impacts to enable downstream impact attribution

Technical Approach

- Build upon key strengths of Sandia (modeling & simulation, detection & attribution, risk analysis)
- Develop novel methods and tools in 3 cross-validating thrusts
- 1991 explosion of Mt. Pinatubo as exemplar



Hypothesis

Tracing pathways between source and impacts will increase certainty of attribution and deepen understanding of dependent causal-like relationships

Pathways represent the spatio-temporally evolving chain of physical processes that connects a source to mpacts.

Outcome

Advance climate attribution science by identifying impacts from localized sources.

- Formalize pathways and establish robustness
- Ranking of sources to impact
- Attribution of source characteristics

Mt. Pinatubo
Stratospheric Aerosol Injection

A strong working example:

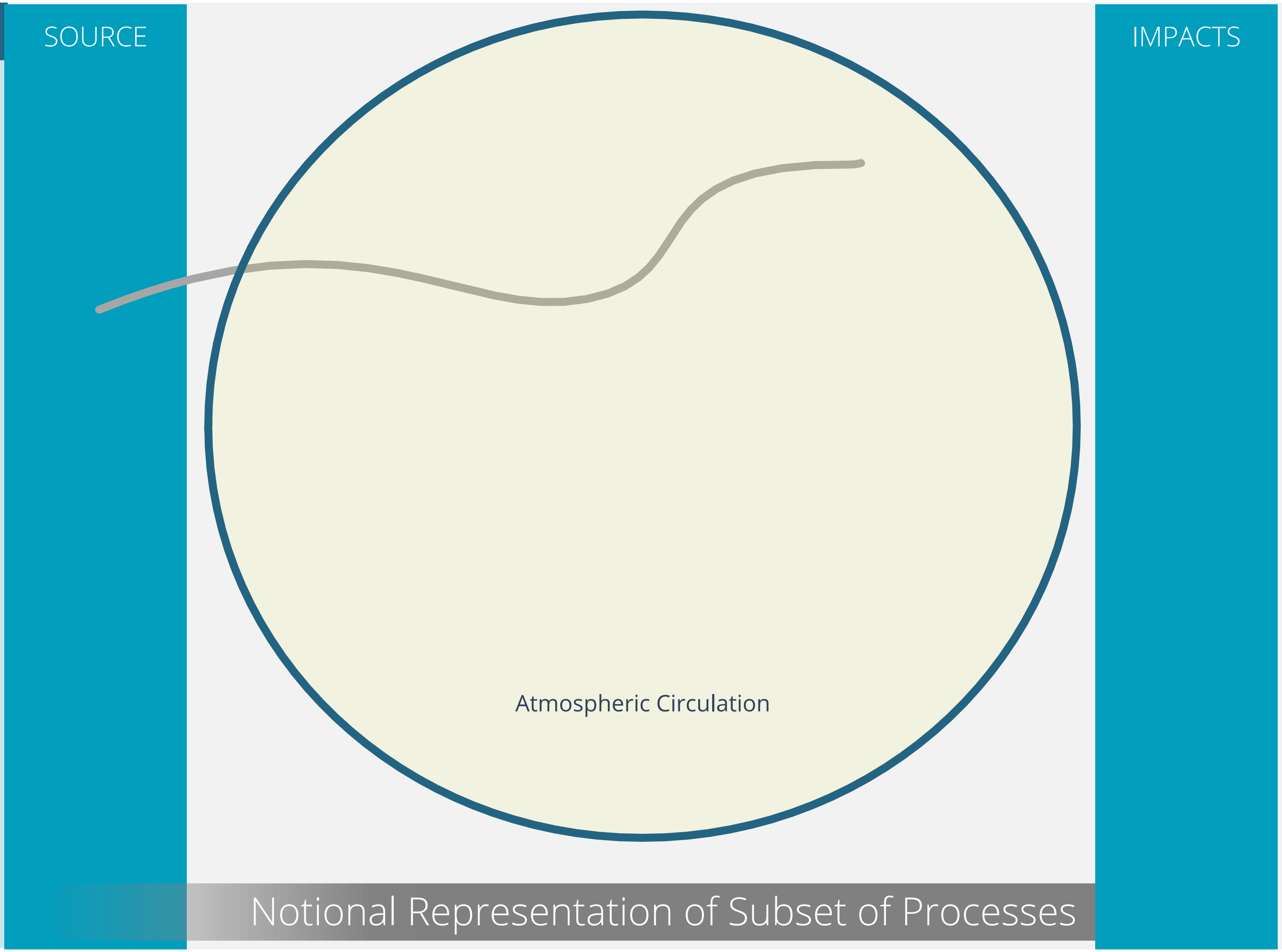
The source is an **analog for a climate intervention technique**

- 1 Tg SO2/yr may produce on the order of 0.1°C cooling (Kravitz et al., 2017)

The source **forcing is external to the feedbacks** within the Earth's climate

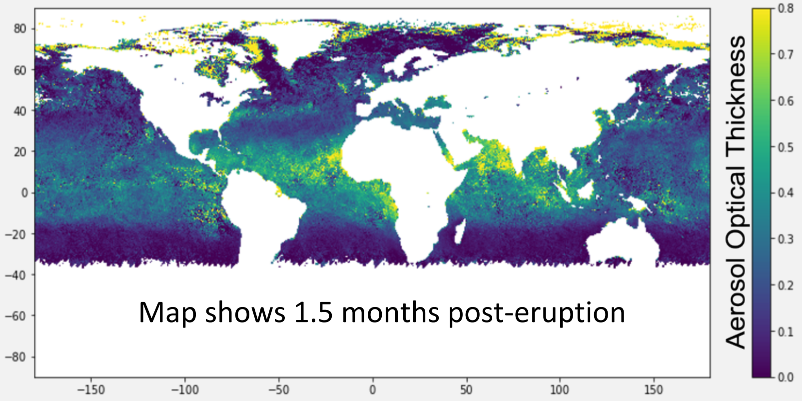
The **impacts are large enough to rise above internal variability** in the simulations

It provides **ample observational data** and, because the impacts and pathways are relatively well characterized it **offers validation of our analysis techniques within the approach**

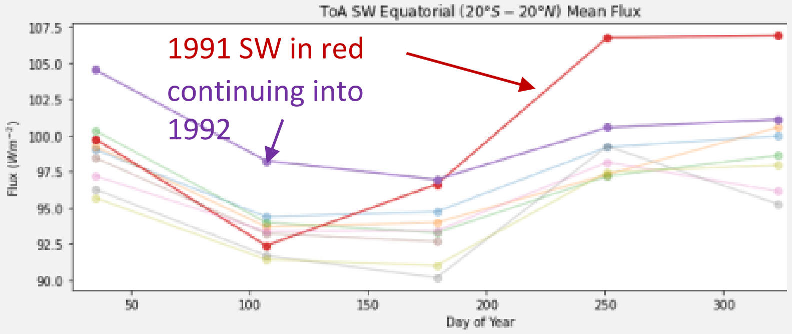


Temperature Impacts from Mt. Pinatubo

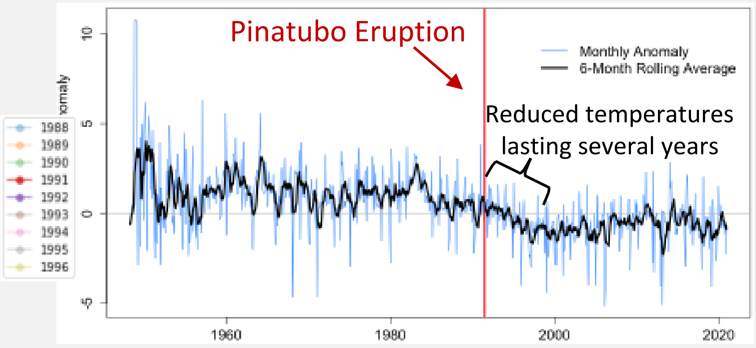
Emission/Aerosol formation
Aerosol Optical depth



Change in Radiation Flux
Net shortwave & longwave



Temperature Change
2m and 50hPa



	Stratosphere	Troposphere	NH Winter Warming
MECHANISM	• Absorption of terrestrial LW and incident solar near-IR radiation by sulfate aerosols • Confounding factors: ozone changes	• Reduction in solar radiation at surface due to scattering by sulfate aerosols • Water vapor feedback [Soden et al., 2002] • Confounding factors: ENSO, LW radiation, aerosol indirect effects, climate dynamics	• Tropical stratospheric warming intensifies polar vortex and induces a strong positive phase in the NAO • Disputed: natural variability or forced by eruption
IMPACT	• Observed ~3 C warming at 50-hPa, [Angell 1997]	• Observed 0.4 C global mean cooling [McCormick et al., 1995]	• Warming in NH first winter after tropical eruption [Robock & Mao, 1992]

Energy Exascale Earth System Model (E3SM)

- US DOE funded Earth system model.
- Collaboration among 8 National Labs and 12 academic institutions.
- Designed to run on DOE's advanced computing platforms and address energy-relevant science questions.

