



Design of Experiments and Machine Learning (DoE & ML)-Based Approach to Better Capture Uncertainty in Future Climate Projections

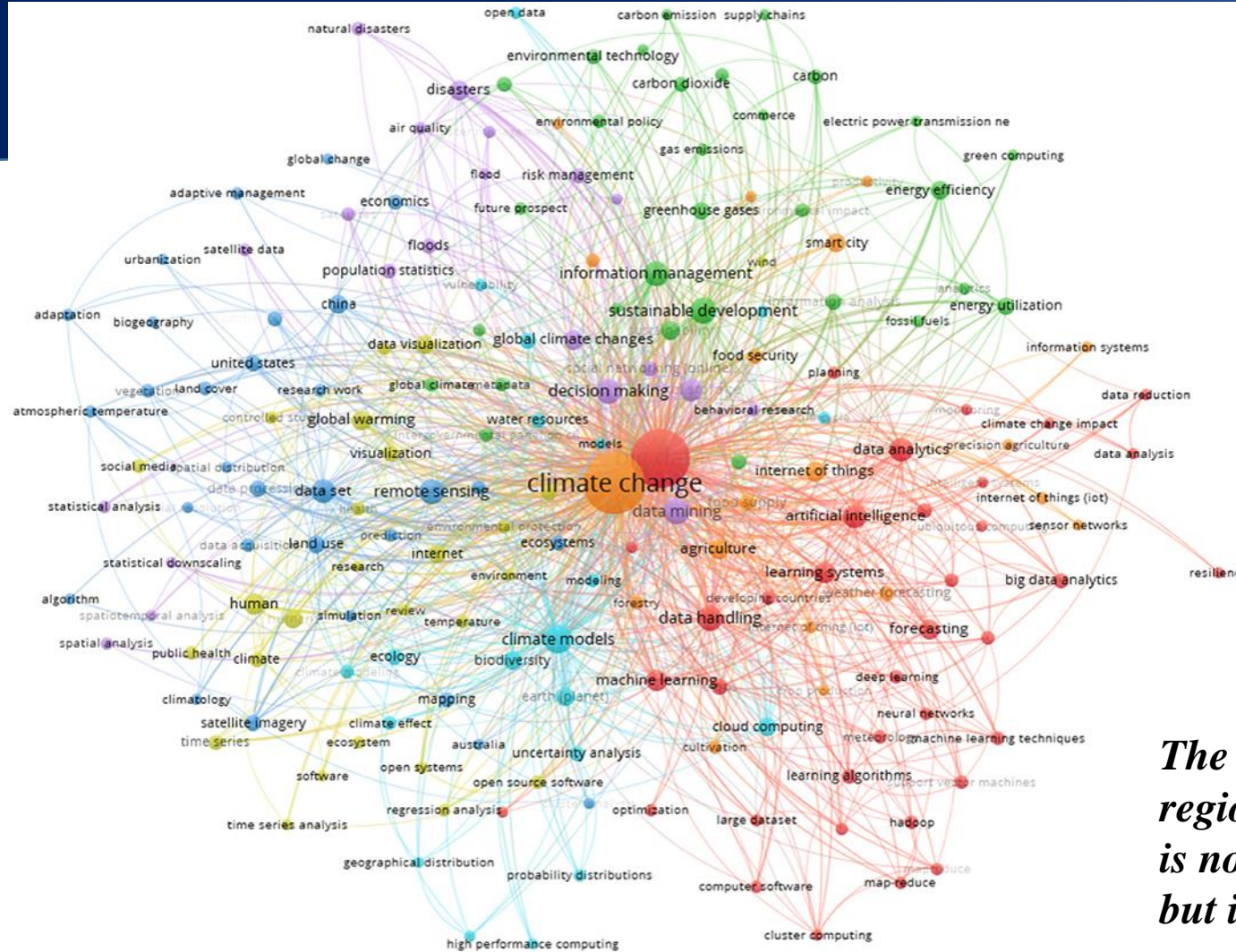
Carla A. Vivacqua

Federal University of Rio Grande do Norte, Brazil

Priscilla Mooney

Alok Samantaray

Norwegian Research Centre, Norway



The next frontier of regional climate modelling is not in producing more data, but in producing more information

Source: Sebestyen et al (2021) (<https://doi.org/10.3389/fenvs.2021.619092>)