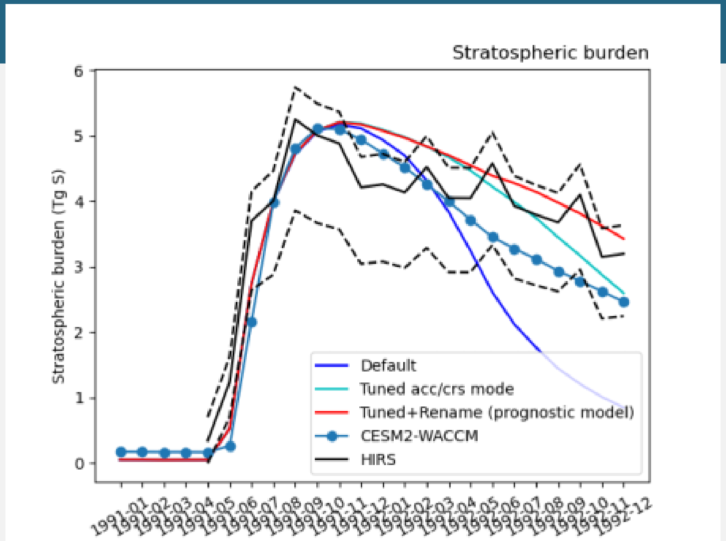
- Capabilities to model Mt. Pinatubo eruption and impacts.
  - Resolves stratosphere: 72 vertical levels with 0.1hPa (64 km) model top.
  - Includes key process models for clouds, ozone, and aerosols [Golaz, et al. 2019].

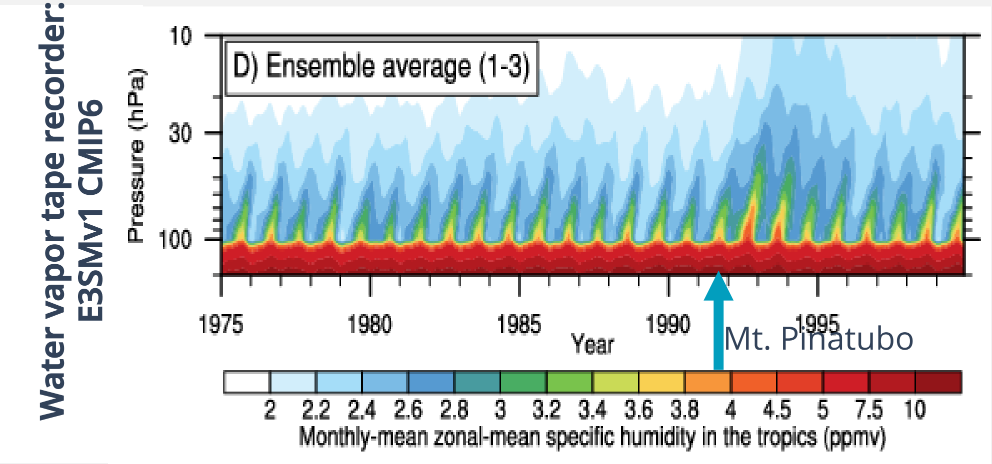
www.e3sm.org

Simulating Mt. Pinatubo

- Implemented prognostic volcanic aerosols in E3SMv2
  - Change aerosol coarse and accumulation mode properties to match stratospheric size distributions
  - Evaluate OH consumption on sulfate production

- Evaluating E3SM simulation capabilities
  - Assess Quasi-Biennial Oscillation (QBO) and realism of E3SM's stratospheric circulation
  - Evaluate E3SM's Brewer Dobson circulation via age-of-air tracers and water vapor tape recorder





X. Liu- Impacts of Model Resolution on Simulated Aerosol and Aerosol Effects on Climate Over East Asia With NCAR CESM, Wed. 6-8 pm.

Team: B. Wagman, H. Brown (SNL), X. Liu, A. Hu (TAMU), C. Jablonowski, J. Hollowed (U Michigan)

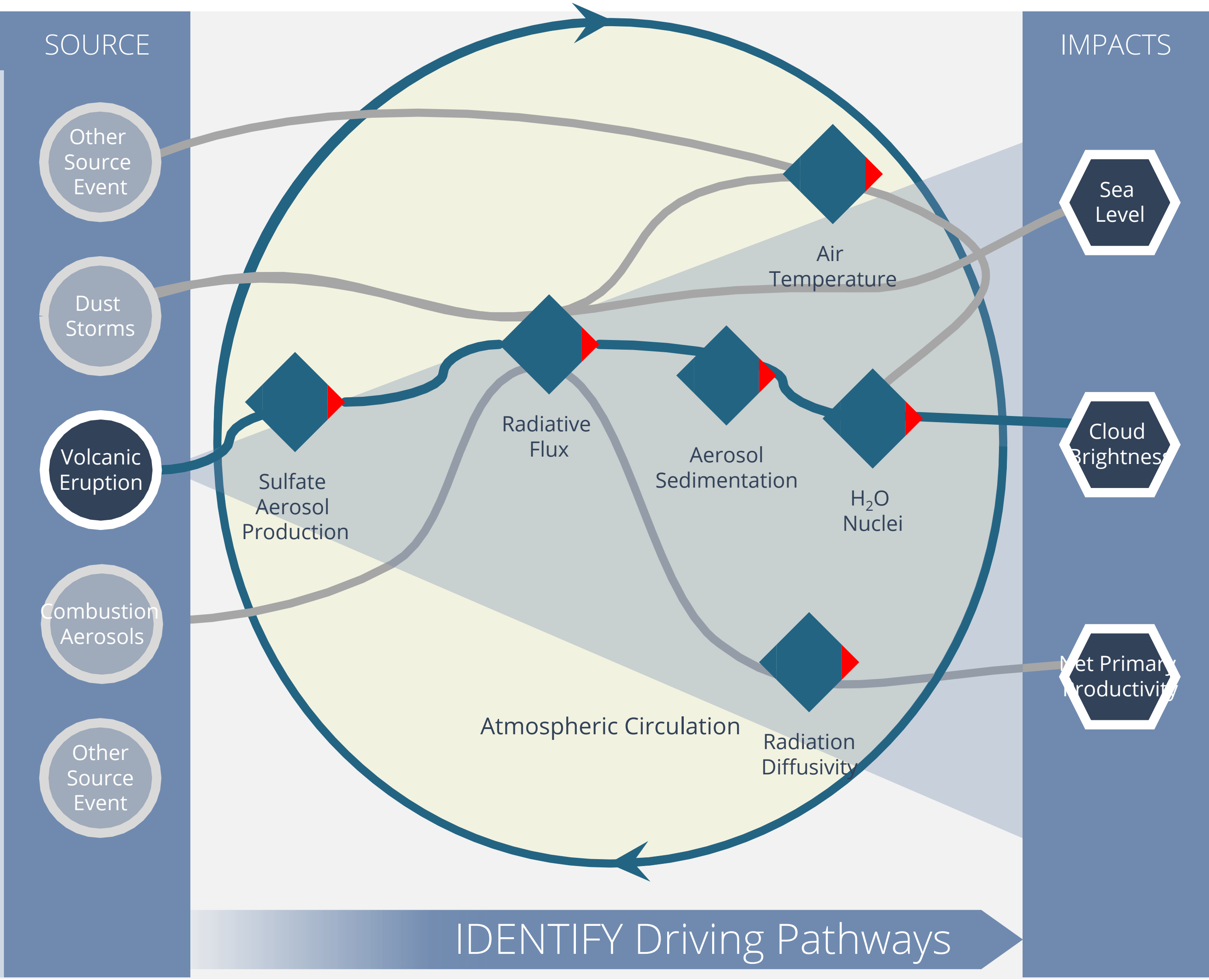
SIMULATED Pathways
finding impacts from a source

Research Composition:

Statistical Approaches: use and extend **sensitivity analyses** and **Random Forest Regression** to identify and rank physical pathways while establishing susceptibility to initial conditions and E3SM representations

Computational Monitoring: instrument E3SM with **tracers and profiling** capabilities to enable pathway detection; additionally deploy **novel simulation strategies** to elucidate Mt. Pinatubo impacts on climate

Anomaly Detection: define pathways through extensions to existing anomaly detection algorithms in order to **spatio-temporal**ly trace **anomaly** progression

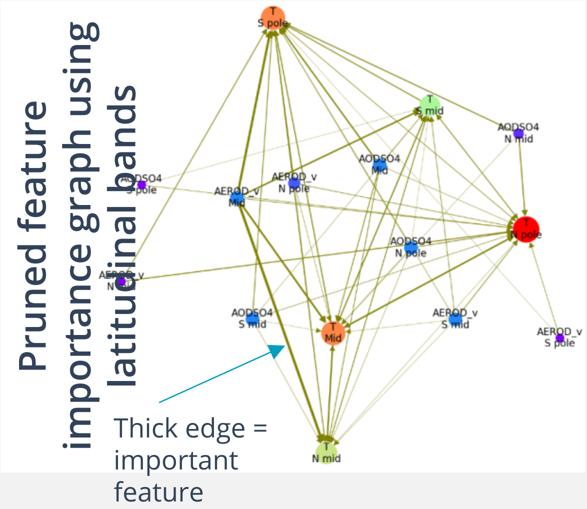


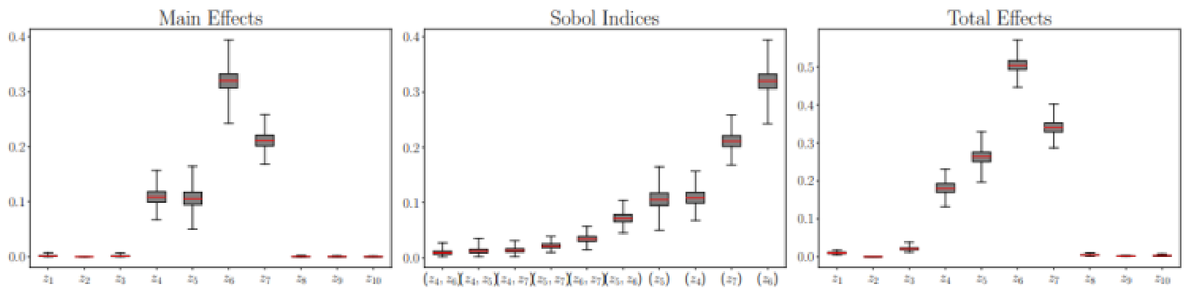
Statistical Methods

- **Random Forest Regression**
  - Use feature importances from Random Forests (RFs) to identify weighted pathways from one variable (or feature) to another

- **Sensitivity Analyses**
  - Determine robustness of pathways to changes in SAI-/model-related inputs
  - Two kinds of sensitivity studies using different resolutions of E3SM:

1) Vary eruption elements, e.g., mass, time, height, location (how Mt. Pinatubo could have been – "how low can you go")
2) Vary MAM4 aerosol and MG2 cloud microphysics model parameters





Global sensitivity analysis (GSA) results for surface temperature QOI from [Tezaur et al., 2021]

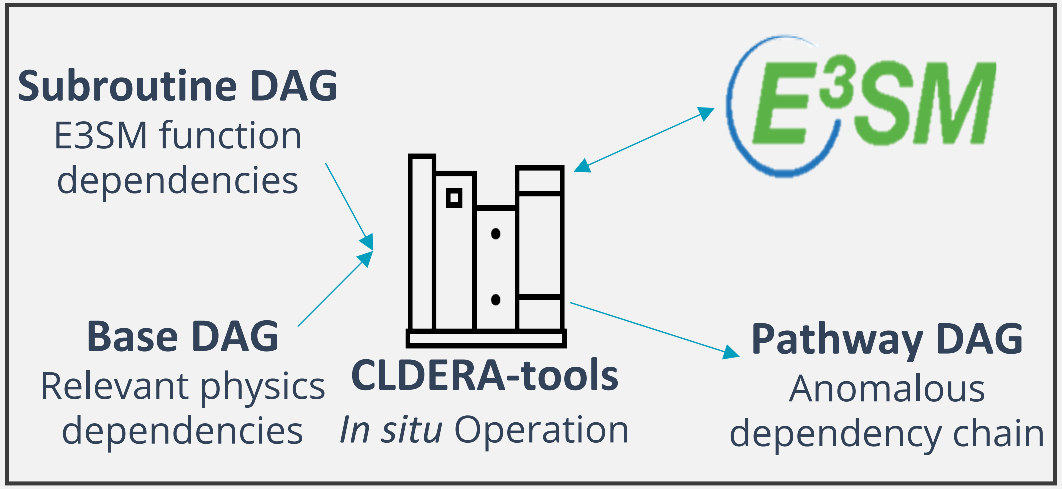
I. Tezaur, Global Sensitivity Analysis Using the Ultra-low Resolution Energy Exascale Earth System Model, Tues. 2:30-4:30 pm.

Team: M. Peterson, I. Tezaur, K. Peterson

Computational Monitoring

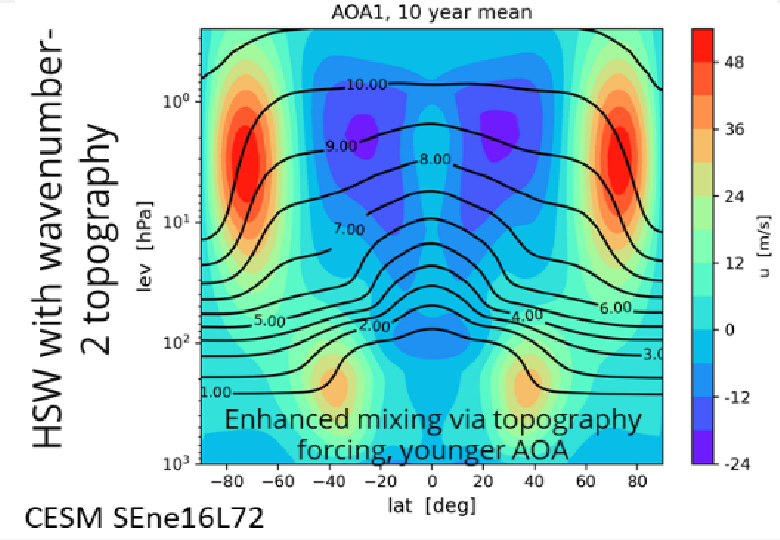
- **Profiling – in situ pathways**
  - Instrument E3SM and develop Directed Acyclic Graphs (DAGs) of process interactions executed within code

A. Steyer- Detecting Physical Pathways With Software Profiling, Tues. 2:30-4:30 pm.



Team A. Steyer, J. Watkins, L. Bertagna, G. Harper, I. Tezaur (SNL)