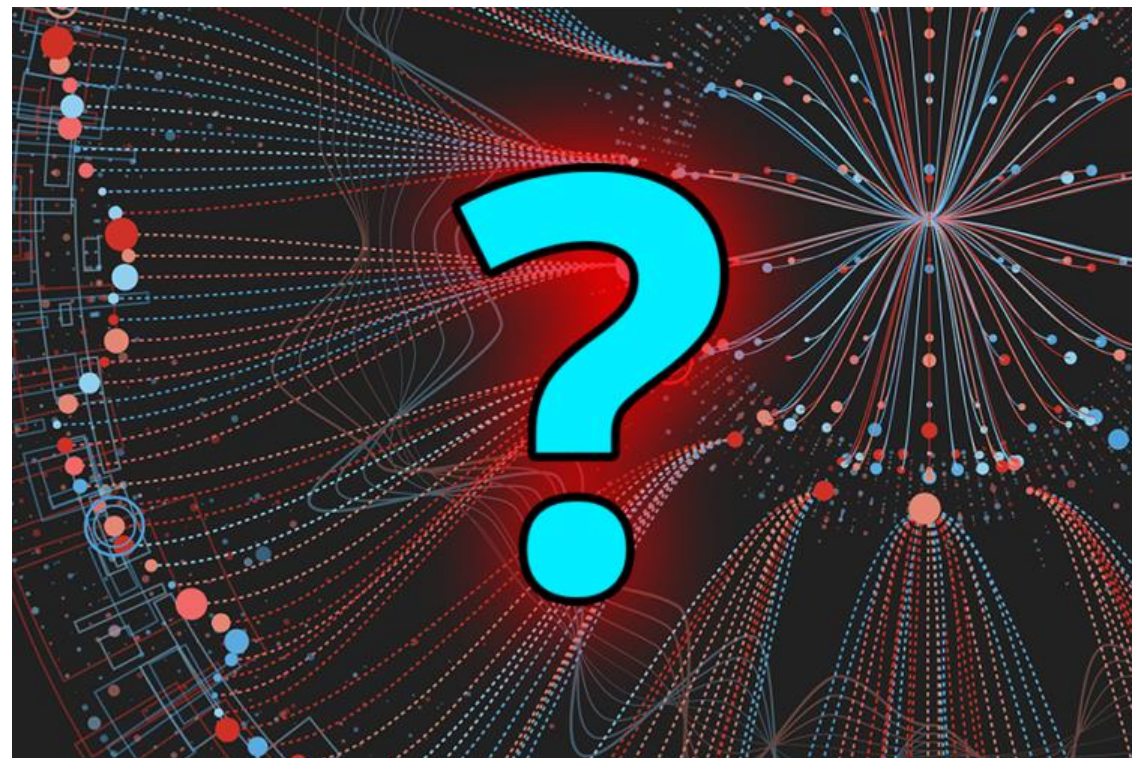
https://no.wikipedia.org/wiki/Fil:A_large_blank_world_map_with_oceans_marked_in_blue.PNG



- Statistician
- Full Professor at Department of Statistics, UFRN/Brazil
- PhD in Industrial Engineering
- Development and application of statistical methods
- Efficiency/cost reduction



- FRONTIER Project
- DoE & ML-based Approach
- Current Results



<https://cdn-wordpress-info.futurelearn.com/wp-content/uploads/8a40fc6-246d-4440-ada8-319dcfda3ca8-754x503.png>