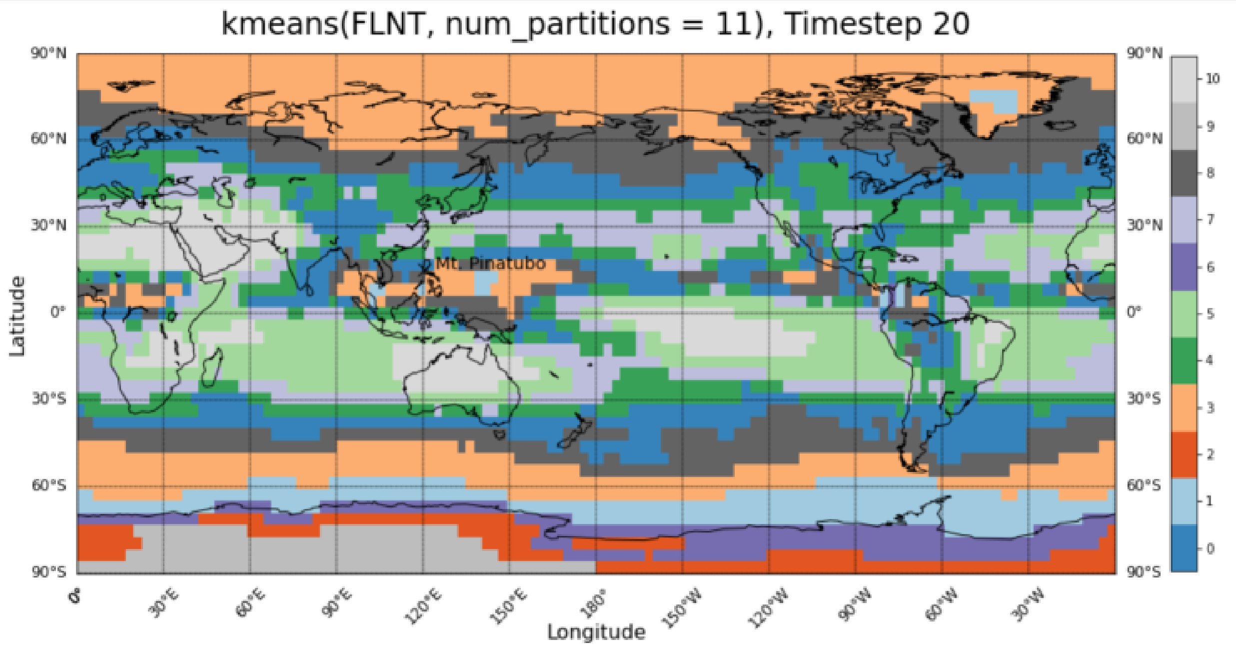
- **Tracers**
  - *Passive*: Evaluate climatological dynamics (like "Age of Air") and changes to those dynamics.
  - *Active*: enable tracking of relative contributions along nodes in pathway.



Team B. Hillman, A. Steyer, I. Tezaur (SNL), C. Jablonowsi, J. Hollowed (U. Michigan)

Anomaly Detection

- Extend existing anomaly (event) detection algorithms [Davis et al., 2021] to enable spatio-temporal anomaly progression, defining pathways
  - Signature-measure-decision-based approach
  - Communication-minimizing
  - Unsupervised (extensions to supervised possible)
  - Leverages “In-Situ Machine Learning for Intelligent Data Capture on Exascale Platforms” (ISML) project (PI: W. Davis)



W. Davis - In Situ Machine Learning for Intelligent Data Capture in HPC Simulations, Wed. 6:00-8:00 pm.

Signature Clustering Approach:

- use compressed ISML signatures for clustering
- use clusters to inform pathway analyses

Team: W. Davis, M. Carlson, I. Tezaur (SNL)

OBSERVED Pathways
finding impacts from a source

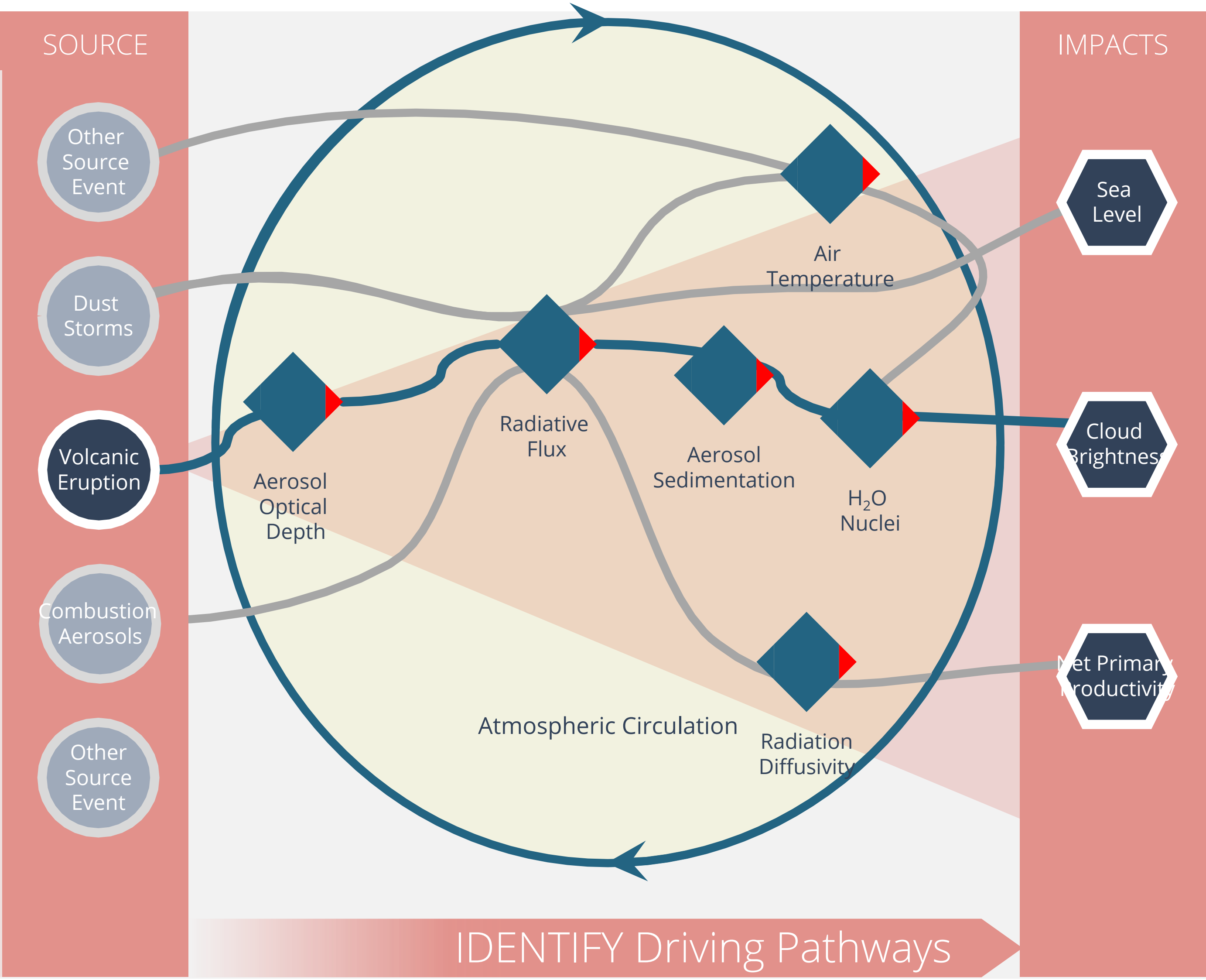
Research Composition:

Data Fusion: strategically source and fuse relevant data of varying resolutions and fidelities to create a **"near-global" picture** of the relevant processes

Change Point: identify the underlying fundamental **shifts** in **climate processes**

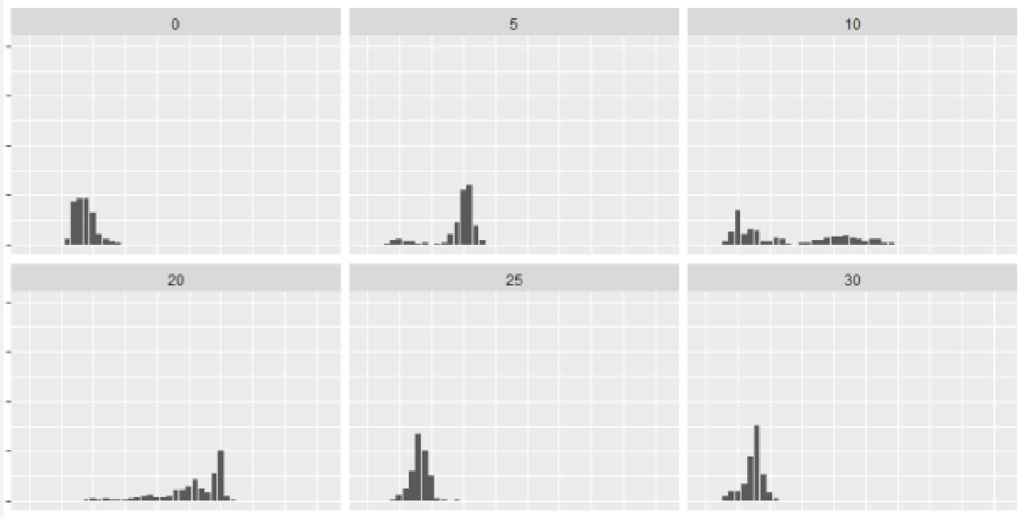
Space-Time Statistical Methods: adapt **Bayesian hierarchical approaches** to represent process dependencies and their dynamic spatio-temporal evolution over the first 3-4months post eruption

Hybrid Statistical & Deep Learning Methods: develop hierarchical statistical approaches **embed**ded with **DL** techniques to trace secondary and tertiary temporally-lagged effects



Change Point Methods

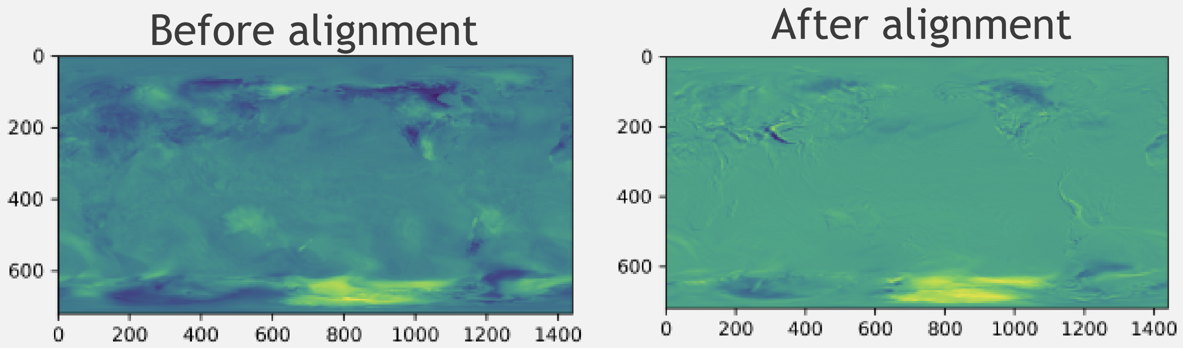
- **Space-time methods** stablish climate shifts and relationship to distance from source.

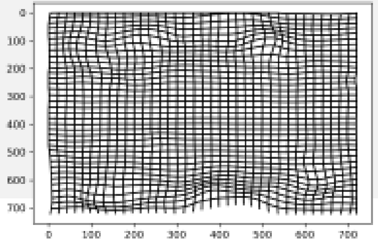


Posterior distribution of estimated spatially-varying changepoint at different latitudes (using simulated data)

- **Elastic Functional Warping Methods** can be leveraged to identify change points in the underlying spatio-temporal process by isolating shifts that are not due to expected weather patterns.

Difference between time 0 to time N





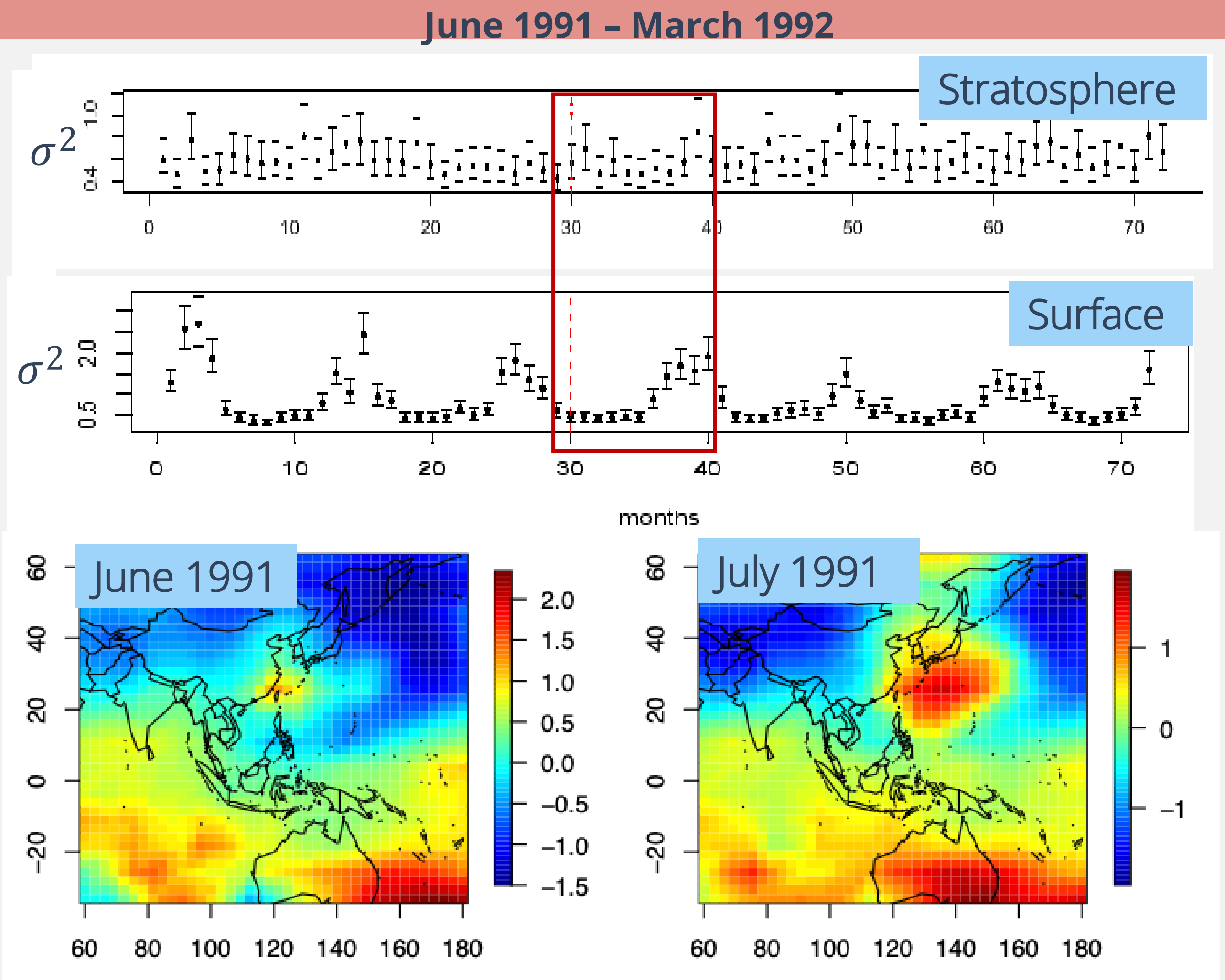
The diffeomorphism gives a measure of the bending/ stretching to align spatial data at time 0 and N.

Team: B. Li, S. Jun (U Illinois), D. Tucker, L. Shand (SNL)

Spatio-Temporal Statistics

- Dynamic spatio-temporal models are intuitive and interpretable ways to represent complex, dependent, physical processes that change across space and time simultaneously.