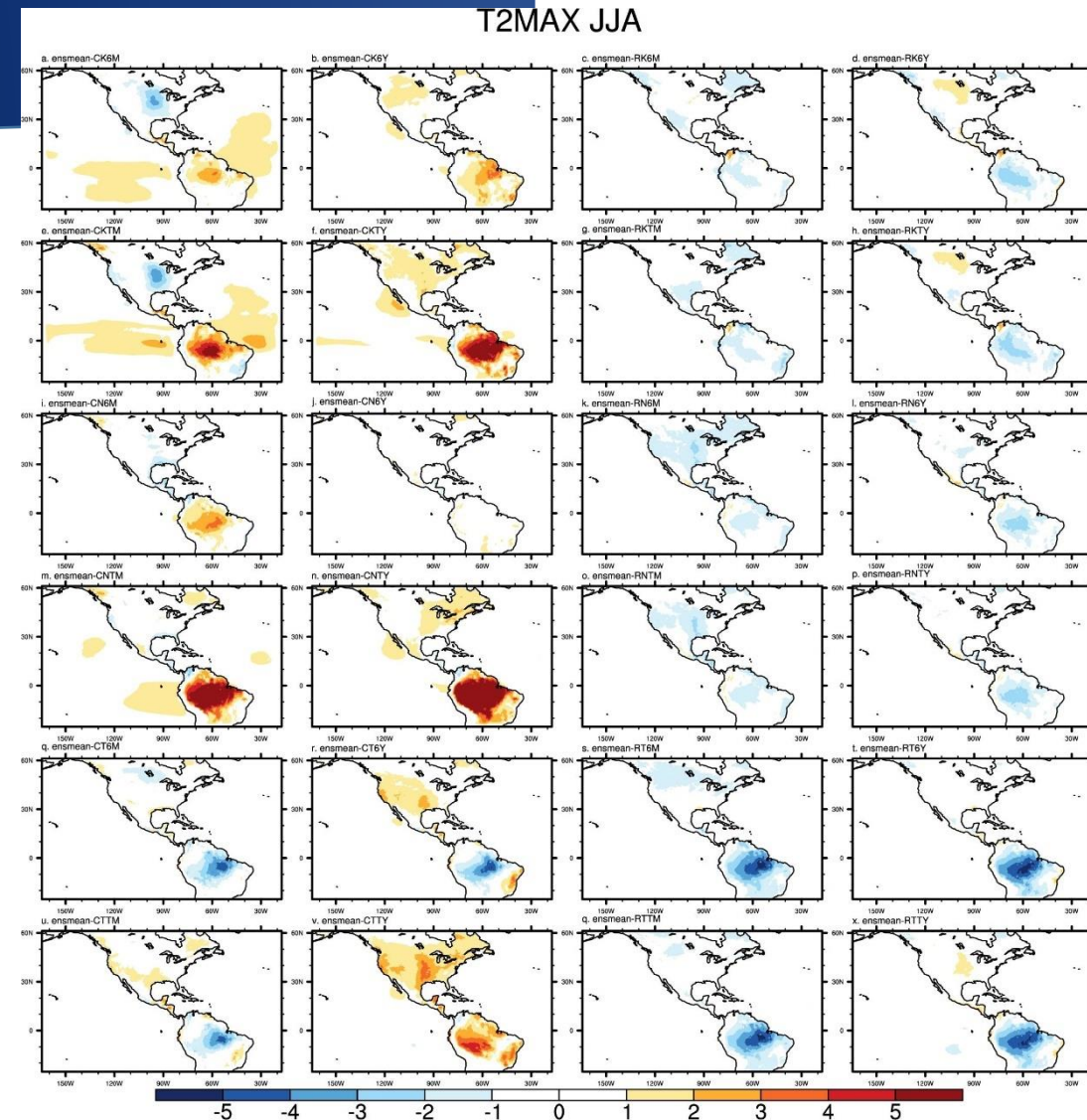
The big data and climate FRONTIER:
making sense of the explosive increase in climate data
through smart designs and big data methods

- Project Leader: Priscilla A. Mooney
 - Partners
 - Objective:
 - To **mitigate future challenges** associated with the exponential increase in climate data expected over the next decade **using efficient design processes and Big Data methods** to ensure effectiveness in *data production* and *data analysis*.

- 3 main scientific questions:
 - Does increasing model resolution add value for simulating extreme events?
 - Can the number of performance metrics be reduced?
 - Can the ‘ensemble of opportunity’ be replaced by something better?

Key tool

- **Methodology:**
 - Statistical experimental design and analysis
- **Benefits:**
 - Help identify WHICH characteristics impact model performance and WHY
 - Currently used approaches focus on WHAT models perform better and a deeper understanding of WHICH characteristics and WHY may require a great effort
 - Allow estimation of the performance of missing ensemble members in simulations