- Implementation of Finley et al. (2012) to reanalysis temperatures at multiple pressure levels reveal evidence of Mt Pinatubo impact.



Time-varying variance parameter estimates with red line representing time of event (above) and estimated spatial dynamic components (below) of the dynamic model applied to stratospheric temps.

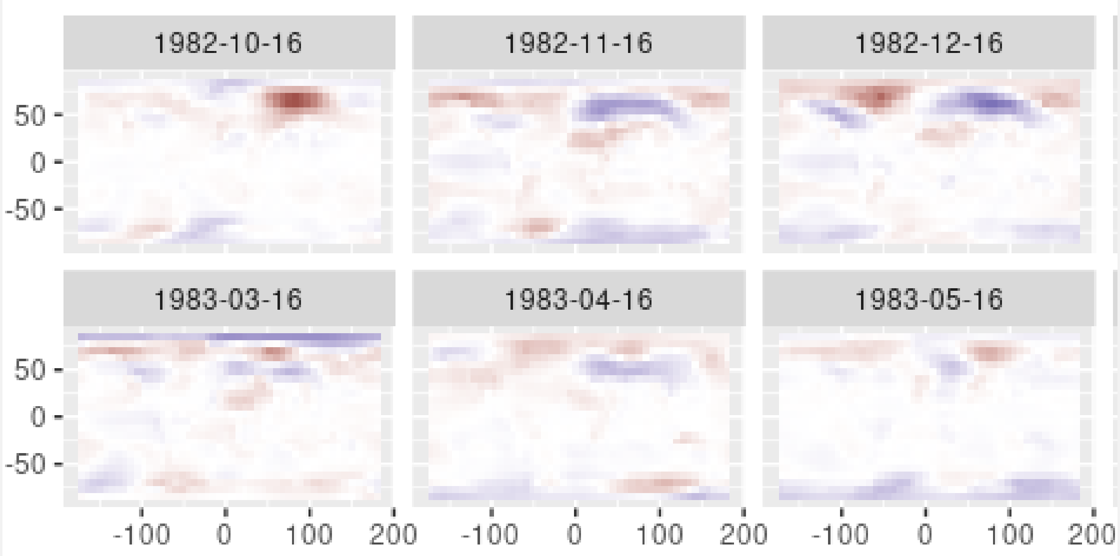
Team: K. Goode, B. Hillman, G. Huerta, J. Li, L. Shand (SNL)

Hybrid Stat-DL Methods

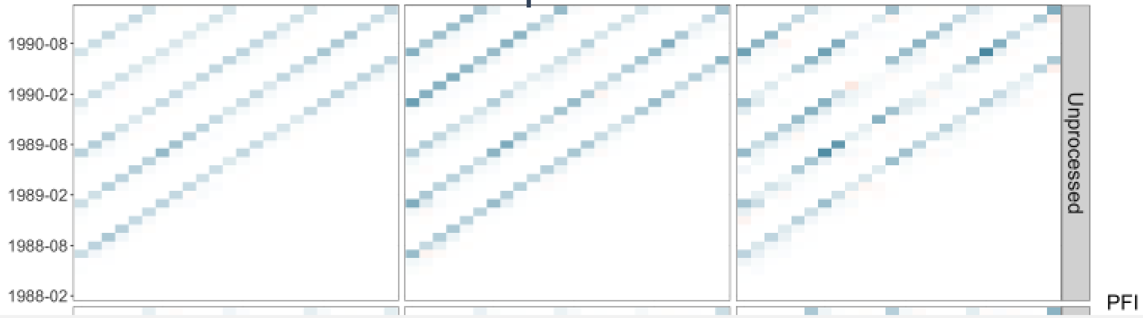
- **Echo-State Networks (ESNs)**
  - Use reservoir computing to efficiently estimate recurrent neural networks (RNN).
  - Represent dependent processes occurring over different time scales by increasing interpretability through introducing spatially dependent parameters.
  - McDermott and Wikle [2018]

- **Permutation Feature Importance**
  - Novel method to increase explainability of existing ESNs.
  - Leverages Goode et al. [2020].

Quadratic ESN Residuals – Need for increased spatial fidelity to correctly account for high latitude temperatures



PFI – identifying how ESN's make use of historical temperatures by investigating the influence of lagged temperature on forecast temperatures



Team: K. Goode, B. Hillman, G. Huerta, J. Li, L. Shand (SNL)

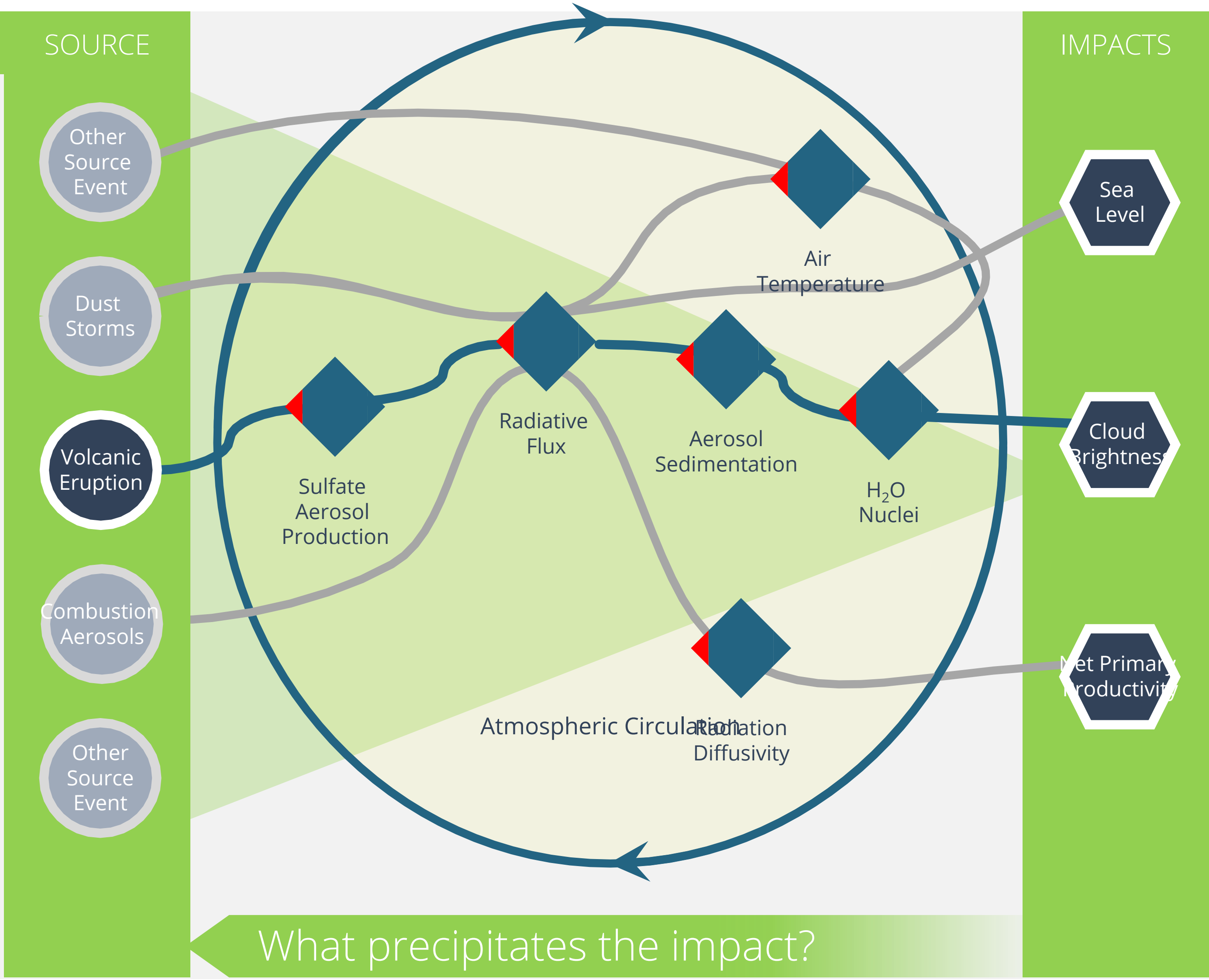
ATTRIBUTION
finding predominant source driving impact

Research Composition

Enhanced Fingerprinting: use pathway information to expand **multi-variate analyses**, sharpening the signal-to -noise ratios and enabling significant correlations between source-impact

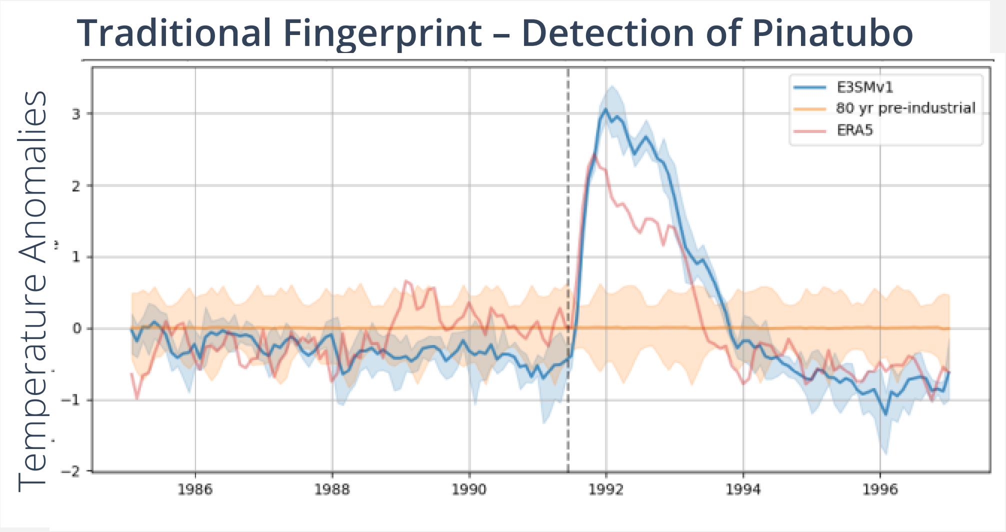
Inverse Optimization: develop deep operator neural networks (DONNs) to model parts of E3SM to enable **PDE-constrained optimization** without intrusion into the E3SM code directly; pathways will act as penalty terms or constraints

Causal Modeling: identify causal relationships and dominant pathways by developing **causal networks** through iterative independence tests and the resulting directed acyclic graphs



Enhanced Fingerprinting

- A fingerprint is a spatial and/or temporal pattern that highlights an impact
- Commonly, a fingerprint is the first principal component or empirical orthogonal function (EOF) from a singular value decomposition of a data matrix.
- Detection involves identifying if the signal projected back onto the EOF goes out of 2σ limits.
- Attribution involves examining the magnitude of the fingerprint in a regression formulation and checking its significance.



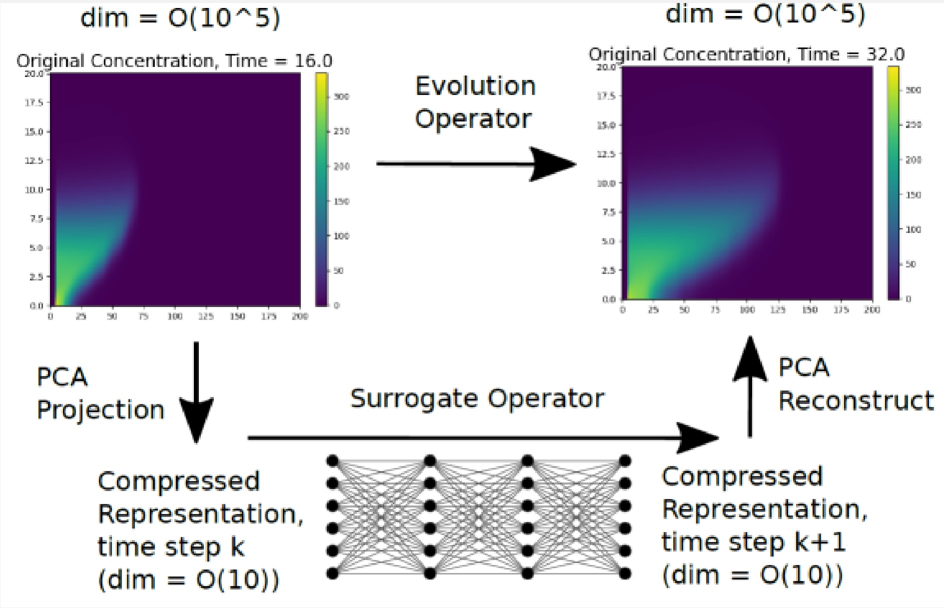
Leverages Wagman et al. 2021

K. Chowdhary, E3SM Atmosphere Surrogate Construction Using Data-driven Reduced Order Modeling, Wed. 6:00-8:00

Team: B. Wagman, K. Chowdhary, L. Swiler (SNL); K. Marvel (Columbia)

Inverse Optimization

- The identification of magnitude, height, and location of the source is a large-scale, PDE constrained inverse problem.
    - Find source parameters θ which minimize the difference between aerosol concentrations (c_{sim}) and observations (c_{obs}) subject to the atmospheric physics constraints
- **Core idea**: develop deep operator neural networks (DONNs) to model parts of E3SM to enable PDE-constrained optimization without intrusion into the E3SM code directly.
- Add pathway information as penalty terms in the objective and/or additional constraints.



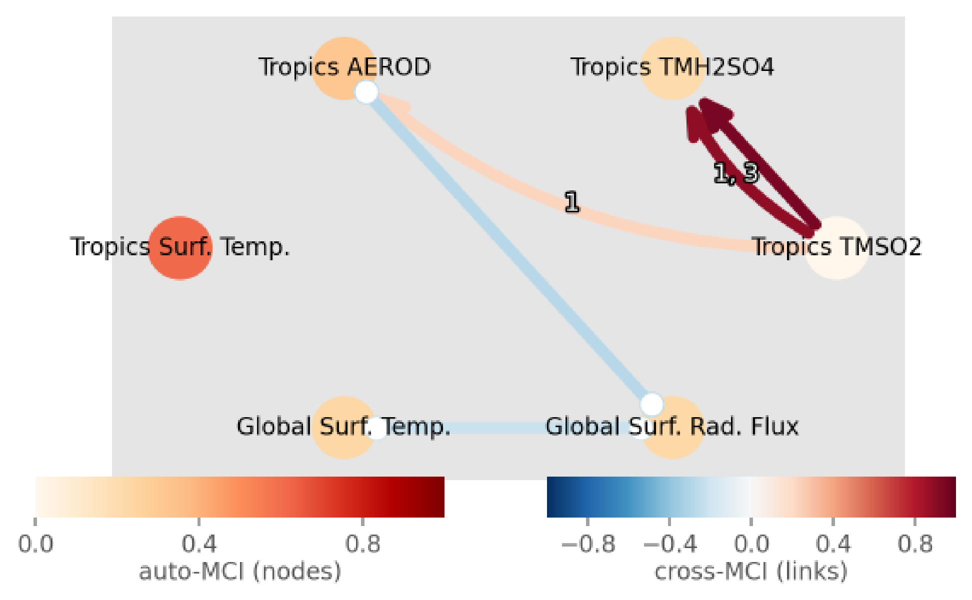
Advection-diffusion reaction with Gaussian plume source

$$\frac{\partial c}{\partial t} - \kappa \nabla^2 c + \boldsymbol{v} \cdot \nabla c - S\boldsymbol{e}_y \cdot \nabla c = R(c) + f$$

M. Gulian Facilitating Atmospheric Source Inversion via Deep Operator Network Surrogates, Wed. 10:30-12:30

Causal Modeling

- Establish relationships between a set of variables to confirm a hypothesized casual relationship
- Causal network learning
  - Fits a graphical model of covariates by iterative conditional independence tests.
  - PCMCI algorithm: Peter Clark Momentary Conditional Independence. [Runge, 2019]
  - Allows for and detects contemporaneous dependencies.