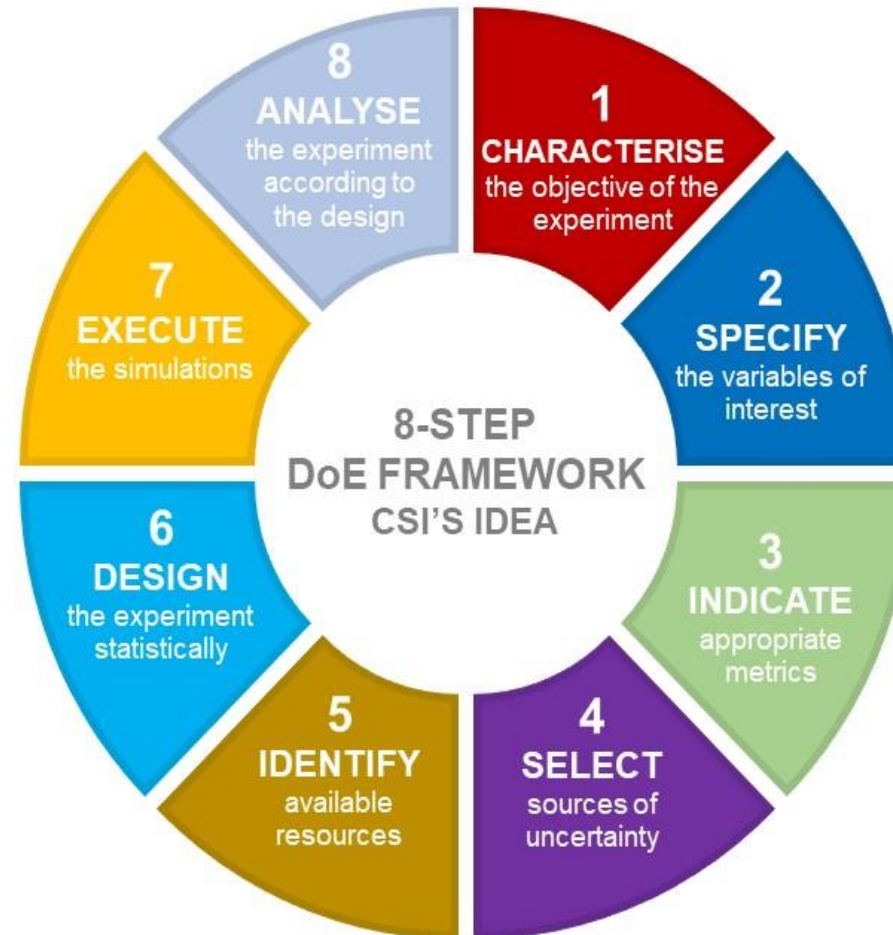


Bruyere (2014)

DoE Framework: Climate Statistics Investigation'S IDEA CSI'S IDEA

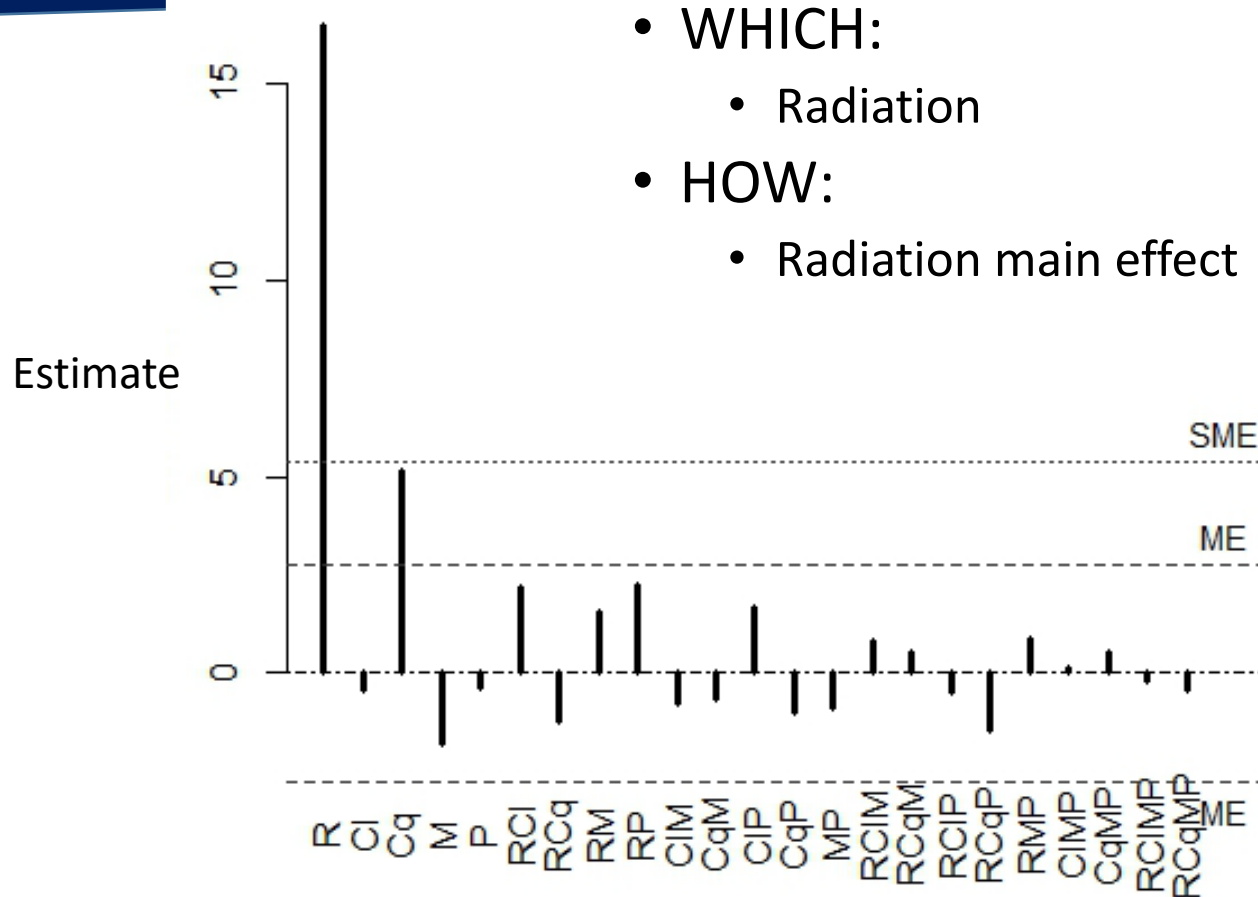
8-step framework based on
Design of Experiments (DoE)