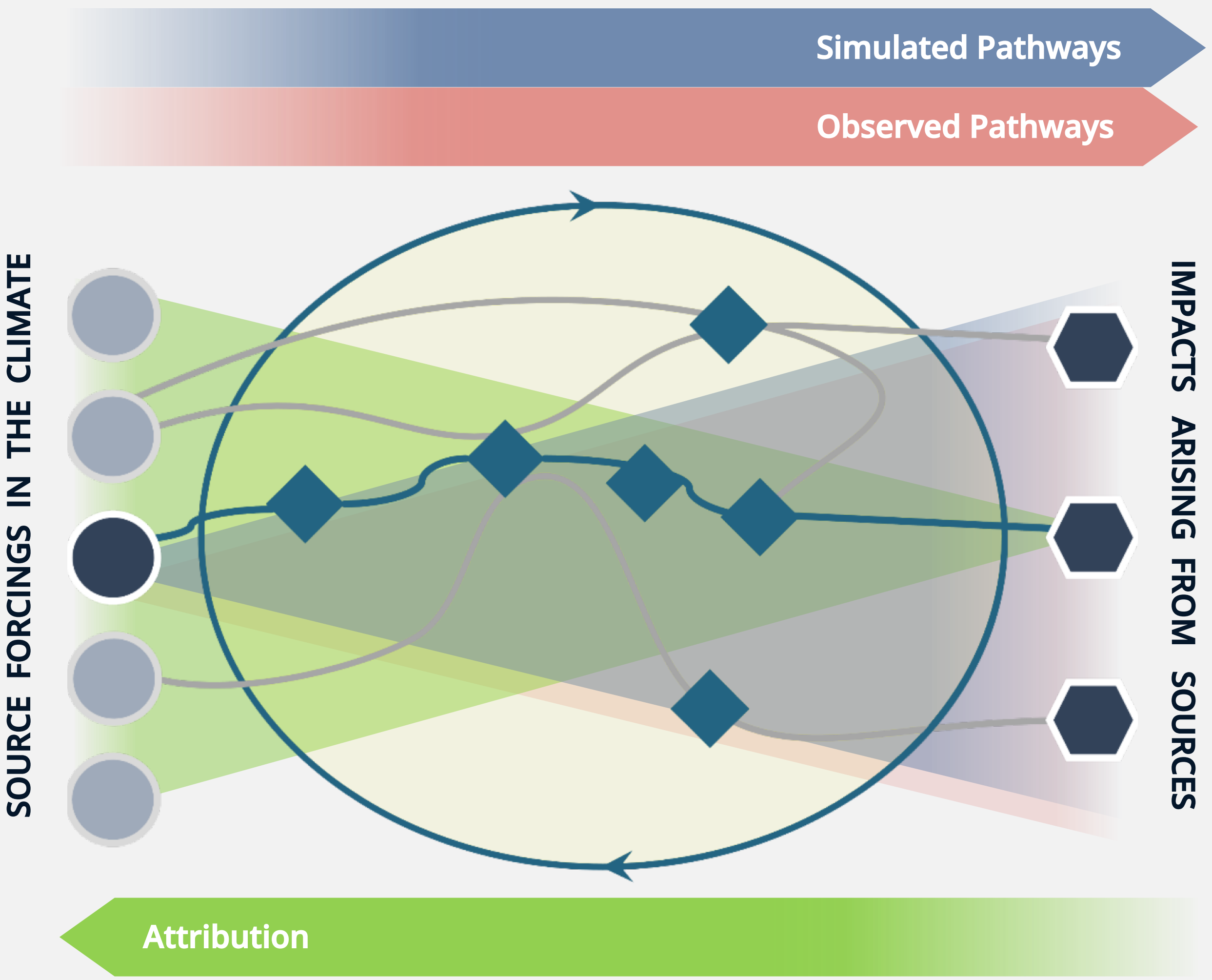


- Tropical SO2 causes Tropical H2SO4 (over a 1 and 3 month lag)
- Tropical SO2 causes a change in optical aerosol depth (AEROD)
- Tropical AEROD shows a negative contemporaneous dependence with global surface radiation flux.

J. Nichol Global Multivariate Causal Discovery for the Analysis of Emergent Properties in Earth System Models, Wed. 10:30-12:30

Team: J. Nichol, M. Smith, L. Swiler (SNL)

Summing CLDERA



OUTCOMES

Tools to discover and represent pathways, and analyses to establish pathway robustness to changing conditions.
Cross-validation using simulated and observed pathways will inform areas for model improvement and new measurements.

Contributory ranking of sources to specific impacts using pathways.
Capability enables robust risk analysis and offers the potential to guide future climate actions.

Attribution of source characteristics using inverse optimization methods.
Will provide credible methodology to deter unilateral implementation of climate interventions.

Beginning-to-end attribution in the climate system
Tracing evolving chains of physical processes to enable attribution of climate impacts from a localized source.

References

- Caldwell, P., et al., *Statistical significance of climate sensitivity predictors obtained by data mining.* Geophysical Research Letters, 2014. **41**.
- Finley, Andrew, Banerjee, Sudipto,Gelfand, Alan. (2012). Bayesian Dynamic Modeling for Large Space–Time Datasets Using Gaussian Predictive Processes. Journal of Geographical Systems. 14. 29-47.
- Golaz, J.-C., et al., The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution. Journal of Advances in Modeling Earth Systems, 2019. 11(7): p. 2089-2129.
- Goode, K., Ries, D., and Zollweg, J. "Explaining Neural Networks with Functional Data Using PCA and Feature Importance". AAAI 2020 Fall Symposium on AI in the Government and Public Sector. November 13-14, 2020. https://arxiv.org/abs/2010.12063.
- Hall, A. et al., *Progressing emergent constraints on future climate change*. Nature Climate Change, 2019. **9,** p 269-278.
- Hasselmann, K., *Multi-pattern fingerprint method for detection and attribution of climate change.* Climate Dynamics, 1997. **13**(9): p. 601-611.
- Hegerl, G.C. and G.R. North, *Comparison of Statistically Optimal Approaches to Detecting Anthropogenic Climate Change.* Journal of Climate, 1997. **10**(5): p. 1125-1133.
- Kravitz, B., et al., *First Simulations of Designing Stratospheric Sulfate Aerosol Geoengineering to Meet Multiple Simultaneous Climate Objectives.* Journal of Geophysical Research: Atmospheres, 2017. **122**(23): p. 12,616-12,634.
- Marvel, K., M. Biasutti, and C. Bonfils, *Fingerprints of external forcings on Sahel rainfall: aerosols, greenhouse gases, and model-observation discrepancies.* Environmental Research Letters, 2020: p. Medium: ED; Size: Article No. 084023.
- McDermott, Patrick & Wikle, Christopher. (2018). Deep echo state networks with uncertainty quantification for spatio-temporal forecasting. Environmetrics. 30. 10.1002/env.2553.
- National Academies of Sciences Engineering Medicine (NASEM), *Reflecting Sunlight: Recommendations for Solar Geoengineering Research and Research Governance*. 2021, Washington, DC: The National Academies Press. 328.
- Nowack, P., et al., *Causal networks for climate model evaluation and constrained projections.* Nature Communications, 2020. **11**(1): p. 1415.
- Runge, J., et al., *Inferring causation from time series in Earth system sciences.* Nature Communications 2019. **10**(1).
- Tezaur, I., et al., Global sensitivity analysis of ultra-low resolution E3SM. Submitted to JAMES, 2022.
- Wagman, B. M., L. P. Swiler, K. Chowdhary, B. Hillman. "The Fingerprints of Stratospheric Aerosol Injection in E3SM." LDRD project 224535 final report, Sept. 2021.
- Williamson, MS et al., *Emergent constraints on climate sensitivities.* Reviews of Modern Physics, 2021. **93** (2)