Steps 6 and 8 are the
backbones of the framework

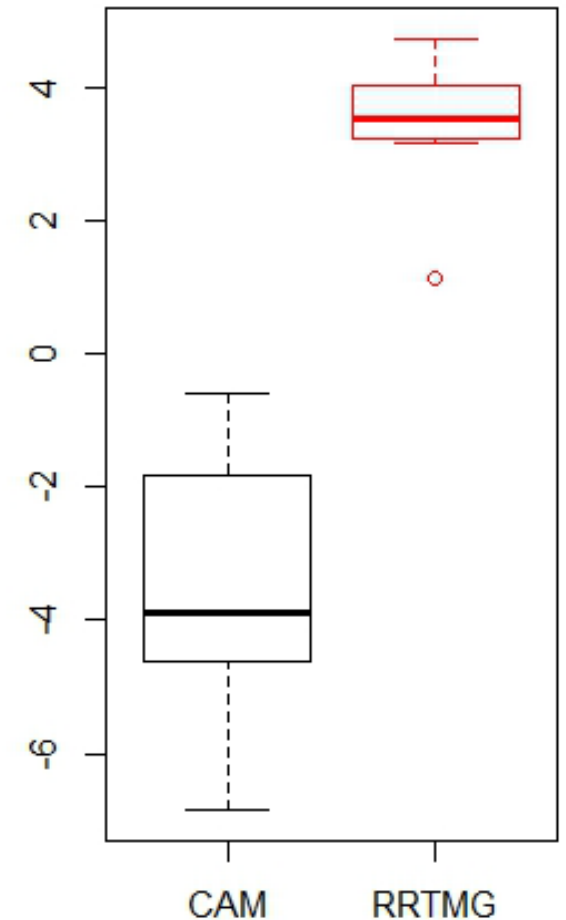


Expected result:
Capture uncertainty
with few well-chosen
ensemble
members

Statistical Data Analysis



Principal Component



- WHAT:
 - RRMTG
- WHY:
 - Improves overall all metrics

Reduction of experimental effort

Objective, Region	Total of Configurations	Total of Simulations	Reduction (%)
<i>Precipitation, RN-Brazil</i>	<i>64</i>	<i>32</i>	<i>50,0</i>
<i>Cyclone tracks, India</i>	<i>512</i>	<i>64</i>	<i>87,5</i>
<i>Several variables, Norway</i>	<i>243</i>	<i>11</i>	<i>95,5</i>
<i>Precipitation, India</i>	<i>2.268</i>	<i>32</i>	<i>98,6</i>

Conclusions

- **Ongoing quest for adequate climate and weather modeling through ensemble experiments needs to be embedded in a statistical design and analysis framework that rigorously and efficiently compares different models against observed or reanalysis data.**
 - A large range of configurations can be evaluated in a controlled way.
 - This enables to scrutinize the role of different sources of uncertainty, guiding the choice of model configuration tailored to one's specific needs.
- **When appropriately applied, statistically designed and analysed experiments can considerably contribute to reducing computational effort and to advancing the knowledge regarding weather and climate modeling across different regions.**



Contact information

carla.vivacqua@ufrn.br

